

СТРАТИФИЦИРАН СЛУЧАЕН ПРИМЕРОК

*Сџево Ѓорѓиев*¹

Множеството чишто елементи задоволуваат одредени карактеристики кои се предмет на истражување се нарекува *популација*. Елементи на популацијата можат да бидат индивидуи, домаќинства, претпријатија, маркети, училишта и слично. На пример, кога би сакале да го истражине придонесот на претпријатијата кои се занимаваат со трговија на мало во економијата на една држава, популацијата ќе ја сочинуваат сите претпријатија кои се занимаваат со трговија на мало. Спроведувањето на ова истражување со целосен опфат, односно со испитување на сите елементи од популацијата, бара ангажирање на голем број на ресурси и средства, што не е секогаш возможно. Заради тоа, најчесто се избира подмножество од популацијата кое го нарекуваме *примерок*. Примерокот како подмножество од популацијата треба на најдобар начин да ги претставува останатите елементи од популацијата кои не се вклучени во примерокот. Изборот на елементи од популацијата кои ќе бидат елементи на примерокот, може да се направи на повеќе начини, но според најосновната поделба на методите за избор на примерок можеме да ги поделиме во две групи:

- Веројатносни методи за избор на примерок;
- Неверојатносни методи за избор на примерок. [1]

Примероците кои се избираат со помош на веројатносни методи за избор на примерок ги нарекуваме *веројатносни примероци*. Веројатносниот примерок уште се нарекува и *случаен примерок*. При веројатносен примерок, секој елемент на популацијата има ненулта веројатност да биде вклучен во примерокот, која ја нарекуваме *веројатносноста на избор*. Придружувањето на веројатност на избор на секој елемент од популацијата треба да биде објективно и популацијата да биде прецизно дефинирана. Доколку предмет на истражување се вработени во трговија на мало во една држава во текот на 2023 година, тоа значи дека популацијата ја сочинуваат оние вработени во трговија на мало,

кои се вработени во претпријатијата кои се занимаваат со трговија на мало во државата во наведениот период. Во овој случај може да се примени веројатносен метод за избор на примерок, заради тоа што популацијата е прецизно дефинирана со познат број на елементи.

Примероците кои се избираат со помош на неверојатносни методи за избор на примерок уште ги нарекуваме и *неслучајни примероци* или *намерни (со процена) примероци*. При неслучаен примерок, секој елемент од популацијата нема можност да учествува при изборот на примерокот врз основа на кој ќе се спроведе истражувањето. Всушност, изборот на елементи за примерокот од популацијата е врз основа на субјективна проценка на лицето кое го спроведува истражувањето. На пример, доколку спроведуваме истражување за карактеристиките на луѓето коишто посетуваат театар и притоа испраќаме лице кои ќе ги анкетира посетителите во одреден ден. Затоа што не знаеме кои лица во текот на тој ден ќе посетат театар, изборот на лицата кои ќе бидат анкетирани е врз основа на проценка на анкетарот кој го спроведува истражувањето, што всушност е субјективна проценка. За разлика од случајниот примерок, неслучајниот примерок не бара популацијата да е прецизно дефинирана. Според тоа, применувањето на техники од неверојатносни методи за избор на примерок е можно и во случај на општа категоризација на популацијата и во случај кога е специјална категоризација на популацијата. Со други зборови, овие техники можеме да ги користиме и во истражување кое сакаме да го спроведеме на вработени во претпријатија кои се занимаваат со трговија на мало и во истражување кое се спроведува за вработени во претпријатија во трговија на мало во одредена држава во текот на 2023 година. Во првиот случај, елементи на популацијата се сите вработени во претпријатија кои се занимаваат со трговија на мало, односно бројот на елементи во популацијата е „бесконечен“. Во вториот случај, елементи на популацијата се сите вработени во претпријатија кои се занимаваат со трговија на мало со конкретно определена држава, во текот на 2023 година, односно со ваквата спецификација сме се ограничиле на популација која се состои од конечен број на елементи. Техниките кои се користат при

неверојатносни методи за избор на примерок ни овозможуваат да избереме примерок од популација која се состои од „беско-нечен“ број на елементи. Овие начини за избор на примерок, најчесто се користат за спроведување на пилот истражувања со цел да се добијат некои нови идеи и начини на истражување, кои ќе бидат предмет на тестирање за тоа дали можат и како можат во иднина да се имплементираат во системот на истражување.

