

## ТЕОРИЈА УЗОРАКА - ЗАДАЦИ СА ВЕЖБИ

1. Показати да важи  $MSE(\hat{\theta}) = D(\hat{\theta}) + b^2(\hat{\theta})$ .
2. Нека су  $X_1, X_2, \dots, X_n$  независне случајне величине из исте расподеле са дисперзијом  $\sigma^2$  и очекивањем  $\mu$ .
  - а) Одредити пристрасност оцене (за  $\sigma^2$ )  $S_n^2 = \sum_{k=1}^n (X_k - \bar{X}_n)^2$ .
  - б) Одредити непристрасну оцену за  $\sigma^2$ .
3. Марко има 7 јабука, Петар 2, Јован 2 и Саша 6. Да ли је оцена укупног броја јабука непристрасна, ако се узорак бира тако да је вероватноћа да су Марко и Петар у узорку  $\frac{1}{3}$ , вероватноћа да су Марко и Јован у узорку  $\frac{1}{2}$ , вероватноћа да су Петар и Јован  $\frac{1}{6}$ , док је вероватноћа избора свих осталих узорака једнака 0. (Као оцена укупног броја јабука користи се статистика  $\hat{t} = N\bar{X}_n$ ).
4. Испитати да ли је боља оцена укупне суме обележја ( $\hat{t} = N\bar{X}_n$ ) на основу узорка обима 2, или на основу узорка обима 3, из популације  $\{1, 2, 3, 4\}$  ако се користи прост случајан узорак без понављања.
5. Колико има простих случајних узорака без понављања, а колико са понављањем, ако се вади узорак обима 7 из популације која садржи 20 јединици?
6. Из скупа  $\{1, 2, 3, 4, 5, 6, 7, 8\}$  ваде се узорци обима 2, и то тако да сваки који садржи непаран број има вероватноћу 0. Они узорци који садрже 2 имају вероватноћу обрнуто пропорционалну другом елементу узорка, а они који садрже 4 имају вероватноћу  $\frac{1}{8}$ . Испитати непристрасност оцене средње вредности и одредити средње квадратну грешку.
7. Из популације  $\{1, 2, \dots, 100\}$  извадити 15 простих случајних узорака без понављања. На сваком од њих наћи оцену укупне суме обележја и испитати који од узорака је најрепрезентативнији, тј. где је реализована вредност статистике  $\hat{t} = N\bar{X}_n$  најближа стварној вредности.
8. Нека је  $Z = \sum_{i=1}^N a_i I_i$ , где су  $a_1, a_2, \dots, a_N$  константе,
  - а  $I_i$  индикатори укључења  $i$ -те јединке у узорак,  $i \in \{1, 2, \dots, N\}$ .
  - Одредити очекивање и дисперзију случајне величине  $Z$ .
9. Дата је популација  $\{1, 2, 3, 4, 5, 6, 7, 8\}$  и размотрен је следећи план узорковања:

$S$	$\{1, 3, 5, 6\}$	$\{2, 3, 7, 8\}$	$\{1, 4, 6, 8\}$	$\{2, 4, 6, 8\}$	$\{4, 5, 7, 8\}$
$P(S)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

- а) Наћи вероватноће укључења  $\pi_i$  за сваки елемент  $i$ .
- б) Одредити расподелу оцене  $\hat{t} = 8\bar{X}_n$ .

### 1 ПРОСТ СЛУЧАЈАН УЗОРАК

1. У кафићу постоје четири врсте колача - баклаве, тулумбе, екери и шампите. Десеторо људи је купило баклаву, двадесеторо тулумбу, двадесет и петоро еклер, а петнаесторо шампиту. На случајан начин се одаберу 2 врсте колача и бележи се број људи који су купили ту врсту.
  - а) Оценити укупан број особа које су купиле колач, и одредити непристрасност те оцене.
  - б) Испитати да ли су узорачка дисперзија и узорачка стандардна грешка непристрасне оцене дисперзије и стандардне грешке на читавој популацији.
2. Узет је прост случајан узорак од 10 ученика од 100 ученика трећег разреда и бележене су њихове оцене из математике. Забележени су резултати: (4, 5, 5, 2, 3, 1, 3, 4, 4, 5). Оценити просечну оцену из математике, а затим израчунати оцену дисперзије те оцене.
3. Показати да је код простог случајног узорка без понављања  $s_n^2$  непристрасна оцена за  $s^2$ .
4. Одредити расподелу случајне величине  $\mu_i$  која представља број појављивања  $i$ -те јединке популације, ( $i \in \{1, 2, \dots, N\}$ ), у простом случајном узорку са понављањем.

