Универзитет у Београду
Математички факултет

Марија Минић

Оцењивање параметара код узорковања методом рангираних скупова

— мастер рад —

ментор:
проф. др Весна Јевремовић

чланови комисије:
проф. др Слободанка Јанковић
др Марко Обрадовић

Београд, 2015.

Апстракт

Метода рангираних скупова је статистичка метода узорковања која се користи када су трошкови мерења обележја високи, било финансијски или временски. Идеја је да се јединице узорка рангирају помоћу обележја које је лако и јефтино мерити, а повезано је са обележјем од интереса. Одатле се добија узорак које најбоље осликава популацију, а на коме се врше стварна мерења. Постоји обимна литература о оценама параметара из узорка добијеног методом рангираних скупова. Многе од тих оцена, као што је случај са оценом средње вредности, су непристрасне и имају мању дисперзију у односу на оцене из простог случајног узорка. У раду су приказане оцене параметара из узорка рангираних скупова за нормалну расподелу и Пуасонову расподелу. Изведене су оцене параметара степене расподеле из узорка рангираних скупова различитим методама и упоређиване су са одговарајућим оценама из простог случајног узорка. Оцене из узорка рангираних скупова су се показале као прецизније и тачније.

Кључне речи: метода рангираних скупова, оцењивање параметара, степена расподела

Abstract

Ranked set sampling (RSS) is statistical technique for data collection that can be used when observations are costly or time-consuming. On the other hand ranking of the observations without actual measurement can be done relatively easily. Methods for obtaining ranked set samples are described and the structural differences between ranked set samples and simple random samples (SRS) are discussed. It is shown that estimates based on RSS are generally more efficient than their counterparts based on SRS. Estimates for mean, variance and parameters of Normal and Poisson distribution are shown. Different estimates of Power function distribution parameters, based on SRS and RSS, are computed and compared. Estimates based on RSS generally have smaller bias and MSE.

Key words: Ranked set sampling, parameter estimation, power function distribution

Садржај

Поглавље 1

Увод

Да би се спровело квалитетно испитивање одређене појаве, неопходно је што квалитетније прикупљање података о тој појави. Притом се могу прикупљати подаци о свим јединицима посматраног скупа или само о делу скупа.

Истраживања која обухватају читав статистички скуп често захтевају пуно времена и стварају велике трошкове, поготово када се ради о великом основном скупу. У неким ситуацијама испитивање свих јединица у статистичком скупу није ни могуће. Због тога су се у статистици развиле методе које на основи познавања једног дела скупа, омогућују доношење закључака о целом скупу. Део скупа на коме се прикупљају подаци назива се узорак.

Да би закључци о некој појави на основу узорка били валидни, узорак мора бити репрезентативан. Репрезентативност узорка постиже се исправним избором елемената основног скупа тако да изабрани елементи по својим основним карактеристикама наликују на основни скуп, односно да узорак буде умањена слика основног скупа. Уколико се направи добар избор узорка, обезбеђују се прецизне, поуздане и корисне информације о целом статистичком скупу, уз уштеду времена, трошкова и напора. Према томе, један од најважнијих корака у осмишљавању и реализацији научног истраживања је избор узорка.

Пре самог избора узорка неопходно је направити план узимања узорка у којем је дефинисано како се бирају јединице узорка. Најчешће коришћене методе за избор узорка су: прост случајан узорак, стратификовани узорак и кластер узорак.

Код простог случајног узорка сваки n-точлани подскуп популације има исту вероватноћу са којом може бити изабран у узорак. Најчешће је потребно имати листу свих елемената популације како би се случајним избором могло одабрати n јединица. Избор јединица се може вршити помоћу случајних бројева (користећи таблице или рачунаре) или извлачењем бројева из кутије. Прост случајан узорак је често основа за сложеније планове узорковања.

Код формирања стратификованог узорка, популација се најпре дели у стратуме (слојеве) из којих се затим узимају прости случајни узорци. Стратуми се најчешће формирају од елемената који су слични у односу на обележје које се испитује, при чему се сваки елемент популације налази у тачно једном стратуму. Стратификација, у општем случају, повећава прецизност у односу на прост случајан узорак.

Код кластер узорка елементи популације се групишу у кластере (групе), који постају нове узорачке јединице. У следећем кораку се бира прост случајан узорак кластера, а затим се испитују сви елеменати одабраних кластера, или се поново узима прост случајан узорак елемената из одабраних кластера. Разлика између кластера и стратума што су јединице у стратуму хомогене, а у кластеру хетерогене. Кластер узорковање, у општем случају, смањује прецизност.

С обзиром на то да сваки од представљених планова узорковања има своје предности и мане, стална тежња у теорији узорака је испитивање и унапређивање нових планова узорковања којим би се добиле прецизније оцене параметара који се испитују. Један од мање познатих планова узорковања, али у последње време веома проучаван, је узорковање методом рангираних скупова[1].

У другом поглављу приказане су основе узорковања методом рангираних скупова.

У трећем поглављу дате су оцене параметара добијене из узорка рангираних скупова. Ове оцене су приказане упоредо са оценама из простог случајног узорка да би се оценила ефикасност узорковања методом рангираних скупова. Дате су оцене средње вредности и дисперзије из узорка рангираних скупова. Такође су оцењени параметри нормалне расподеле, као пример непрекидне расподеле и параметар Пуасонове расподеле, као пример дискретне расподеле. Разматрају се оцене из узорка рангираних скупова добијене различитим методама и издвајају су најефикасније.

У четвртом поглављу параметри степене расподеле су израчунати из узорка рангираних скупова и из простог случајног узорка. Оцене су изведене различитим методама: метод момената, комбиновани метод, модификовани метод максималне веродостојности, метода медијане, метода квантила, метода најмањих квадрата и као модификована непристрасна оцена са најмањом дисперзијом. Посматран је случај када су оба параметра степене расподеле непознати и када је један параметар познат. Све

[1] енгл. ranked set sampling (RSS)

добијене оцене су упоређене помоћу Монте Карло симулација. Закључци добијени из овог поглавља су нови резултати. Неки од њих су презентовани на конференцији „Conference on Applied Statistics 2014” у Рибну (Словенија) у оквиру рада „Estimating parameters using Ranked Set Sampling”([14]), као и на симпозијуму „V симпозијум - математика и примене” у Београду у оквиру рада „Узорковање методом рангираних скупова” ([13]).

У последњем, петом поглављу, дати су закључци о ефикасности методе рангираних скупова, њеним предностима и манама. Посебан акценат је стављен на ефикасност новодобијених оцена степене расподеле.

Поглавље 2

Метода рангираних скупова

Метода рангираних скупова се први пут помиње у раду Мекинтајера из 1952. године ([16]). Мекинтајеров приступ је био више интуитиван и испрва није изазвао интересовање научне јавности. Такахаси и Вакимото су у свом раду из 1968. године дали детаљнија теоријска објашњења ([23]), након чега је препознат сав значај ове методе. Данас је литература о методи рангираних скупова веома обимна, али оставља места и за даља истраживања.

Код узорковања методом рангираних скупова најпре се бира релативно велики узорак, али се мерење обележја од интереса за истраживање врше само на малом подузорку. Идеја је да се првобитни узорак рангира по неком обележју који је лако и јефтино мерити, да би права мерења (која најчешће захтевају време и новац) била вршена само на одређеном броју јединица.Уколико се из популације бира прост случајан узорак може се десити да у узорак уђу само екстреми. Идеја методе рангираних скупова је да у узорак уђу што разноврсније јединице, али без великих додатних трошкова.

Најпре се из популације бира прост случајан узорак обима m. Овакав узорак се назива сет. Ове јединице се рангирају по величини на основу обележја које је једноставно и јефтино мерити. Издваја се најмања јединица, а остале јединице се одбацују. Затим се бира нови прост случајан узорак обима m. Јединице се опет рангирају по помоћном обележју, али се сада издваја јединица која је друга по реду, а остале се одбацују. Поступак се понавља m пута, тако што у i-том понављању бира i-та јединица у рангираном низу. На тај начин се добија m јединица које улазе у коначан узорак. У табели 2.1 изабране јединице обележене су црвеном бојом. Овај поступак назива се циклус.