1. ВИДОВИ ВЕРОЈАТНОСНИ ПРИМЕРОЦИ

Како што видовме претходно, веројатносниот примерок е карактеристичен по тоа што секој елемент од популацијата од која се избира примерокот има ненулта веројатност да биде вклучен во примерокот, односно на секој елемент од популацијата му се придружува веројатност на избор. Притоа разликуваме повеќе видови на веројатносни примероци, односно случајни примероци, и тоа:

- 1) Прост случаен примерок;
- 2) Систематски случаен примерок;
- 3) Стратифициран случаен примерок;
- 4) Кластер примерок;
- 5) Повеќестепен примерок.

Во продолжение ќе дадеме краток опис на примероците кои спаѓаат во оваа група, а во наредниот параграф детално ќе го разгледаме стратифицираниот случаен примерок.

1) *Прост случаен примерок* е примерокот во кој секој елемент од популацијата има подеднаква можност да биде избран во примерокот, односно сите елементи на популацијата имаат еднаква *веројатност на избор*. Популацијата во овој вид на примерок се состои од конечен број на елементи кои можат да се претстават во листа, која ја нарекуваме *примерочна рамка*. Притоа, елементите од популацијата не треба да се преклопуваат. Исто така, елементите на популацијата мора да се хомогени, односно да задоволуваат одредена карактеристика која е предмет на истражување. Според тоа, добиваме дека е потребно популацијата да биде добро дефинирана, односно кои карактеристики треба да ги има еден елемент, за да биде вклучен во популацијата којашто е предмет на истражување.

2) *Системајскиој случаен примерок* исто така се користи за хомогена популација. За разлика од простиот случаен примерок, овој вид на примерок се карактеризира со тоа што секој елемент од популацијата нема еднаква веројатност на избор. Во овој случај, елементите на примерокот ги избираме со таканаречен *чекор на избор*. Чекорот на избор може да биде претставен во минути, редни броеви или зафатнина на простор. На пример, елемент на примерокот е секој петти елемент од рамката, или елемент на примерокот се лица на одреден граничен премин така што времето поминато помеѓу едно и друго анкетирано лице е 15 минути. За овој вид на примерок, поволно е да биде формирана примерочната рамка, но и без неа може да биде спроведено истражувањето. Ќе го илустрираме методот за избор на систематски примерок. Нека големината на систематскиот случаен примерок е n . Нека

$$k = \begin{cases} \frac{N}{n}, \text{ ако } \frac{N}{n} \in \mathbb{Z} \\ \left[\frac{N}{n} \right] + 1, \text{ ако } \frac{N}{n} \notin \mathbb{Z} \end{cases},$$

каде што N е бројот на елементи во популацијата, n бројот на елементи во примерокот, а со $\left[\frac{N}{n} \right]$ го означуваме најголемиот цел број помал од $\frac{N}{n}$. На случаен начин избираме еден цел број од интервалот $[1, k]$ кој го означуваме со R . [2] Тогаш, во систематскиот случаен примерок влегуваат сите елементи од популацијата чишто редни броеви согласно нашето нумерирање се во множеството

$$S = \{R + lk \mid l \in \mathbb{N}\}.$$

3) *Сирајифициран случаен примерок* е примерок кој се користи за истражувања во кои популацијата е хетерогена, односно секој елемент на популацијата не ги задоволува сите карактеристики кои се предмет на истражување. Всушност елементите на популацијата се разликуваат еден од друг по одредена карактеристика. Поради тоа, се формираат подгрупи од популацијата

кои се хомогени, односно сите елементи на подгрупата задоволуваат исти карактеристики. Да забележиме дека, карактеристиките според кои се формирани подгрупите од популацијата се дел од карактеристиките кои се предмет на истражување. Вака формираните подгрупи ги нарекуваме *сѝраѝуми*. ([1],[2],[4])