5. У кутији се налази 100 папирића на којима се налазе каро, пик, херц и треф. Извлачи се 10 папирића, тако што се папирић извуче, забележи се знак, а затим се врати у кутију. Добијен је узорак (К,П,П,Х,К,Т,Т,П,Х,Х).
- Ако каро има вредност 1 динар, пик 2, херц 3, а трефа четири, оценити укупну суму која се налази у кутији.
  - Одредити оцену броја пикова.
  - Одредити непристрасне оцене дисперзија ових оцена.
6. Показати да је код простог случајног узорка са понављањем  $\frac{N}{N-1}s_n^2$  непристрасна оцена за  $s^2$ .
7. Одредити обим простог случајног узорка (без понављања) потребан за оцењивање пропорције неког својства обележја популације.
8. Истраживач жели да оцени пропорцију деце са плавим очима у вртићу од 4000 деце. Прихватљиво одступање му је 5% са ризиком 8%. Одредити обим узорка потребан истраживачу. Ако је истраживач већ радио истраживање у другом вртићу и добио пропорцију  $\frac{1}{3}$ , који му је онда обим узорка потребан?
9. Узет је прост случајан узорак од 10 кућа из популације од 100 кућа. Број становника у кућама из узорка је 2,5,1,4,4,3,2,5,2,3.
- Оценити укупан број становника у популацији и оценити дисперзију те оцене.
  - Одредити просечан број становника по кући и оценити дисперзију те оцене.
  - Наћи приближни 90%-ни интервал поверења за укупан број становника, као и за просечан број становника по кући.
10. Одређује се број дивљих животиња у неком региону који је подељен на 286 области. Изабран је прост случајан узорак од 15 области и дат је број животиња у њима:
- $$1, 50, 21, 98, 2, 36, 4, 29, 7, 15, 86, 10, 21, 5, 4$$
- Наћи узорачку дисперзију, оцену средине, оцену дисперзије те оцене, као и оцену стандардне грешке.
  - Наћи 90%-ни интервал поверења за укупан број животиња.
  - Коју величину узорка треба узети за оцењивање укупног броја животиња тако да одступање не буде веће од 2000 животиња са вероватноћом 90%.
11. Ботаничар жели да оцени број стабала брезе у некој области. Област је подељена на 1000 делова. Познато је из претходних испитивања да је дисперзија броја стабала по области приближно  $s^2 \approx 45$ . Одредити величину простог случајног узорка, потребну да са вероватноћом 95% одступање не буде веће од 500 стабала.
12. Посматрамо популацију обима  $N = 5$ , чији су елементи означени бројевима 1, 2, 3, 4 и 5, и вредности обележја су редом 3, 1, 0, 1 и 5. Размотримо принцип простог случајног узорковања за узорак обима  $n = 3$ .
- Одредити све могуће узорке обима 3 и вероватноће да сваки од тих узорака буде изабран.
  - За сваки узорак, израчунати узорачку средњу вредност, узорачку дисперзију, оцену дисперзије узорачке средине и медијану.
  - Показати да је средина узорка непристрасна оцена средине обележја популације, и да је узорачка дисперзија непристрасна оцена дисперзије коначне популације.
13. У фајлу *deca.txt* дати су подаци о броју деце у свакој од 51 улица у неком месту. Наћи 95%-ни интервал поверења за укупан број деце користећи прост случајан узорак обима 7.
14. За које од следећих планова простог случајног узорковања ће бити дата најпрецизнија оцена средње вредности популације? Претпоставимо да свака популација има дисперзију  $s^2$ .
- Прост случајан узорак обима 400, добијен из популације обима 4000.
  - Прост случајан узорак обима 30, добијен из популације обима 300.
  - Прост случајан узорак обима 3000, добијен из популације обима 300000000.
15. Оцењује се пропорција елемената популације који задовољавају одређену особину. Показати да важи  $\frac{D(\hat{p})}{E(\hat{p})} = \frac{N-n}{n(N-1)}p(1-p)$ , где је  $\hat{p}$  удео елемената из узорка који задовољавају дату особину.