1	$X_{(1)}$	X_2	X_3	$\dots$	X_m
2	X_1	$X_{(2)}$	X_3	$\dots$	X_m
3	X_1	X_2	$X_{(3)}$	$\dots$	X_m
	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
m	X_1	X_2	X_3	$\dots$	$X_{(m)}$

Табела 2.1: Узорак методом рангираних скупова настао из једног циклуса величине m

Циклуси се могу понављати док се не добије жељени обим узорка. Уколико се циклус понови k пута, добија се узорак обима $n = mk$. На издвојеним јединицама врши се мерење обележја које је предмет истраживања. Уколико је, на пример, потребан узорак обима $n = km$, где је $m < k$, потребно је изабрати циклус у коме се бира m јединица и поновити га k пута. У том случају укупан број јединица који се бирају је m^2k. У табели 2.2 приказан је узорак изабран методом рангираних скупова са k циклуса величине m.

циклус 1	$X_{(1)1}$	$X_{(2)1}$	$X_{(3)1}$	$\dots$	$X_{(m)1}$
циклус 2	$X_{(1)2}$	$X_{(2)2}$	$X_{(3)2}$	$\dots$	$X_{(m)2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
циклус к	$X_{(1)k}$	$X_{(2)k}$	$X_{(3)k}$	$\dots$	$X_{(m)k}$

Табела 2.2: Узорак методом рангираних скупова настао из k циклуса где је сваки величине m

Помоћно обележје се често узима према мишљењу стручњака или процене особе са искуством. На пример, уколико је потребно оценити просечан принос одређене биљке на неком газдинству искусан радник може да упореди принос на мањим парцелама без стварног мерења. Заправо, и сам Мекинтајер ([16]) је развио метод рангираних скупова за примену у пољопривреди.

Рангирање се може извршити и на основу неког другог обележја за које је из литературе или искуства познато да корелира са обележјем које се посматра, а које је лакше за мерење. На пример, Самави ([25], стр. 6) у својој студији је мерио ниво билирубина код новорођенчади. Да би се одредио ниво билирубина, потребно је узети крв и анализирати у лабораторији. Међутим, како бебе са вишим нивоом билирубина имају жућу боју лица, беоњача и груди, Самави је уз помоћ датог критеријума најпре рангирао бебе, а затам вршио лабараторијска мерења на мањем подузорку.

Ако помоћно обележје рангира јединице на исти начин као и главно у питању је савршено рангирање. Међутим, то не мора бити случај. Визуалне процене или процене стручњака могу бити погрешне, а ако се рангирање врши по неком другом обележју, корелација између главног и помоћног обележја није идеална. Такво рангирање се назива несавршено. Уколико је рангирање савршено, јединице из узорка се обележавају са $X_{[i]j}$, а уколико је несавршено, са $X_{(i)j}$.

Обим и број циклуса су блиско повезани са начином рангирања јединица. Већи обим циклуса даје више информација. Међутим, код несавршеног рангирања, повећањем циклуса повећава се и грешка. Треба узети у обзир и да велики обим циклуса захтева и узимање великог почетног узорка и рангирање великог броја јединица. Зато, при избору обима и броја циклуса, треба обратити пажњу колико је рангирање близу савршеном, колика је његова цена, и да ли је лако извести га.

Поред основног модела развијене су многе варијације методе рангираних скупова. Величина сета, m, не мора бити константна. Неки сетови су природно неједнаки: број путника на различитим линијама градског аутобуса или број пацијената у чекаоници код доктора. Једно решење је да се величина свих сетова смањи на величину најмањег сета, али се тада беспотребно губе информације.

Није неопходно да се у крајњем узорку налази исти број јединица сваког ранга. Уколико је број јединица сваког ранга исти узорак је уравнотежен, у супротном је неуравнотежен. У неким случајевима неуравнотежен узорак даје прецизније оцене.

Модификација основне методе према којој се број јединица одређеног ранга које се бирају за узорак одређује на основу стандардног одступања ранга назива се метода рангираних скупова са оптималном алокацијом [1]. Да би се ова техника применила потребна је а приори оцена стандардног одступања ранга. Међутим, ова информација је ретко позната унапред. Ипак, ова техника се може применити и уколико је једино познато да је расподела асиметрична. Уколико је, на пример, расподела асиметрична удесно, стандардна девијација се повећава са повећањем ранга([17]).

Мутлак ([15]) је предложио узорковање методом рангираних скупова

[1]енгл. optimal ranked set sampling (ORSS)

помоћу медијана [2]. Нека је расподела посматраног обележја унимодална и симетрична око θ. Параметар θ се може оценити тако што се у сваком од m узорака из једног циклуса издваја медијана. Међутим, овакав модел неће бити добар избор уколико је потребно оценити дисперзију обележја. У том случају је боље користити уравнотежени узорак.

Узорак се може изабрати тако да се користе само екстреми [3] ([20]). Овај метод се спроводи тако што се из узорака наизменично бирају најмања и највећа вредност. Метода рангираних скупова помоћу екстрема се користи код симетричних расподела.

У овом раду је посматран само уравнотежен узорак рангираних скупова.

[2] енгл. median ranked set sampling (MRSS)

[3] енгл. extreme ranked set sampling (ERSS)

Поглавље 3

Оцене параметара

У овом поглављу биће приказани неки од досадашњих резултата у области оцењивања параметара из узорка рангираних скупова. Главни циљ је приказати однос оцена параметара из узорка рангираних скупова и простог случајног узорка.

Дефиниција 3.0.1 *Пристрасност оцене $\hat{\theta}$ параметра θ је разлика између очекиване вредности од $\hat{\theta}$ и θ, тј. $Bias(\theta) = E(\hat{\theta}) - \theta$. Оцена чија је пристрасност 0 је непристрасна оцена и задовољава $E(\hat{\theta}) = \theta$ за све θ.*

Дефиниција 3.0.2 *Средњеквадратна грешка [1] оцене $\hat{\theta}$ параметра θ је функција од θ дефинисана са $E(\hat{\theta} - \theta)^2$ и обележава се са $MSE(\hat{\theta})$.*

Декомпозиција MSE се може извршити на следећи начин $MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = Bias^2(\hat{\theta}) + D(\hat{\theta})$. Према томе, за непристрасну оцену важи:

$$MSE(\hat{\theta}) = D(\hat{\theta}).$$

Квалитет две оцене се може упоредити посматрањем количника њихових средњеквадратних грешака. Уколико су две оцене непристрасне, овај количник је једнак количнику дисперзија и дефинише се као релативна ефикасност.

[1] енгл. mean square error (MSE)

Дефиниција 3.0.3 *Уколико су $\hat{\theta}_1$ и $\hat{\theta}_2$ две непристрасне оцене параметра θ, релативна ефикасност*[2] *је*

$$RE(\hat{\theta}_1, \hat{\theta}_2) = \frac{D(\hat{\theta}_1)}{D(\hat{\theta}_2)}$$

Нека, на даље, важи да је дат узорак рангираних скупова случајне промењиве X обима $n = mk$.

$$\begin{array}{ccccc} X_{(1)1} & X_{(2)1} & X_{(3)1} & \dots & X_{(m)1} \\ X_{(1)2} & X_{(2])} & X_{(3)2} & \dots & X_{(m)2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{(1)k} & X_{(2)k} & X_{(3])} & \dots & X_{(m)k} \end{array}$$

Случајна промењива X има функцију расподеле F и густину расподеле f. Јединице у истој колони имају исту заједничку расподелу. Обележимо са $F_{[i]}$ и $f_{[i]}$ функцију и густину те расподеле.

3.1 Непараметарски приступ

У овом поглављу, нису дате никакаве претпоставке о расподели обележја које посматрамо.

3.1.1 Оцена средње вредности

Пре него што пређемо на оцену средње вредности прикажимо једну особину неких механизама рангирања која ће касније бити од користи.