4) Подгрупите од популацијата кои природно се наметнуваат и ги групираат елементите на популацијата се нарекуваат *класѝери*. Всушност, група од елементи кои припаѓаат на одредено географско подрачје се нарекува *класѝер*. ([1],[2]) Примерокот кој се состои од кластери се нарекува *класѝер ѝриимерок*. Овој вид примерок се користи кога елементите на популацијата се распространети на широко географско подрачје. Популацијата е поделена на подгрупи наречени кластери врз база на нивната географска распределба. На некој начин, оваа поделба е слична на поделбите на популацијата кои ги спомнавме претходно. На пример, популацијата во една држава при кластер примерок ја делиме на кластери кои се всушност градовите, популацијата во градовите ја делиме во кластери што се всушност општините, итн. Кластерите треба да бидат хомогени помеѓу себе согласно карактеристиките кои се предмет на истражување. Да забележиме дека хомогеноста на кластерите, не значи дека е задолжително елементите на кластерот да се хомогени. Со други зборови, доколку популација се градовите, а кластери училиштата, не мора да значи дека секое училиште задоволува исти карактеристики. При спроведување на истражување со кластер примерок, најважно е да забележиме дека популацијата се дели на кластери. Изборот на примерокот се прави на ист начин како кај простиот случаен примерок или како кај систематскиот случаен примерок, со тоа што во случајов на случаен начин се избираат кластерите. Она по што кластер примерокот се разликува од останатите примероци, а воедно и она што е негова најзначајна карактеристика е тоа што сите елементи на кластерот без разлика дали станува збор за личности, домаќинства, училишта и слично, се анкетирани.

5) При избор на *комбиниран (ѝовеќееѝаѝен)* примерок се комбинираат две или повеќе техники кои се користат за избор на веројатносни примероци. Овој вид примерок се користи кога

елементите на популацијата се распространети на пошироко географско подрачје и не е возможно да се избере соодветен примерок со примена само на една од претходно споменатите техники. Всушност, можеме да кажеме дека во овој случај избираме примерок од примерокот. Елементите од популацијата кои ќе бидат анкетирани, се добиваат со повеќестепен избор, користејќи повеќе техники за избор. [1]

Вообичаено, во првата етапа, популацијата се дели на кластери и на случаен начин се избираат кластери. Кластерите од поделбата се хомогени помеѓу себе, но елементите на кластерите можат да бидат хетерогени. За да се избегне хетерогеноста, формираме стратуми во кластерите кои се претходно избрани на случаен начин. Според тоа, кластерите ги нарекуваме првостапни, а стратумите второстапни единици. Формирањето на овие стратуми може да се изведе со помош на техники за избор на кластер примерок или со техники за избор на стратифициран примерок, во зависност од природата на истражувањето. Во некои случаи е потребно елементите на одреден стратум да се поделат и на помали единици, како на пример, трговските центри по продавници, зградите по станови и слично. [3]

2. СТРАТИФИЦИРАН СЛУЧАЕН ПРИМЕРОК

Како што видовме во претходниот параграф, стратифицираниот случаен примерок се базира на формирањето на хомогени подгрупи од популацијата, односно групирање на елементите коишто задоволуваат исти карактеристики по групи наречени стратуми.

Критериумите според кои се формираат стратумите произлегуваат од самата природа и целта на истражувањето за кое се формира примерокот. Поради тоа, критериумите за стратифицирање се разликуваат од едно до друго истражување, па дури и кога се работи за истражувања од тесно поврзани области. Најчесто употребувани критериуми за креирање на стратуми се пол, возраст, социоекономски статус и слично. Формирањето на стратумите може да се толкува како разбивање на популацијата на дисјункти подмножества, наречени стратуми.

Елементите на стратифициран случаен примерок се избираат на случаен начин од секој стратум од популацијата. На тој начин примерокот ја наследува особината за поделба по стратуми, по што го добил и името. Определувањето на бројот на елементи во секој од стратумите, може да се направи на два начина:

- Пропорционално распоредување – бројот на предвидени елементи за примерокот се распоредува пропорционално по стратуми во однос на бројот на елементите во стратумите на популацијата;

- Рамномерно распоредување – бројот на предвидени елементи на примерокот се распоредува рамномерно на секој стратум, односно бројот на елементи во секој стратум е еднаков на количникот помеѓу бројот предвидените елементи на примерокот и вкупниот број на стратуми.

При стратифициран случаен примерок неопходно е да се формира примерочна рамка, односно да се определат карактеристиките кои треба да бидат задоволени за нешто да биде елемент на популацијата која ќе биде предмет на истражување. Во зависност од природата на истражувањето, согласно примерочната рамка, се определува кои карактеристики ќе бидат вклучени при формирање на подгрупите од популацијата, односно стратумите. Исто така, покрај примерочна рамка, потребно е да се изготви и *рамка на сѐраиум*, за секој стратум од популацијата. Рамките мора да бидат комплетно определени, за да може да се поседуваат сите потребни информации и за да може да се исфрлат сите елементи од популацијата кои поради одредена оправдана причина не можат да бидат дел од примерокот.