## 2 УЗОРКОВАЊЕ СА НЕЈЕДНАКИМ ВЕРОВАТНОЋАМА

1. Наћи *Hansen – Hurwitz*-ова оцену укупне суме обележја популације ако се користи прост случајан узорак са понављањем, а затим и дисперзију и *Hansen – Hurwitz*-ову оцену дисперзије те оцене.
2. Претпоставимо да имамо узорак са понављањем и да је  $p_i$ , вероватноћа избора  $i$ -те јединке у једном покушају, позната. Наћи  $\pi_i$ , вероватноћу укључења  $i$ -те јединке у узорак и  $\pi_{ij}$  - заједничку вероватноћу укључења јединки  $i$  и  $j$  ( $i, j \in \{1, 2, \dots, N\}$ ).
3. Популација се састоји од 15 јединки чије су величине 23, 30, 41, 26, 53, 60, 28, 52, 113, 72, 80, 35, 42, 38 и 52. Изабрати узорак са вероватноћама пропорционалним величини обима 3 користећи Лахиријев метод.
4. Изабран је узорак са вероватноћама пропорционалним величини, са понављањем, обима 3 из популације величине 10. Дате су вредности обележја изабраних елемената и вероватноће избора:

$i$	1	2	3
$x_i$	3	10	7
$p_i$	0.06	0.2	0.1

- а) Наћи оцену укупне суме обележја популације користећи *Hansen – Hurwitz*-ову оцену.
  - б) Оценити дисперзију те оцене.
  - в) Наћи оцену укупне суме обележја популације користећи *Horvitz – Thompson*-ову оцену.
  - д) Оценити дисперзију те оцене.
5. Ако се бира узорак обима  $n$  без понављања, показати да важи:
    - а)  $\sum_{i=1}^N \pi_i = n$ .
    - б)  $\sum_{j=1, j \neq i}^N \pi_{ij} = (n-1)\pi_i$ , за свако фиксирано  $i$ ,  $i = \overline{1, n}$ .
  6. За испитивање загађености 320 језера на површини  $80km^2$  изабран је узорак од 4 језера. Прво језеро из узорка бирано је два пута, а остала два по једном. Концентрација загађености у та три језера у узорку је редом 2, 5 и 10, а величине тих језера (у  $km^2$ ) су редом 1.2, 0.2 и 0.5. Наћи *Hansen – Hurwitz*-ову оцену средњег загађења по језеру у посматраној популацији, као и оцену дисперзије добијене оцене.
  7. Из популације коју чине три поља на којима се узгаја пшеница бира се узорак обима 2 са вероватноћама пропорционалним величинама, са понављањем. У следећој табели су дати подаци о количини произведене пшенице на сваком пољу и вероватноће избора сваког поља.

$i$	1	2	3
$x_i$	11	6	25
$p_i$	0.3	0.2	0.5

Налажењем свих могућих узорака одредити вероватноће укључења сваког елемента  $\pi_i$ , ( $i = 1, 2, 3$ ), *Hansen – Hurwitz*-ову и *Horvitz – Thompson*-ову оцену за укупну производњу пшенице.

8. У фајлу *radnici.txt* дати су подаци о броју радника у производњи у 10 фабрика у индустријској зони. Из ове популације избрати узорак обима 3 са понављањем са вероватноћама избора пропорционалним броју радника у фабрици. Користећи добијени узорак одредити *Hansen – Hurwitz*-ову оцену укупне производње.