Дефиниција 3.1.1 *За механизам рангирања се каже да је доследан уколико за свако x важи следећа једначина*

$$F(x) = \frac{1}{m}\sum_{i=1}^{m} F_{[i]}(x).$$

Теорема 3.1 *Следећи механизми рангирања су доследни [25]:*

i) савршено рангирање без помоћног обележја

[2] енгл. relative efficiency

ii) несавршено рангирање без помоћног обележја

ii) рангирање са помоћним обележјем

Доказ:

i)

Уколико је рангирање савршено, јединице из i-те колоне, $i = 1, ..., m$, имају густину расподеле као и i-та статистика поретка. Другим речима важи једнакост $f_{[i]}(x) = f_{(i)}(x)$. Како је

$$f_{(i)}(x) = \frac{m!}{(i-1)!(m-i)!} F(x)^{(i-1)}[1 - F(x)]^{(m-i)} f(x)$$

може се израчунати да је

$$f(x) = \frac{1}{m} \sum_{i=1}^{m} f_{(i)}(x),$$

и према томе,

$$f(x) = \frac{1}{m} \sum_{i=1}^{m} f_{[i]}(x).$$

тј. савршено рангирање без помоћног обележја је доследно.

ii)

Код несавршеног рангирања без помоћног обележја јединице узорка из i-те колоне немају густину $f_{(i)}$. Ипак, функција расподеле i-те колоне се може записати у облику

$$F_{[i]}(x) = \sum_{s=1}^{m} p_{si} F_{(s)}(x),$$

где је са p_{si} обележена вероватноћа да s-та статистика поретка буде оцењена рангом s. Уколико су вероватноће исте за сваки циклус, $\sum_{s=1}^{m} p_{si} = \sum_{i=1}^{m} p_{si} = 1$. Према томе,

$$\begin{aligned} \frac{1}{m} \sum_{i=1}^{m} F_{[i]}(x) &= \frac{1}{m} \sum_{i=1}^{m} \sum_{s=1}^{m} p_{si} F_{(s)}(x) \\ &= \frac{1}{m} \sum_{s=1}^{m} \left(\sum_{i=1}^{m} p_{si} \right) F_{(s)}(x) = F(x) \end{aligned}$$

Овим је доказано да је несавршено рангирања без помоћног обележја доследно.

iii)

Код рангирања са помоћним обележјем, уместо по обележју X, рангирање се врши на основу обележја Y, које је у вези са обележјем X. Нека је $Y_{(i)}$ i-та статистика поретка обележја Y. Нека је са $f_{X|Y_{(i)}}(x|y)$ обележена условна функција густине за X под условом $Y_{(i)} = y$ и $g_{(i)}(y)$ је маргинална функција густине $Y_{(i)}$. Тада је

$$f_{[i]}(x) = \int f_{X|Y_{(i)}}(x|y)g_{(i)}(y)dy$$

Одатле,

$$\begin{aligned} f(x) = & \int \sum_{i=1}^{m} \tfrac{1}{m} f_{X|Y_{(i)}}(x|y)g_{(i)}(y)dy \\ & = \tfrac{1}{m} \sum_{i=1}^{m} f_{[i]}(x) \end{aligned}$$

Другим речима, овакав механизам рангирања је доследан. □

Сада је могуће дати оцену средње вредности из узорка рангираних скупова.

Средња вредност обележја оцењена из узорка рангираних скупова дата је још на почетку развоја ове методе. Мекинтајер ([16]) је желео да нађе ефикаснији и бољи начина за процену просечног приноса пашњака. Такахаси и Вакимото ([23]) су касније дали и особине ове оцене.

Теорема 3.2 *Дата је случајна величина X чија је средња вредност μ, а дисперзија σ^2. Нека је изабран узорак методом рангираних скупова и нека је механизам рангирања доследан. Оцена средње вредности* [3] *дата је са*

$$\hat{\mu}_{RSS} = \frac{1}{mk} \sum_{i=1}^{m} \sum_{j=1}^{k} X_{(i)j}. \tag{3.1}$$

За ову оцену важи

i) $\hat{\mu}_{RSS}$ *је непристрасна оцена, односно* $E(\hat{\mu}_{RSS}) = \mu$

ii) $D(\hat{\mu}_{RSS}) \leq \frac{\sigma^2}{mk}$ *где једнакост важи само у случају да је механизам рангирања потпуно насумичан.*

[3] Ознака RSS (Ranked Set Sample) у формули означава да је оцена добијена из узорка рангираних скупова

Доказ:

i)

$$E(\hat{\mu}_{RSS}) = \frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k}E(X_{(i)j}) = \frac{1}{m}\sum_{i=1}^{m}E(X_{(i)1})$$

Последња једнакост је тачна зато што $X_{(i)1}, X_{(i)2}, ..., X_{(i)k}$ имају заједничку расподелу $F_{[i]}$. Даље,

$$E(\hat{\mu}_{RSS}) = \frac{1}{m}\sum_{i=1}^{m}\int xdF_{[i]}(x) = \int xd\frac{1}{m}\sum_{i=1}^{m}F_{[i]}(x) = \int xdF(x) = \mu$$

Последња једнакост важи зато што је рангирање доследно. Овим је тврдња доказана.

ii)

$$\begin{aligned} D(\hat{\mu}_{RSS}) &= \tfrac{1}{(mk)^2}\sum_{i=1}^{m}\sum_{j=1}^{k}D(X_{(i)j}) = \tfrac{1}{m^2k}\sum_{i=1}^{m}D(X_{(i)1}) \\ &= \tfrac{1}{mk}\left(\tfrac{1}{m}\sum_{i=1}^{m}(E(X_{(i)1})^2 - (E(X_{(i)1}))^2)\right) \\ &= \tfrac{1}{mk}\left(\mu_2 - \tfrac{1}{m}\sum_{i=1}^{m}(E(X_{(i)1}))^2\right) \end{aligned}$$

Са μ_2 је обележен други моменат случајне величине X. Из Коши-Шварцове неједнакости следи

$$\frac{1}{m}\sum_{i=1}^{m}(EX_{(i)1})^2 \geq \left(\frac{1}{m}\sum_{i=1}^{m}(EX_{(i)1})\right)^2 = \mu^2$$

где једнакост важи ако је $E(X_{(1)1}) = E(X_{(2)1}) = ... = E(X_{(m)1})$, што је случај када је механизам рангирања потпуно насумичан. Одатле, $D\hat{\mu}_{RSS} \leq \frac{\sigma^2}{mk}$. $\square$

Ова теорема се може уопштити на произвољну функцију $h(x)$. Детаљније о томе се може наћи у [25].

На основу наведене теореме, могуће је упоредити оцене средње вредности добијене из узорка рангираних скупова и из простог случајног узорка. Нека је са $\hat{\mu}_{SRS}$ обележена оцена која је добијена методом момената из простог случајног узорка[4] , обима $n = mk$. Као што је познато

[4]Ознака SRS (Simple Random Sample) означава да је оцена добијена из простог случајног узорка

ова оцена је непристрасна и њена дисперзија је $\frac{\sigma^2}{km}$. Оцене $\hat{\mu}_{SRS}$ и $\hat{\mu}_{RSS}$ могу се упоредити помоћу релативне ефикасности.

$$RE(\hat{\mu}_{SRS}, \hat{\mu}_{RSS}) = \frac{D(\hat{\mu}_{SRS})}{D(\hat{\mu}_{RSS})}$$

На основу резултата $ii)$ из претходне теореме важи

$$RE(\hat{\mu}_{SRS}, \hat{\mu}_{RSS}) \geq 1$$

Како је релативна ефикасност већа од јединице, оцена средње вредности из простог случајног узорка је мање ефикасна од оцене из узорка рангираних скупова.