Елементите на примерокот се избираат на случаен начин, со исти методи како и кај простиот случаен примерок. Во случај кога е невозможно да се формира примерочна рамка, а со тоа е и невозможно да се формира рамка на стратум, елементите на примерокот се избираат со исти методи како и кај систематскиот случаен примерок. Стратифицираниот случаен примерок, е од големо значење во случај кога се разгледува хетерогена популација. Со помош на овој вид на примерок, при хетерогена популација, можеме да избереме соодветен примерок чиј елементи добро ја опишуваат таа популација. При хетерогена популација,

тешко можеме да избереме примерок кој добро ја опишува популацијата, ако тоа го правиме со прост случаен примерок или со систематски случаен примерок.

Пример 1. Спроведуваме истражување за структурата и влијанието на малите, средните и големите претпријатија во економијата на една држава. Општо е познато дека влијанието и структурата на наведените групи на претпријатија е различна, па користењето на прост случаен примерок во овој случај не е повољно. Да претпоставиме дека може да се примени прост случаен примерок. Тогаш, елементите на примерокот се избираат на случаен начин од популацијата. Затоа што станува збор за хетерогена популација, може да се случи да не биде избран ниту еден елемент од подгрупа од популацијата со елементи кои се единствени во популацијата и задоволуваат одредена карактеристика која е предмет на истражување. Според тоа, добиваме дека примерокот не е соодветен, односно немаме елементи во примерокот со чија помош би оценувале одредена карактеристика која е предмет на истражување. Од истите причини, и систематскиот случаен примерок не е соодветен примерок за истражување на кое одговара хетерогена популација.

Определувањето на бројот на стратуми зависи од многу фактори како што се: потешкотијата при креирање на примерочните рамки на стратумите и трошоците за стратифицирање, [2]. Најважно што треба да се забележи е дека колку повеќе информации имаме за популацијата, толку подобри стратуми можеме да формираме. Често собираме премилинарни информации за карактеристиките на популацијата според кои врз база на експертско мислење и анализа на истите ги креираме стратумите. Основните информации за карактеристиките на популацијата наједноставно ги добиваме со помош на неверојатносен примерок. Исто така, при креирање на стратумите можеме да користеме помошни информации од споредување на истражувањето од претходните години или од други истражувања кои се во корелација со истражувањето за кое треба да креираме стратуми.

Кога станува збор за стратифициран случаен примерок, зборуваме за примерок кој бара вложување на големи напори во процесот за избор, многу потрошено време, како и големи тро-

шоци за спроведување на истражување со овој вид на примерок. Исто така, како потешкотија и слабост при избор на овој вид на примерок се јавува фактот дека погрешно избраните карактеристики за формирање на стратуми може да предизвикаат пропаѓање на целото истражување и неупотребливост на добиените податоци.

2. МАТЕМАТИЧКА РЕПРЕЗЕНТАЦИЈА НА СТРАТИФИЦИРАН СЛУЧАЕН ПРИМЕРОК

Како што видовме, при стратифициран случаен примерок вршине разбивање на популацијата на дисјункти подмножества, наречени стратуми. Разгледуваме хетерогена популација Π со N елементи, поделена во H стратуми со N_h елементи во стратум h , $h = 1, 2, \dots, H$. Притоа, е исполнето равенството

$$N_1 + N_2 + \dots + N_H = N.$$

Со n_h да го означиме бројот на елементи од примерокот во стратум h , $h = 1, 2, \dots, H$, кои се избираат на случаен начин од N_h елементи од стратум h од популацијата Π . Според тоа, за бројот на елементи во примерокот, n , добиваме $n = n_1 + n_2 + \dots + n_H$.