9. Дата је популација од четири прашуме датих површина, и бројеви тигрова који живе у њима. Вади се узорак од 2 прашуме. Оценити укупан број тигрова, одредити дисперзију те оцене и наћи непристрасну оцену те дисперзије ако је узорковање вршено:

- Са понављањем (Користити *Hansen – Hurwitz*-ову оцену),
- Без понављања (Користити *Horvitz – Thompson*-ову оцену),

и извучен је узорак (1,2).

$i$	1	2	3	4
$P(km^2)$	100	200	300	500
$p_i$	11	20	23	54

10. Популацију чини база *trees* која садржи податке о 31 дрвету.

- Изабрати прост случајан узорак обима 10 помоћу функције *sample()* и оценити средњу вредност обележја *Volume* користећи добијени узорак. Затим поновити поступак 1000 пута и за тако добијене вредности одредити сумарне статистике и нацртати хистограм. Упоредити оцену средње вредности добијену на основу једног узорка и ону добијену симулацијом.
- Изабрати узорак са понављањем са вероватноћама пропорционалним обиму стабла (променљива *Girth*) и оценити средњу вредност обележја *Volume* користећи *Hansen – Hurwitz*-ову оцену.

### 3 СТРАТИФИКОВАН УЗОРАК

1. У популацији од 6 елемената, вредности 0, 1 и 2 су у првом стратуму, а 4, 6 и 11 у другом стратуму. У узорак су изабрани елементи 0 и 2 из првог стратума, а 6 и 11 из другог.

- Одредити оцену средње вредности обележја популације, дисперзију те оцене и оцену те дисперзије.
- Упоредити добијену дисперзију са дисперзијом оцене средине добијене на основу простог случајног узорка без понављања.
- Оценити укупну суму обележја популације, дисперзију те оцене и оцену те дисперзије.

2. Испитује се број претрчаних метара на физичком ученика једног разреда. У разреду има 112 ученика, од чега 59 девојчица и 53 дечака. Посматрано је 7 дечака и 8 девојчица.  
Дечаки: 879, 810, 789, 567, 900, 870, 777.  
Девојчице: 450, 234, 679, 456, 239, 555, 560, 467.

- Оценити укупан број метара које су претрчала деца тог разреда и оценити дисперзију те оцене. Упоредити са оценом добијеном из простог случајног узорка без понављања.
- Оценити просечан број претрчаних метара и оценити дисперзију те оцене.

3. Популација од 10 елемената подељена је на два стратума са елементима 3, 5, 8, 23, 2 и 34, 11, 2, 1, 12. Вади се стратификован узорак обима 4, тако што се из сваког стратума вади прост случајан узорак са понављањем обима 2. У узорак су изабране јединке 3 и 5 из првог стратума, а 2 и 12 из другог.

- Оценити укупну суму обележја популације, израчунати дисперзију те оцене и оцену те дисперзије
- Оценити средњу вредност обележја популације, израчунати дисперзију те оцене и оцену те дисперзије.

4. а) Популација је подељена на три стратума чије су величине  $N_1 = 123$ ,  $N_2 = 102$  и  $N_3 = 180$  и одговарајуће дисперзије су  $s_1^2 = 116$ ,  $s_2^2 = 143$  и  $s_3^2 = 170$ . Ако се бира стратификован узорак обима 10, одредити величину узорка који се вади из сваког стратума користећи пропорционални избор.

- Извести формулу за одређивање обима узорака који се узимају из сваког стратума ако се користи Нојманова метода, а затим одредити величину узорка који се узима из сваког стратума за популацију и дати обим узорка из дела а).

5. Нека је популација од 200 елемената подељена у три стратума са  $N_1 = 60, N_2 = 80$  и  $N_3 = 60$  и  $s_1 = 2s_2 = 4s_3$ . Ако се бира стратификован узорак обима 30, одредити величину узорка који се вади из сваког стратума користећи:

- а) пропорционални избор;
- б) Нојманов избор.