3.1.2 Оцена дисперзије

Нека су дати прост случајан узорак обима mk

$$(X_{11}, X_{21}, ..., X_{m1}, X_{12}, ..., X_{mk})$$

и узорак рангираних скупова истог обима

$$(X_{(1)1}, X_{(2)1}, ..., X_{(m)1}, X_{(1)2}, ..., X_{(m)k}).$$

Оцена дисперзије из простог случајног узорка је

$$s^2 = \frac{1}{mk-1} \sum_{i=1}^{m} \sum_{j=1}^{k} (X_{ij} - \hat{\mu}_{SRS})^2$$

где је $\hat{\mu}_{SRS} = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{k} X_{ij}$. Ова оцена је непристрасна и има дисперзију

$$D(s^2) = \frac{2\sigma^4}{mk-1}$$

По узору на ову оцену, оцена дисперзије из узорка рангираних скупова је

$$\hat{\sigma}^2_{RSS} = \frac{1}{mk-1} \sum_{i=1}^{m} \sum_{j=1}^{k} (X_{(i)j} - \hat{\mu}_{RSS})^2, \tag{3.2}$$

где је $\hat{\mu}_{RSS} = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{k} X_{(i)j}$. У циљу одређивања очекивања и дисперзије ове оцене, потребно је увести следеће ознаке. Нека је $X_1, X_2, ..., X_m$ прост случајан узорак где свако X_i има очекивање μ и дисперзију σ^2. Са $X_{(i)}$ је обележена i-та статистика поретка. Ако су c и k произвољне константе, тада важи

$$\sum_{i=1}^{m}(X_{(i)}-c)^k=\sum_{l=1}^{m}(X_l-c)^k$$

Одатле,

$$\sum_{i=1}^{m}E(X_{(i)}-c)^k=mE(X-c)^k$$

Као посебан случај, добија се да је

$$\sum_i \mu_{(i)}=m\mu$$

и

$$\sum_i \sigma_{(i)}^2=m\sigma^2-\sum_i \tau_{(i)}^2 \tag{3.3}$$

где је $\mu_{(i)}=EX_{(i)}$, $\sigma_{(i)}^2=DX_{(i)}$ и $\tau_{(i)}=\mu_{(i)}-\mu$. Сада, очекивање оцене $\hat{\sigma}_{RSS}^2$ је:

$$\begin{aligned}
E(\hat{\sigma}_{RSS}^2) &= E\left(\tfrac{1}{mk-1}\sum_{i=1}^{m}\sum_{j=1}^{k}(X_{(i)j}-\hat{\mu}_{RSS})^2\right)\\
&= \tfrac{k}{mk-1}\left(\sum_i E(X_{(i)}^2)-mE(\hat{\mu}_{RSS}^2)\right)\\
&= \tfrac{k}{mk-1}\left(\sum_i(\sigma_{(i)}^2+\mu_{(i)}^2)-m(D(\hat{\mu}_{RSS})+\mu^2)\right)
\end{aligned}$$

Када се у овом изразу замени дисперзија $D(\hat{\mu}_{RSS})$ која је израчуната у теореми 3.1

$$D(\hat{\mu}_{RSS})=\frac{1}{mk}\left(\mu_2-\frac{1}{m}\sum_{i=1}^{m}\mu_{(i)}^2\right),$$

добија се

$$\begin{aligned}
E(\hat{\sigma}_{RSS}^2) &= \frac{\sum_i \sigma_{(i)}^2}{m}+\frac{k\left(\sum_i \mu_{(i)}^2-m\mu^2\right)}{mk-1}\\
&= \sigma^2+\frac{\sum_i \tau_{(i)}^2}{m(mk-1)}
\end{aligned}$$

Дакле, оцена $\hat{\sigma}_{RSS}^2$ није непристрасна. Међутим, из (3.3), следи да је

$$\frac{\sum_i \tau_{(i)}^2}{m(mk-1)}\leq\frac{\sigma^2}{(mk-1)}$$

Према томе, оцена $\hat{\sigma}_{RSS}$ је асимптотски непристрасна, било да m или k теже бесконачности. Десна страна горње неједнакости је пристрасност оцене $\frac{(mk-1)s^2}{mk}$ из простог случајног узорка

$$\frac{\sigma^2}{(mk-1)} = \sigma^2 - E\left(\frac{(mk-1)s^2}{mk}\right)$$

Пристрасност оцене $\hat{\sigma}_{RSS}$ није већа од пристрасности оцене $\frac{(mk-1)s^2}{mk}$, мада није могуће уклонити је на исти начин.

Дисперзија оцене $\hat{\sigma}^2_{RSS}$ је

$$\begin{aligned} D(\hat{\sigma}^2_{RSS}) = \ & \tfrac{k}{(mk-1)^2}\big\{\left(\tfrac{mk-1}{mk}\right)^2 \sum_i \mu_{4(i)} + 4\sum_i \tau^2_{(i)}\sigma^2_{(i)} \\ & +4\left(\tfrac{mk-1}{mk}\right)\sum_i \tau_{(i)}\mu_{3(i)} + \tfrac{4k}{(mk)^2}\sum\sum_{i<j}\sigma^2_{(i)}\sigma^2_{(j)} \\ & +\tfrac{2(k-1)-(mk-1)^2}{(mk)^2}\sum_i \sigma^4_{(i)}\big\} \end{aligned}$$

где је $\mu_{k(i)} = E(X_{(i)} - \mu_{(i)})^k$. Уколико је рангирање потпуно насумично $\mu_{k(i)} = E(X-\mu)^k = \mu_k$ за свако i. Уколико се убаци у претходну једнакост добија се

$$D(\hat{\sigma}^2_{RSS}) = \frac{\mu_4}{mk} - \frac{(mk-3)\sigma^2}{mk(mk-1)} = D(s^2)$$

Оцена дисперзије добијена из узорка рангираних скупова, не мора бити ефикаснија од оцене добијене из простог случајног узорка, нарочито ако је обим узорка мали. Међутим, за велике узорке може се доказати да је оцена из узорка рангираних скупова ефикаснија од оцене добијене из простог случајног узорка. Прецизније, важи следећа теорема.

Теорема 3.3 *Нека су изабрани прост случајан узорак и узорак рангираних скупова обима mk из популације чије обележје има четврти моменат. Тада постоји n^*, такво да за $mk > n^*$ важи:*

$$\frac{MSE(\hat{\sigma}^2_{SRS})}{MSE(\hat{\sigma}^2_{RSS})} = \frac{D(\hat{\sigma}^2_{SRS})}{MSE(\hat{\sigma}^2_{RSS})} \geq 1.$$

Доказ теореме може се наћи у ([22]). Прецизност оцене из узорка рангираних скупова може се повећати било повећањем величине циклуса m или броја циклуса k. У пракси, значајно повећање величине циклуса није могуће, па се прибегава повећању броја циклуса.

3.2 Параметарски приступ

У литератури се могу наћи оцене параметара из узорка рангираних скупова за многе расподеле. У овом поглављу биће приказане оцене за параметаре нормалне и Пуасонове расподеле.

3.2.1 Нормална расподела

За оцену параметра μ код нормалне расподеле може се узети оцена (3.1), а за оцену параметра σ^2, постоје побољшања оцене (3.2). Најпре ћу приказати различите оцене дисперзије нормалне расподеле за случај када је узорак рангираних скупова биран из једног циклуса. Затим, у другом делу, приказаћу оцене дисперзије из узорка који је биран из више циклуса, а које су настале као комбинација оцена добијених из узорка са једним циклусом.

Оцена дисперзије: Један циклус

Нека је изабран узорак рангираних скупова из једног циклуса обима m. Оцена $\hat{\sigma}^2_{RSS}$ није погодна за коришћење код малих узорака. Дефинишимо параметар $\nu_{(i)}$ на следећи начин:

$$\nu_{(i)} = \frac{\mu_{(i)} - \mu}{\sigma} = \frac{\tau_{(i)}}{\sigma}.$$

Слично, у општем случају

$$\nu_{k(j)} = \frac{\mu_{k(j)}}{\sigma^k}.$$

Следећа оцена је непристрасна ([26]).

$$\widehat{\sigma}^2_{RSS,mod} = \frac{\sum_{i=1}^{m}(X_{(i)} - \widehat{\mu}_{RSS})^2}{m - 1 + \frac{1}{m}\sum_{i=1}^{m}\nu^2_{(i)}} \tag{3.4}$$

Дисперзија ове оцене је

$$\begin{aligned} D(\hat{\sigma}^2_{RSS,mod}) &= \frac{\sigma^4}{(m-1+\frac{1}{m}\sum_{i=1}^{m}\nu_{(i)})^2}\Big\{\left(\tfrac{m-1}{m}\right)^2\sum_{i=1}^{m}\nu_{4(i)} \\ &+4\sum_{i=1}^{m}\nu^2_{(i)}\nu_{2(i)} + 4\tfrac{m-1}{m}\sum_{i=1}^{m}\nu_{(i)}\nu_{3(i)} \\ &+\tfrac{4}{m^2}\sum\sum_{i<j}\nu_{2(j)}\nu_{2(i)} - \left(\tfrac{m-1}{m}\right)^2\sum_{i=1}^{m}\nu^2_{2(i)}\Big\} \end{aligned}$$

За случај нормалне расподеле дисперзија ове оцене је мања од дисперзије оцене из простог случајног узорка, $D(s^2) = 2\sigma^4/(m-1)$, за $m = 2, 3, 4$ и 5 ([26]).