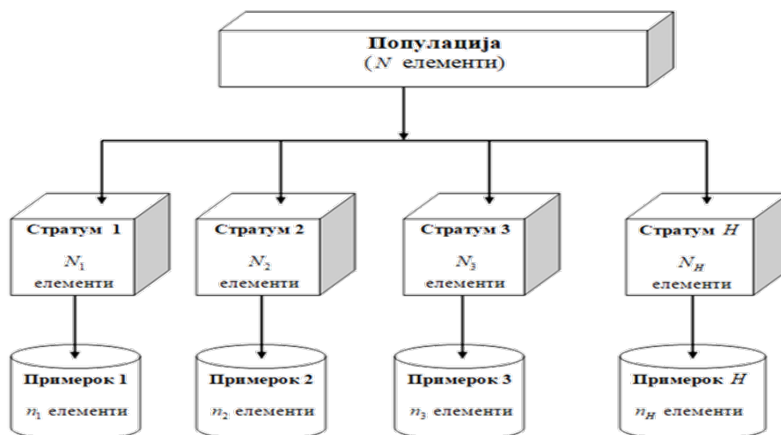
Со S_h го означуваме множеството со n_h елементи од стратум h од популацијата Π . Ако со S го означиме множеството на примерокот, тогаш е исполнето $S_h \subset S$ и

$$S = \bigcup_{h=1}^H S_h, \text{ така што } S_i \cap S_j = \emptyset, \text{ за } i \neq j.$$

Исто така, да забележиме дека $n_h \leq N_h$, односно $n \leq N$. Од сето ова, добиваме дека при избор на стратифициран случаен примерок, избираме H прости случајни примероци, по еден од секој стратум. Всушност, при избор на елементи во стратум h од примерокот, избираме прост случаен примерок со n_h елементи од популација со N_h елементи, [3].

При секое примерочно истражување, главната цел е врз основа на избраниот примерок да се донесат одредени заклучоци

за карактеристиките на популацијата. Како помошна алатка за донесување на тие заклучоци се користат оценувачи.



Слика 1. Дијаграм на стратифициран случаен примерок

Изборот на соодветен оценувач за оценување на одредена карактеристика на елементите на популацијата е сложена постапка која зависи од природата на истражувањето и бројот на податоци со кои располагаме. Сепак, за оценување на одредени карактеристики на елементите од популацијата постојат општо познати оценувачи за кои е докажано дека се добри оценувачи за тие карактеристики. Притоа, една од поважните алатки при оценување се тежинските фактори со чија помош добиените резултати од примерокот можеме да ги обопштиме на ниво на популацијата која ја разгледуваме. Во продолжение ќе го воведеме поимот за примерочен тежински фактор и ќе го определиме неговиот облик во случај на прост и стратифициран случаен примерок.

3. ПРИМЕРОЧЕН ТЕЖИНСКИ ФАКТОР

Кај случајните примероци, на секој елемент на популацијата му се придружува веројатност на избор. Веројатноста на избор на $i \in \Pi$ е еднаква на веројатноста од множество со N елементи да се извлечат n елементи, така што i да биде во множеството извлечени n елементи. Според тоа,

$$\pi_i = P\{\text{елементот } i \text{ да биде во примерокот}\} = \frac{C_{N-1}^{n-1}}{C_N^n}, [2].$$

При случајни примероци, π_i се познати пред спроведување на истражувањето и претпоставуваме $\pi_i > 0$, за секое $i = 1, 2, \dots, N$. Дефинираме *примерочен тежински фактор* како реципрочна вредност од веројатноста на избор, односно

$$w_i = \frac{1}{\pi_i}, [2].$$

Анализирајќи го примерочниот тежински фактор на елементот i , можеме да заклучиме дека со негова помош го добиваме бројот на елементи од популацијата кои не се вклучени во примерокот претставени преку елементот i од примерокот.

Да забележиме дека при прост случаен примерок, веројатностите на избор за секој елемент i од популацијата се еднакви помеѓу себе, односно $\pi_i = \frac{n}{N}$, а според тоа и примерочните тежински фактори се еднакви помеѓу себе, односно $w_i = \frac{1}{\pi_i} = \frac{N}{n}$. Да забележиме дека при прост случаен примерок е исполнето

$$\sum_{i \in S} w_i = \sum_{i \in S} \frac{N}{n} = N.$$

Во случај на стратифициран случаен примерок, примерочните тежински фактори се дефинираат на ниво на стратум. Веројатноста на избор на $j \in \Pi$ од стратум h е еднаква на веројатноста од множеството со N_h елементи да се изберат n_h , така што j да припаѓа на множеството со n_h извлечени елементи. Според тоа,

$$\pi_{hj} = P\{\text{елементот } j \text{ да биде во примерокот}\} = \frac{C_{N_h-1}^{n_h-1}}{C_{N_h}^{n_h}},$$

односно

$$w_{hj} = \frac{1}{\pi_{hj}} = \frac{C_{N_h}^{n_h}}{C_{N_h-1}^{n_h-1}} = \frac{N_h}{n_h}.$$