6. За испитивање просечне недељне потрошње бензина град је подељен на 4 дела, који се посматрају као стратуми. Забележена је потрошња бензина за протеклу недељу код сваког возача из узорка и добијени су следећи подаци:

Стратум 1:  $N_1 = 3750, n_1 = 50, \bar{x}_{n1} = 12.6, s_{n1}^2 = 2.8$   
 Стратум 2:  $N_2 = 3272, n_2 = 45, \bar{x}_{n2} = 14.5, s_{n2}^2 = 2.9$   
 Стратум 3:  $N_3 = 1387, n_3 = 30, \bar{x}_{n3} = 18.6, s_{n3}^2 = 4.8$   
 Стратум 4:  $N_4 = 2475, n_4 = 30, \bar{x}_{n4} = 13.8, s_{n4}^2 = 3.2$ .

- а) Оценити средњу недељну потрошњу за цео град.
- б) Наћи 95%-ни интервал поверења добијене оцене.
- б) Ако треба изабрати стратификовани случајни узорак обима 1000, одредити величину узорка који се вади из сваког стратума по пропорционалном распореду.

7. Средњошколци су подељени у три групе по успеху у школи са циљем да се испита колико су заинтересовани за позориште. У фајлу *posete.txt* дати су подаци о броју посета позоришту за ученике, као и којој групи по успеху припадају. Изабрати стратификован узорак без понављања обима 20, користећи пропорционални избор, а затим оценити просечан број посета позоришту за средњошколце, дисперзију те оцене и оцену те дисперзије. Затим урадити то исто користећи Нојманов избор и упоредити добијене дисперзије.

#### 4 КОЛИЧНИЧКО ОЦЕЊИВАЊЕ

1. Узет је прост случајан узорак без понављања обима 30 из велике популације и добијени су подаци  $\bar{y}_n = 5$  и  $\bar{x}_n = 18$ , где је  $x$  обележје од интереса, а  $y$  помоћно обележје. Такође, познато је да је  $\bar{y} = 6$ . Наћи оцену количника  $\hat{R}$  и количничку оцену средње вредности обележја  $x$ .

2. Агент за некретнине жели да процени релативну промену у просечној цени кућа за период од две године. Посматрао је 20 кућа од укупно 1000 за које је задужен. Дате су цене кућа у хиљадама евра ове године и одговарајуће цене пре две године за сваку од 20 посматраних кућа:

Пре две године: 290, 53, 300, 60, 110, 311, 150, 150, 240, 190, 32, 180, 90, 351, 241, 300, 230, 84, 140, 155;

Сада: 255, 50, 275, 66, 109, 321, 113, 126, 215, 177, 24, 179, 65, 339, 314, 271, 235, 74, 135, 154.

Оценити релативну промену у просечној цени.

3. Дата је популација обима 4 и вредности обележја које се испитује ( $x$ ) и помоћног обележја ( $y$ ):

$x$ : 3, 5, 7, 9

$y$ : 1, 2, 2, 3

На основу простих случајних узорака обима 2 и 3 испитати да ли је количничка оцена количника популације непристрасна.

4. Дате су следеће вредности:

Школа	1	2	3	4
Број деце која имају 5 из математике	200	300	100	250
Број посматраних разреда	4	5	3	4

Истраживач жели посматрањем прве и треће школе да донесе закључак о укупном броју деце која имају пет из математике. Одредити количничку оцену, испитати да ли је непристрасна, а затим упоредити са оценом добијеном на основу простог случајног узорка без понављања. Која је боља?

5. Дати су подаци за прост случајан узорак без понављања обима 10 из популације обима 53 у коме је  $x$  обележје које се испитује, а  $y$  помоћно обележје:

$x_i$	8.75	2.55	9.00	1.10	7.50	5.00	1.15	3.40	2.00	1.25
$y_i$	8.35	1.50	10.00	0.60	7.50	7.95	0.95	4.40	1.00	0.50

Познато је да је популацијска средина обележја  $y$  једнака  $\bar{y} = 3.33$ .

- а) Наћи количничку оцену средње вредности за обележје  $x$ .
- б) Оценити дисперзију добијене количничке оцене.
- в) Наћи 95%-ни интервал поверења за средину обележја популације.

6. У граду који има 28753 кућа, циљ је да се испита просечан рачун за струју. Посматране су четири куће и забележен је број чланова, као и рачуни за струју за сваку од кућа.