Још једна оцена дисперзије нормалне расподеле у случају једног циклуса добија се уколико се оцена средње вредности добијена методом максималне веродостојности, $\hat{\mu}_{RSS}$, замени најбољом линеарном непристрасном оценом из узорка рангираних скупова [5].

$$\widehat{\mu}_{blue} = \frac{\sum\limits_{i=1}^{m} X_{(i)}/\nu_{2(i)}}{\sum\limits_{i=1}^{m} 1/\nu_{2(i)}}$$

Тада ће оцена дисперзије нормалне расподеле имати облик:

$$\widehat{\sigma}_1^2 = a_m \sum_{i=1}^{m} \frac{(X_{(i)} - \widehat{\mu}_{blue})^2}{\nu_{2(i)}}, \tag{3.5}$$

$$a_m = \frac{1}{m - 1 - \sum\limits_{j=1}^{m} \frac{\nu_{(j)}^2}{\nu_{2(j)}}},$$

где је константа a_m је изабрана тако да оцена буде непристрасна.

Општи случај оцене (3.5) је

$$\widehat{\sigma}_2^2(b) = g_m(b) \sum_{i=1}^{m} b_i (X_{(i)} - \widehat{\mu}_{blue})^2$$

где је

$$g_m(b) = \left[\sum_{j=1}^{m} b_j \left(\nu_{(j)}^2 + \nu_{2(j)} - \frac{1}{\sum\limits_{l=1}^{m} \frac{1}{\nu_{2(l)}}} \right) \right]^{-1},$$

а b_j су позитивне константе биране тако да оцена буде непристрасна. За $b_j = 1/\nu_{2(j)}$ добија се оцена (3.5). Дисперзија ове оцене је

$$\begin{aligned} D(\hat{\sigma}_2^2(b)) &= \sigma^4 g_m^2(b) \{ \sum_{i=1}^{m} b_i^2 E \left(\tfrac{X_{(i)} - \hat{\mu}_{blue}}{\sigma} \right)^4 \\ &+ \sum\sum_{j \neq i} b_j b_i E \left[\left(\tfrac{X_{(j)} - \hat{\mu}_{blue}}{\sigma} \right)^2 \left(\tfrac{X_{(i)} - \hat{\mu}_{blue}}{\sigma} \right)^2 \right] \} - \sigma^4 \end{aligned}$$

[5]енгл. best binear unbaised estimator (BLUE)

Дисперзија оцене $\hat{\sigma}_1^2$ добија се заменом параметра $b_j = 1/\nu_{2(j)}$. Константе b_j се могу бирати за свако m тако да минимизују дисперзију.

Поред ове три оцене, Ју и сарадници ([26]) су у свом раду посматрали и оцену добијену нумеричким методама из функције веродостојности[6], $\hat{\sigma}^2_{RSS,mle}$. Упоредили су оцене за разне вредности m,а на основу релативне ефикасности. Оцене $\hat{\sigma}_1^2$ и $\hat{\sigma}_2^2$ се понашају приближно исто, са тим да оцена $\hat{\sigma}_2^2$ има предност, с обзиром да је могуће бирати параметре b_j тако да се минимизује дисперзија. Оцене $\hat{\sigma}_1^2$ и $\hat{\sigma}_2^2$ су ефикасније од s^2 јер имају мању дисперзију. Ове оцене су боље и у односу на $\hat{\sigma}^2_{RSS,mod}$. Ефикасност је већа како се m повећава. Са друге стране, оцена $\hat{\sigma}^2_{RSS,mle}$ је ефикаснија у односу на $\hat{\sigma}_1^2$ и $\hat{\sigma}_2^2$. Ипак, израчуната пристрасност за ову оцену је релативно велика, па оцена методом максималне веродостојности није препоручљива. Дакле, оцена $\hat{\sigma}_2^2(b)$ са оптималним b_j најбоља је оцена у односу на остале наведене оцене. Детаљне табеле можете наћи у ([26]).

Оцена дисперзије: Више циклуса

У овом делу приказаћу оцене дисперзије нормалне расподеле из узорка рангираних скупова кад се узорак бира из више циклуса.

Нека је $(X_{(1)1}, X_{(2)1}, ..., X_{(m)1}, X_{(1)2}, ..., X_{(m)k})$ узорак рангираних скупова обима mk, где је m величина циклуса, а k број циклуса. Као и у случају једног циклуса за оцену дисперзије је могуће узети оцену (3.2)

$$\widehat{\sigma}^2_{RSS} = \frac{1}{mk-1}\sum_{i=1}^{m}\sum_{j=1}^{k}(X_{(i)j} - \widehat{\mu}_{RSS})^2.$$

Како је ова оцена пристрасна, модификацијом се добија непристрасна оцена:

$$\widehat{\sigma}^2_{RSS,mod} = \frac{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k}(X_{(i)j} - \widehat{\mu}_{RSS})^2}{mk - 1 + \frac{1}{m}\sum\limits_{i=1}^{m}\nu^2_{(i)}} \tag{3.6}$$

Поред ове, постоје и нове оцене. При рачунању нових оцена параметра σ^2 могуће је користити два приступа.

У првом приступу из сваког циклуса се рачуна оцена параметра, па са за укупну оцену узима аритметичка средина добијених оцена. Нека је узорак биран помоћу k независних циклуса величине m. Оцену дисперзије из j-тог циклуса за $j = 1, 2, ..., k$, биће обележене са $\hat{\sigma}_j^2$. Како

[6]енгл. maximum likelihood estimator (MLE)

је ових k оцена независно и једнако расподељено, оцена дисперзије из целог узорка рангираних скупова је

$$\hat{\sigma}^2_{RSS,ukupno} = \frac{1}{k}\sum_{j=1}^{k}\hat{\sigma}^2_{(j)}$$

За оцену $\hat{\sigma}^2_{(j)}$ може се узети нека од предходно предложених оцена: $\hat{\sigma}^2_{RSS,mod}$, $\hat{\sigma}^2_1$ и $\hat{\sigma}^2_2(b)$. Међутим, ниједан од предложених оцена нема задовољавајућа својства ([26]).

У другом приступу, узорак рангираних скупова се посматра као унија m класа, свака величине k, где је i-та класа i-та статистика поретка, $i = 1, 2, ..., m$. Дакле, узорак се састоји од m подузорака, $(X_{(i)1}, X_{(i)2}, ..., X_{(i)k})$, $1 \leq i \leq m$, где су подузорци међусобно независни. Декомпозиција оцене (3.6) се може извршити на следећи начин

$$\begin{aligned}\hat{\sigma}^2_{RSS,mod,ukupno} &= \frac{1}{mk-1+\sum\limits_{i=1}^{m}\nu^2_{(i)}}\left\{\sum_{i=1}^{m}\sum_{j=1}^{k}(X_{(i)j}-\overline{X}_{(i)})^2 + k\sum_{i=1}^{m}(\overline{X}_{(i)}-\hat{\mu}_{RSS})^2\right\}\\ &= \frac{1}{mk-1+\sum\limits_{i=1}^{m}\nu^2_{(i)}}\left\{(k-1)\sum_{i=1}^{m}V_i + k\sum_{i=1}^{m}(\overline{X}_{(i)}-\hat{\mu}_{RSS})^2\right\}\end{aligned}$$