И во овој случај, примерочниот тежински фактор на елементот j од стратум h го претставува бројот на елементи од популацијата

во стратум h кои не се вклучени во примерокот, претставени преку елементот j од примерокот во стратум h . Да забележиме

дека
$$\sum_{h=1}^H \sum_{j \in S} w_{hj} = \sum_{h=1}^H \sum_{j \in S} \frac{N_h}{n_h} = \sum_{h=1}^H N_h = N.$$
 За разлика од примерочните

тежински фактори при прост случаен примерок, примерочните тежински фактори при страти-фициран случаен примерок можат меѓусебно да се разликуваат. Во случај на стратифициран случаен примерок, сите примерочни тежински фактори се еднакви помеѓу себе, ако во секој стратум има еднаков број на елементи и за распоредувањето на елементите на примерокот се користи рамномерна распределба на елементите.

При секое статистичко истражување важно е да определиме начин, како да ги комбинираме различните информации од примерокот со цел да добиеме еден оценувач кој ќе биде доволно добар и со негова помош да можеме да ги извлечеме потребните информации за параметарот кој го оценуваме. Оценувач на аритметичка средина на популација.

Пример 2. Разгледуваме популација од шест студенти. Бројот на сите можни случајни примероци со четири елементи е еднаков на бројот на комбинации од 6 елементи од класа 4, односно

$C_6^4 = \frac{6!}{4! \cdot 2!} = 15.$ Според тоа, добиваме дека веројатноста на избор

во овој случај е еднаква на $\pi = \frac{4}{6} = \frac{2}{3},$ а тежинскиот фактор е

еднаков на $w = \frac{3}{2}.$ Ако се подели популацијата на два стратума

така што елементи на Стратум 1 се студентите од 1 до 3, а елементи на Стратум 2 се студентите од 4 до 6, сите можни стратифицирани случајни примероци со големина 4, така што се избираат по двајца студенти од секој стратум е производ на бројот на прости случајни примероци од Стратум 1 и бројот на прости случајни примероци од Стратум 2, односно $C_3^2 \cdot C_3^2 = 9.$

Множествата стратифицирани случајни примероци кои ги исполнуваат дадените својства се дадени во следната табела:

Стратифициран случаен примерок

Примерок	{1,2, 4,5}	{1,3, 4,5}	{2,3, 4,5}	{1,2, 4,6}	{1,3, 4,6}	{2,3, 4,6}	{1,2, 5,6}	{1,3, 5,6}	{2,3, 5,6}
Стратум 1	{1,2}	{1,3}	{2,3}	{1,2}	{1,3}	{2,3}	{1,2}	{1,3}	{2,3}
Стратум 2	{4,5}	{4,5}	{4,5}	{4,6}	{4,6}	{4,6}	{5,6}	{5,6}	{5,6}

Во овој случај имаме дека веројатноста на избор е еднаква на $\pi_{hj} = \frac{2}{3}$, а тежинскиот фактор е еднаков на $w_{hj} = \frac{3}{2}$ за секое $h=1,2$ и секое $j=1,2,3,4,5,6$.

Користењето на веројатносни примероци при споведување на одредено истражување е од големо значење доколку сакаме добиените резултати да ги подигнеме на ниво на популација. Сето тоа произлегува од теоријата на веројатност според која и ги пресметуваме примерочните тежински фактори со чија помош можеме да ги обопштиме добиените резултати од примерокот на ниво на популацијата.

ЛИТЕРАТУРА

- [1] A. Mohsin, *A Manual for Selecting Sampling Techniques in Research*, University of Karachi, Iqra Univeristy, 2016.
- [2] L. Sharon, *Sampling Design and Analysis*, Second edition (Advanced Series), State University, Arizona, 2010
- [3] С. Ѓорѓиев, *Теорија на примерок со посебен осврт кон стратифициран случаен примерок*, дипломска работа, ПМФ, 2020
- [4] Montana State University | John. J. Borkowski
<https://math.montana.edu/jobost446/documents/ho3a.pdf>

¹Универзитет „Св. Кирил и Методиј“, Скопје
Природно-математички факултет, Скопје
ул. „Архимедова“ бр. 3, 1000 Скопје, Р. Северна Македонија
e-mail: stevo.gjorgiev@gmail.com

Примен: 11.4.2023

Поправен: 21.9.2023

Одобрен: 25.9.2023

Објавен на интернет: 1.12.2023