Прва кућа: 1 члан, рачун од 1281 динар;

Друга кућа: 3 члана, рачун од 4375 динара;

Трећа кућа: 2 члана, рачун од 2333 динара;

Четврта кућа: 4 члана, рачун од 5789 динара.

Укупан број становника у граду је 70351.

Наћи количничку оцену просечног рачуна за струју, као и оцену дисперзије те оцене. Колико је укупно новца потрошено на струју?

7. У популацији обима 15 познате су вредности помоћног обележја  $y$ : 2.3, 3.0, 4.1, 2.6, 5.3, 6.0, 2.8, 5.2, 11.3, 7.2, 8.0, 3.5, 4.2, 3.8, 5.2. Из популације је одабран прост случајан узорак обима 5 и добијени су следећи резултати:

$x_i$	40	40	70	82	52
$y_i$	4,1	5,2	7,2	8,0	5,2

Наћи количничку оцену средње вредности обележја  $x$  и оцену дисперзије те оцене.

8. Дати су подаци за два стратума исте величине из популације величине 12.

$h$	$x_{h_i}$	$y_{h_i}$
1	2,2,3,4,6,7	6,6,7,7,8,8
2	10,11,12,14,15,16	10,11,16,17,18,18

Из сваког стратума бирају се у узорак по 3 јединке. Вредности јединки укључених у узорак дате су у табели:

$h$	$x_{h_i}$	$y_{h_i}$
1	2,3,7	6,7,8
2	10,14,15	10,17,18

Наћи комбиновану и посебну количничку оцену средине обележја популације.

## 5 РЕГРЕСИОНО ОЦЕЊИВАЊЕ

1. Из популације која садржи 1200 јединки, одабран је прост случајан узорак обима 20 и забележене су следеће информације о обележју које се испитује ( $x$ ) и помоћном обележју ( $y$ ):

$$\bar{x}_n = 20,$$

$$\bar{y}_n = 15,$$

$$\sum_{i \in S} (x_i - \bar{x}_n - b_0(y_i - \bar{y}_n))^2 = 100.$$

Такође, из претходних испитивања познато је да је  $b_0 = 2$  и  $\bar{y} = 12$ .

Наћи регресиону оцену средње вредности обележја  $x$  на популацији, као и оцену дисперзије те оцене.

2. У квалификацијама за европско првенство учествовало је 40 екипа. Одабран је узорак од 6 екипа, и дати су подаци о броју постигнутих кошева на 10 утакмица, као и број кошаркаша у тиму виших од 2 метра.

Број кошева: 451, 345, 444, 378, 399, 421.

Виши од 2m: 6, 4, 4, 6, 5, 7.

Познато је да је укупно на турниру 211 играча виших од 2m. Регресионом методом оценити укупан број постигнутих кошева на десет утакмица, ако је регресиони коефицијент:

- а) Непознат, и оцењује се из узорка;
- б) Познат,  $b_0 = 20$ . У овом случају одредити оцену дисперзије оцене укупног броја кошева.

3. Истраживач жели да оцени укупан број радно способног становништва у 100 градова. У фајлу *stanovnici.txt* налазе се информације о броју радног способног становништва за 25 градова. Прва колона представља податке прикупљене пре пет година, а друга податке који се односе на ову годину. Познато је да је пре пет година било укупно 168740 радно способних становника у ових 100 градова.

Наћи регресиону оцену укупног броја радно способних становника.

4. Тест из математике на пријемном у једној школи полагао је 486 ученика, који су затим полагали тест из српског. Простим случајним узорковањем одабрано је 10 студената и забележен је њихов број поена на тесту из српског. Познато је да просечан број поена на тесту из математике за свих 486 ученика једнак 52. Оценити просечан број поена из српског, регресионом оценом, ако су добијени следећи подаци:

Српски: 65, 78, 52, 82, 92, 89, 73, 98, 56, 75.

Математика: 39, 43, 21, 64, 57, 47, 28, 75, 34, 52.