где је $\overline{X}_{(i)} = \frac{1}{k}\sum\limits_{j=1}^{k}X_{(i)j}$ средња вредност класе i-тих статистика, а $V_i = \frac{1}{k-1}\sum\limits_{j=1}^{k}(X_{(i)j}-\overline{X}_{(i)})^2$ дисперзија у i-тој класи. Израз $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k}(X_{(i)j}-\overline{X}_{(i)})^2$ је дисперзију унутар класе i-тих статистика, а $k\sum\limits_{i=1}^{m}(\overline{X}_{(i)}-\hat{\mu}_{RSS})^2$ је дисперзију између класа i-тих статистика. Дисперзија оцене је

$$\begin{aligned}D(\hat{\sigma}^2_{RSS,mod,ukupno}) &= \frac{k\sigma^4}{\left(mk-1+\frac{1}{m}\sum\limits_{i=1}^{m}\nu^2_{(i)}\right)^2}\Big\{\Big(\frac{mk-1}{mk}\Big)^2\sum_{i=1}^{m}\nu_{4(i)}\\ &+4\sum_{i=1}^{m}\nu^2_{(i)}\nu_{2(i)} + 4\frac{mk-1}{mk}\sum_{i=1}^{m}\nu_{(i)}\nu_{3(i)}\\ &+\frac{4k}{(mk)^2}\sum\sum_{i<j}\nu_{2(i)}\nu_{2(j)} + \frac{2(k-1)+(mk-1)^2}{(mk)^2}\sum_{i=1}^{m}\nu_{2(i)^2}\Big\}\end{aligned}$$

Из ове оцене могуће је извести нове оцене користећи само једну врсту дисперзије или комбиновањем обе (унутар класе и између класа).

Оцена помоћу унутаркласне дисперзије

Нека су $\nu_{2(i)}$ познате константе тако да су V_i непристрасне оцене за дисперзију унутар класе i-тих статистика, где важи $E(V_i) = \sigma^2\nu_{2(i)}$. За V_i важи да су међусобно независне са дисперзијом $D(V_i) = \sigma^4\eta_{(i)}$, где је

$$\eta_{(i)} = \frac{\nu_{4(i)} - \nu_{2(i)^2}}{k} + \frac{2}{k(k-1)}\nu_{2(i)}^2$$

Из претходног директно следи да је најбоља линеарна непристрасна оцена за σ^2 преко унутаркласних дисперзија[7], V_i:

$$\hat{\sigma}^2_{RSS,W,mod} = \sum_{i=1}^{m} \frac{\frac{\nu_{2(i)}}{\eta_{(i)}}}{\sum\limits_{j=1}^{m} \frac{\nu_{2(j)}^2}{\eta_{(j)}}} V_i$$

Дисперзија ове оцене је:

$$D(\hat{\sigma}^2_{RSS,Wmod}) = \frac{\sigma^4}{\sum\limits_{i=1}^{m} \frac{\nu_{2(i)}^2}{\eta_{(i)}}}$$

Оцена помоћу међукласне дисперзије

Слично претходном, могуће је формирати нове оцене користећи само дисперзију између класа. Средње вредности из класа $\overline{X}_{(i)}$, $i = 1, 2, ..., m$ су независне са дисперзијом $D(\overline{X}_{(i)}) = D(X_{(i)j})/k$. Аналогно оценама из једног циклуса, могу се издвојити следеће оцене на основу дисперзије између класа:

1. Оцена добијена коришћењем међукласне дисперзије[8] из оцене 3.4 је

$$\hat{\sigma}^2_{RSS,B,mod} = \frac{\sum\limits_{i=1}^{m} (\overline{X}_{(i)} - \hat{\mu}_{RSS})^2}{\frac{m-1}{k} + \frac{mk-m+1}{mk} \sum\limits_{i=1}^{m} \nu_{(i)}^2}$$

[7] Индекс W у називу оцене потиче од речи од *within* (унутар), и односи се на унутаркласну дисперзију.

[8] Индекс B у називу оцене потиче од речи од *between* (између), и односи се на међукласну дисперзију.

Дисперзија ове оцене је

$$\begin{aligned} D(\hat{\sigma}^2_{RSS,B,mod}) &= \frac{\sigma^4}{\left(\frac{m-1}{k}+\frac{mk-m+1}{mk}\sum\limits_{i=1}^{m}\nu^2_{(i)}\right)^2}\Big\{\left(\frac{m-1}{m}\right)^2\sum_{i=1}^{m}\left[\frac{\nu_{4(i)}+3(k-1)\nu_{2(i)}}{k^3}\right] \\ & 4\sum_{i=1}^{m}\frac{\nu^2_{(i)}\nu_{2(i)}}{k}+4\frac{m-1}{m}\sum_{i=1}^{m}\frac{\nu_{(i)}\nu_{3(i)}}{k^2} \\ & \frac{4}{m^2}\sum\sum_{i<j}\frac{\nu_{2(i)}\nu_{2(j)}}{k^2}-\left(\frac{m-1}{m}\right)^2\sum_{i=1}^{m}\frac{\nu^2_{2(i)}}{k^2}\Big\} \end{aligned}$$

2. Оцена добијена коришћењем међукласне дисперзије из оцене σ_1^2 је

$$\hat{\sigma}^2_{RSS,B1} = a_{m,k}\sum_{i=1}^{m}\frac{(\overline{X}_{(i)}-\overline{\mu}_{blue})^2}{\nu_{2(i)}/k}$$

где је

$$a_{m,k} = \left(m-1+k\sum_{i=1}^{m}\frac{\nu^2_{(i)}}{\nu_{2(i)}}\right)^{-1}$$

и

$$\hat{\mu}_{blue} = \frac{\sum\limits_{i=1}^{m}X_{(i)}/\nu_{2(i)}}{\sum\limits_{i=1}^{m}1/\nu_{2(i)}}$$

Дисперзија добијене оцене се добија када се у изразу $D(\hat{\sigma}_{RSS,B2})$ који је приказан у следећој тачки замени $b_i = 1/\nu_{2(i)}$, $i = 1,2,...,m$.

3. Оцена добијена коришћењем међукласне дисперзије из оцене σ_1^2 је

$$\hat{\sigma}^2_{RSS,B2} = g_{m,k}(b)\sum_{i=1}^{m}b_i(\overline{X}_{(i)}-\overline{\mu}_{blue})^2$$

где је

$$g_{m,k}(b) = \left(\sum_{i=1}^{m}b_i(\nu^2_{(i)}+\frac{\nu_{2(i)}}{k}-\frac{1}{k\sum\limits_{j=1}^{m}\nu_{2(j)})}\right)^{-1}.$$

Константе b_i су такве да минимизирају дисперзију $\hat{\sigma}^2_{RSS,B2}$.

Дисперзија ове оцене је

$$\begin{aligned} D(\hat{\sigma}^2_{RSS,B2}(b)) &= \sigma^4 g^2_{m,k}(b)\{\sum_{i=1}^{m}b_i^2E\left(\frac{X_{(i)}-\hat{\mu}_{blue}}{\sigma}\right)^4 \\ &+\sum\sum_{j\neq i}b_jb_iE\left[\left(\frac{X_{[j]}-\hat{\mu}_{blue}}{\sigma}\right)^2\left(\frac{X_{(i)}-\hat{\mu}_{blue}}{\sigma}\right)^2\right]\}-\sigma^4 \end{aligned}$$

Оцене добијене комбинацијом унутаркласне и међукласне дисперзије

Оцене се могу добити и када се комбинују унутаркласна и међукласна дисперзија[9].

$$\hat{\sigma}^2_{RSS,BW}(\alpha) = \alpha\hat{\sigma}^2_{RSS,W} + (1-\alpha)\hat{\sigma}^2_{RSS,B} \tag{3.7}$$

Класа оцена дисперзије нормалне расподеле где је $\hat{\sigma}^2_{RSS,W}$ оцена дисперзије помоћу унутаркласних дисперзија, а $\hat{\sigma}^2_{RSS,B}$ је једна од оцена дисперзије помоћу међукласних дисперзија. За неко $\alpha\epsilon[0,1]$ оцена $\hat{\sigma}^2_{RSS,BW}(\alpha)$ је непристрасна. Оптимално α које минимизује дисперзију оцене је

$$\alpha^* = \frac{D(\hat{\sigma}^2_{RSS,B}) - cov(\hat{\sigma}^2_{RSS,B}, \hat{\sigma}^2_{RSS,W})}{D(\hat{\sigma}^2_{RSS,W}) + D(\hat{\sigma}^2_{RSS,B}) - 2cov(\hat{\sigma}^2_{RSS,B}, \hat{\sigma}^2_{RSS,W})}$$

За сваку комбинацију оцена $\hat{\sigma}^2_{RSS,B}$ и $\hat{\sigma}^2_{RSS,W}$ могуће је наћи α тако да је дисперзија оцене минимална. Оцене добијене комбинацијом унутаркласне оцене $\hat{\sigma}^2_{RSS,W}$ и међукласних оцена $\hat{\sigma}^2_{RSS,B,mod}$, $\hat{\sigma}^2_{RSS,B1}$ и $\hat{\sigma}^2_{RSS,B2}$ су, редом, $\hat{\sigma}^2_{RSS,WBmod}$, $\hat{\sigma}^2_{RSS,WB1}$ и $\hat{\sigma}^2_{RSS,WB2}$. Дисперзија ових оцена је:

$$\begin{aligned} D(\hat{\sigma}^2_{RSS,BW}(\alpha)) &= \alpha^2 D(\hat{\sigma}^2_{RSS,W}) + (1-\alpha)^2 D(\hat{\sigma}^2_{RSS,B}) \\ &-2\alpha(1-\alpha)cov(\hat{\sigma}^2_{RSS,W}, \hat{\sigma}^2_{RSS,B}) \end{aligned}$$

Када се у овај израз убаци оптимална вредност за α

$$D(\hat{\sigma}^2_{RSS,WB}(\alpha^*)) = \frac{D(\hat{\sigma}^2_{RSS,B})D(\hat{\sigma}^2_{RSS,W}) - cov^2(\hat{\sigma}^2_{RSS,B}, \hat{\sigma}^2_{RSS,W})}{D(\hat{\sigma}^2_{RSS,W}) + D(\hat{\sigma}^2_{RSS,B}) - 2cov(\hat{\sigma}^2_{RSS,B}, \hat{\sigma}^2_{RSS,W})}$$

***

Све претходно наведене оцене, Ју и сарадници([26]) упоредили су симулацијом. Оцене $\hat{\sigma}^2_{RSS,W,mod}$,$\hat{\sigma}^2_{RSS,B,mod}$, $\hat{\sigma}^2_{RSS,B1}$ и $\hat{\sigma}^2_{RSS,B2}$ показују лоше особине када се упореде са оценом дисперзије s^2 из простог случајног узорка и са оценом дисперзије $\hat{\sigma}^2_{RSS,mod}$ из узорка рангираних скупова. Међутим, оцене добијене комбинацијом унутаркласних и међукласних дисперзија $\hat{\sigma}^2_{RSS,WBmod}$, $\hat{\sigma}^2_{RSS,WB1}$ и $\hat{\sigma}^2_{RSS,WB2}$, показују боље особине у односу на s^2. Дисперзија s^2 је већа од дисперзија ових оцена, и та

[9] Индекс BW у називу оцене потиче од $between-within$ (између - унутар), и односи се на унутаркласну и међукласну дисперзију.

разлика је већа са порастом броја циклуса. Сличне особине се могу приметити и када се оцене $\hat{\sigma}^2_{RSS,WBmod}$, $\hat{\sigma}^2_{RSS,WB1}$ и $\hat{\sigma}^2_{RSS,WB2}$ упореде са оценом $\hat{\sigma}^2_{RSS,mod}$. Према томе, све оцене добијене комбинацијом унутаркласних и међукласних дисперзија показују добре особине. Што се тиче међусобног односа ове три оцене $\hat{\sigma}^2_{RSS,WBmod}$ показује најгоре особине. У већини случајева $\hat{\sigma}^2_{RSS,WB2}$ је најбоља оцена. Детаљније о овоме, можете наћи у [26].

3.2.2 Пуасонова расподела

Радови везани за оцене параметра из узорка рангираних скупова углавном испитују непрекидне расподеле. Међутим, параметар који се оцењује из узорка рангираних скупова може бити и из дискретне расподеле. Један од ретких радова који испитују оцењивање параметра дискретне расподеле из узорка рангираних скупова је [2].

Као што је познато, за Пуасонову расподелу важи

$$P(X \leq x) = \frac{\lambda^x}{x!}e^{-x}, \quad x = 0, 1, 2, ...$$

где је $E(X) = \lambda$ и $D(X) = \lambda$.

Уколико је $(X_1, ..., X_m)$ прост случајан узорак оцена параметра λ методом максималне веродостојности је

$$\hat{\lambda}_{SRS} = \frac{1}{m}\sum_{i=1}^{m} X_i.$$

Ова оцена је непристрасна и има дисперзију λ/m.

Слично, уколико је узорак $(X_{(1)1}, X_{(2)1}, ..., X_{(m)1})$ биран методом рангираних скупова са једним циклусом, оцена параметра λ је

$$\hat{\lambda}_{RSS} = \frac{1}{m}\sum_{i=1}^{m} X_{(i)1}$$

Оцена из узорка рангираних скупова може се добити и на други начин. За оцену параметра λ може се узети линеарна комбинација вредности из узорка

$$\lambda^* = \sum \gamma_i X_{(i)1} \tag{3.8}$$

где се γ_i бирају тако да $D(\lambda^*)$ буде минимално у класи непристрасних оцена. Нека су $U_{(i)}$ дефинисани као $U_{(i)} = (X_{(i)1} - \lambda)/\lambda$, за $i = 1, 2, ..., m$. С обзиром да Пуасонова расподела не припада фамилији расподела које имају параметре положаја и расприпења, редуковане статистике поретка $U_{(i)}$ немају расподелу слободну од параметра. Према томе, очекивање случајне величине $U_{(i)}$, $\alpha_{i:m}$ и дисперзија случајне величине $U_{(i)}$, $\nu_{i:m}$ зависе од параметра λ. Уколико се та чињеница занемари за тренутак, очекивање оцене λ^* је

$$E(\lambda^*) = \lambda \sum \gamma_i + \sqrt{\lambda}\alpha_{i:m},$$

а дисперзија

$$D(\lambda^*) = \lambda \sum \gamma_i \alpha_{i:m},$$

при чему је због непристрасности $\sum \gamma_i = 1$ и $\sum \gamma_i \alpha_{i:m} = 0$. Да би се добиле вредности γ_i потребно је минимизирати Лагранжову функцију

$$F = \sum \gamma_i^2 \nu_{i:m} - a(\sum \gamma_i - 1) - b(\sum \gamma_i \alpha_{i:m})$$

Решавањем система једначина

$$\frac{\partial F}{\partial \gamma_i} = 0$$

$$\sum \gamma_i = 1$$

$$\sum \gamma_i \alpha_{i:m} = 0$$

за $i = 1, 2, ..., m$ добија се

$$\gamma_i = \frac{(\sum \alpha_{j:m}^2/\nu_{j:m}) - \alpha_{i:m}(\sum \alpha_{j:m}/\nu_{j:m})}{\left[(\sum 1/\nu_{j:m})(\sum \alpha_{j:m}^2/\nu_{j:m}) - (\sum \alpha_{j:m}/\nu_{j:m})^2\right] \nu_{i:m}} \tag{3.9}$$

Дисперзија оцене λ^* је

$$D(\lambda^*) = \lambda \frac{(\sum \alpha_{i:m}^2/\nu_{i:m})}{(\sum 1/\nu_{i:m})(\sum \alpha_{i:m}^2/\nu_{i:m}) - (\sum \alpha_{i:m}/\nu_{i:m})^2}$$

Дакле , уколико се занемари чињеница да $\alpha_{i:m}$ и $\nu_{i:m}$ зависе од параметра λ, λ^* је оцена са најмањом дисперзијом у класи непристрасних оцена.