5. Дата је популација од 12 јединки која је подељена на три стратума. Вредности обележја од интереса и помоћног обележја су дате у следећој табели:

	$x$	$y$
<i>I</i>	4,3,2,1,6	5,5,1,2,7
<i>II</i>	2,3	4,5
<i>III</i>	5,3,1,7,5	6,3,3,2,8

У узорак су одабрани елементи:

	$x$	$y$
<i>I</i>	4,1,6	5,2,7
<i>II</i>	2	4
<i>III</i>	5,3	6,3

а) Ако је познато да је  $b_0 = 2$ , наћи комбиновану оцену средње вредности обележја  $x$ , а затим одредити дисперзију те оцене и оцену те дисперзије.

б) Ако је познато да је  $b_{10} = 2, b_{20} = 1, b_{30} = 3$ , наћи посебну оцену средње вредности обележја  $x$ , а затим одредити дисперзију те оцене и оцену те дисперзије.

6. У фирми ради 500 људи који су подељени на четири категорије према степену образовања. Циљ је да се оцени колико просечно новца запослени месечно издвајају за забаву. У првој категорији је 57 радника, другој 228, трећој 153 и четвртој 82 радника. Посматра се 30 људи и бележе се њихове зараде, као и колико новца су потрошили на забаву.

а) Користећи Нојманов избор, проценити колико би људи требало одабрати из сваке категорије да би се донели исправни закључци о просечној суми новца потрошеној за забаву.

б) Користећи резултате из дела а), извадити стратификован узорак од 30 радника, а затим комбинованом и посебном методом оценити просечну суму новца коју запослени месечно потроши за забаву.

## 6 КЛАСТЕР УЗОРАК

1. Популација од 12 елемената подељена је на три скупине и подаци су дати у следећој табели:

	$x_{ij}$
1	2,4,6
2	3,3,4,5,5
3	1,1,2,4

Изабран је прост случајан узорак од прве и треће скупине. Наћи оцену укупне суме обележја популације, дисперзију те оцене, а затим и количничку оцену укупне суме обележја популације, и дисперзију те оцене.

2. Популација која садржи 100 елемената, издељена је на 10 скупина. Изабран је прост случајан узорак од 3 скупине чије су величине и суме обележја:

$i$	1	2	3
$M_i$	5	20	10
$t_i$	4	12	7

- Наћи непристрасну оцену укупне суме обележја популације и оцену популацијске средине.
- Оценити дисперзију добијене оцене суме.
- Наћи количничку оцену укупне суме обележја популације и оцену популацијске средине.
- Оценити дисперзију добијене оцене суме.

3. Истраживачи желе да процене укупан број оболелих од неке ретке болести у Србији. Посматрани су Борски, Рашки и Моравички округ и забележни су следећи резултати:

Округ	Борски	Рашки	Моравички
Број становника	178718	300234	230748
Број оболелих	1001	2230	1798

Укупно има 29 округа и 7186862 становника. Оценити укупан број оболелих користећи количничку оцену, а затим оценити дисперзију те оцене.

4. Популација која садржи 100 елемената, издељена је на 20 скупина. Изабран је узорак од 5 скупина са вероватноћама пропорционалним величинама скупина (са понављањем), чије су величине и суме обележја дате у следећој табели:

$i$	1	2	3	4	5
$M_i$	5	17	10	12	22
$t_i$	4	11	7	11	23

Наћи *Horvitz – Thompson*-ову оцену средње вредности обележја и оценити дисперзију те оцене.

5. Дати су подаци о оценама из математике за 12 ученика који су подељени у 3 групе:

	Оцене
$I$	2,4,5
$II$	3,3,4,5,2
$III$	1,3,5,1

Бирањем пропорционалним величини групе, у узорак су одабране прва и трећа група. Наћи *Hansen – Hurwitz*-ову оцену просечне оцене из математике и дисперзију те оцене.

## 7 СИСТЕМАТСКИ УЗОРАК

- Популација се састоји од 12 елемената, а обележје је редни број јединке. Формирати систематске узорке обима 4, а затим одредити дисперзију средине систематског узорка.
- Популација се састоји од 900 јединки. Одабрати узорак обима 120 користећи систематски узорак.