Са друге стране, погодно је размотрити најбољу линеарну непристрасну оцену. Када се посматра узорак рангираних скупова, најбоља линеарна нeристрасна оцена је ефикаснија у односу на оцену максималне веродостојности код нормалне, експоненцијалне и других расподела ([21], [3]). Лојд ([9]) је испитивао фамилије непрекидних расподела које имају параметре положаја и расприпења $F((X - \mu)/\sigma)$. Најбоља линеарна непристрасна оцена за вектор параметара $\theta' = (\mu, \sigma)$ заснована на статистикама поретка из једног узорка $X_1, X_2, ..., X_n$ је

$$\theta^* = (A'D^{-1}A)^{-1}A'D^{-1}X$$

са коваријационом матрицом

$$D(\lambda^*) = \sigma^2(A'D^{-1}A)^{-1}$$

где је $A = (1, \alpha)$, $\alpha' = \{\alpha_{i:m}\} = \{E((X_{(i)1} - \mu)/\sigma)\}$, $D = \{\nu_{i:m}\}$ је матрица коваријансе за редуковану статистику поретка $U_{(i)} = X_{(i)} - \mu)/\sigma$; $1' = (1, 1, ..., 1)$ и $X = (X_{(1)}, X_{(2)}, ..., X_{(m)})$. Иако је Пуасонова расподела дискретна и не припада фамилији расподела које имају параметре положаја и распршења, овај начин може бити инспирација за добијање оцене параметра λ. Нека је

$$\begin{aligned}
&\theta' = (\lambda, \sqrt(\lambda)), \mathtt{1} = (1, 1, ..., 1) \quad X' = (X_{(1)1}, X_{(2)1}, ..., X_{(m)1})\\
&\alpha' = \alpha_{i:m} = \{E(\tfrac{X_{(i)1} - \lambda}{\sqrt{\lambda}})\}\\
&V = diag(\nu_{i:m}) = \{D(\tfrac{X_{(i)1} - \lambda}{\sqrt{\lambda}})\}
\end{aligned}$$

Прва компонента случајног вектора θ^* може се записати као

$$\lambda^* = \sum_{i=1}^{m} \delta_i X_{(i)1}$$

где је

$$\begin{aligned}
&\tau_i = \tfrac{c - b\alpha_{i:m}(\lambda)}{\triangle \nu_{i:m}(\lambda)}\\
&c = \tfrac{\sum \alpha^2_{j:m}(\lambda)}{\nu_{i:m}(\lambda)}\\
&b = \tfrac{\sum \alpha_{j:m}(\lambda)}{\nu_{i:m}(\lambda)}\\
&\triangle = (\sum \tfrac{1}{\nu_{j:m}(\lambda)})(\tfrac{\sum \alpha^2_{j:m}(\lambda)}{\nu_{i:m}(\lambda)}) - (\tfrac{\sum \alpha_{j:m}(\lambda)}{\nu_{i:m}(\lambda)})^2
\end{aligned}$$

Ова оцена се поклапа са оценом (3.8) где је γ_i дефинисано са (3.9). Дисперзија оцене λ^* је

$$D(\lambda^*) = \frac{\lambda}{\triangle}\left(\frac{\alpha^2_{j:m}(\lambda)}{\nu_{j:m}(\lambda)}\right).$$

Као што је већ познато λ^* је функција јединица из узорка и константи које зависе од λ. Из тог разлога оцена λ^* не може бити оцена параметра λ. Потребно је наћи замене за $\alpha_{i:m}$ и $\nu_{i:m}$ које не зависе у великој мери од λ. Са тим циљем, најпре је потребно испитати како очекивање и дисперзија редуковане статистика поретка Пуасонове расподеле зависе од λ.

Симулацијом су израчунате вредности $\alpha_{i:m}(\lambda)$ и $\nu_{i:m}(\lambda)$ за величину узорка од 2 до 20 и вредности параметра $\lambda = 1, 3, 5, 10, 15, 20, 25, 30, 35, 40, 45$ и 50. Познато је да за $n \to \infty$ Пуасонова расподела се приближава нормалној. Исто се може очекивати и за очекивање и дисперзију редуковане статистике поретка. Користећи разматрања из симулације оцена λ^* може се модификовати тако да не зависи од параметра λ већ да је функција очекивања и дисперзије стандардизованих статистика поретка из нормалне расподеле. Детаљи ове симулације неће бити приказани али се могу наћи у ([2]).

Нормална модификована оцена из узорка рангираних скупова се може дефинисати као

$$\lambda_N^* = \sum \delta_i X_{(i)}$$

где је

$$\begin{array}{ll} \delta_i = \frac{c_N - b_N \lambda_{i:m}}{\Delta_N \eta_{i:m}} = \frac{1/\eta_{i:m}}{\sum \eta_{j:m}} & \\ c_N = \sum \left(\frac{\gamma_{j:m}^2}{\eta_{j:m}}\right) & b_N = \sum \left(\frac{\gamma_{j:m}}{\eta_{j:m}}\right) = 0 \\ \Delta = \sum \left(\frac{1}{\eta_{j:m}}\right) \sum \left(\frac{\gamma_{j:m}^2}{\eta_{j:m}}\right) & \end{array}$$

При томе су $\alpha_{i:m}(\lambda)$ и $\nu_{i:m}(\lambda)$ замењени са $\lambda_{i:m}$ и $\eta_{i:m}$ који представљају очекивање и дисперзију редукованих статистика поретка нормалне расподеле. Овако дефинисана оцена λ_N^* је функција узорка и константи које не зависе од непознатих параметара, тј λ. За ову оцену израчунато очекивање, пристрасност, дисперзија и средњеквадратна грешка су, редом:

$$E(\lambda_N^*) = \lambda + \sqrt{\lambda} \frac{\sum \alpha_{i:m}(\lambda)/\eta_{i:m}}{\sum 1/\eta_{i:m}}$$

$$Bias(\lambda_N^*) = \sqrt{\lambda} \frac{\sum \alpha_{i:m}(\lambda)/\eta_{i:m}}{\sum 1/\eta_{i:m}}$$

$$D(\lambda_N^*) = \frac{\lambda \sum \nu_{i:m}(\lambda)/\eta_{i:m}^2}{(\sum 1/\eta_{i:m})^2}$$

$$MSE(\lambda_N^*) = \lambda \frac{\sum \nu_{i:m}(\lambda)/\eta_{i:m}^2 - (\sum \alpha_{i:m}(\lambda)/\eta_{i:m})^2}{(\sum 1/\eta_{i:m})^2}$$

Симулацијама су израчунате претходно наведене величине за вредности параметра $\lambda = 1, 3, 5, 10, 20, 40, 60, 80, 100$ и за величину узорка $n = 5, 10, 15, 20$. Пристрасност ове оцене је увек негативна и за фиксиран обим узорка n је приближно константна како λ расте. Вредности дисперзије и средњеквадратне грешке се повећавају како λ расте за фиксни обим узорка.

Да би се упоредила оцена максималне веродостојности из простог случајног узорка $\hat{\lambda}_{SRS}$ и нормална модификована оцена из узорка рангираних скупова λ^*_N потребно је посматрати следеће количнике

$$\frac{D(\hat{\lambda_{SRS}})}{D(\lambda^*_N)} = \frac{(\sum 1/\eta_{j:m})^2}{m\sum \nu_{j:m}(\lambda)/\eta^2_{j:m}}$$

$$\frac{MSE(\hat{\lambda_{SRS}})}{MSE(\lambda^*_N)} = \frac{(\sum 1/\eta_{j:m})^2}{m(\sum \nu_{j:m}(\lambda)/\eta^2_{j:m} - (\sum \alpha_{j:m}(\lambda)/\eta_{j:m})^2}$$

Дисперзија и средњеквадратна грешка оцене $\hat{\lambda}_{SRS}$ су веће од дисперзије и средњеквадратне грешке λ^*_N и посматрани количник се повећава са порастом величине узорка и са порастом параметра Пуасонове расподеле.

Да би се нормална модификована оцене из узорка рангираних скупова λ^*_N упоредила са класичном оценом из узорка рангираних скупова $\widehat{\lambda}_{RSS}$, опет се посматрају исти количници. Дисперзија оцене $\widehat{\lambda}_{RSS}$ је $\frac{\lambda\sum \nu_{i:m}(\lambda)}{m^2}$.

$$\frac{D(\hat{\lambda}_{RSS})}{D(\lambda^*_N)} = \frac{\sum(\nu_{i:m}(\lambda))(\sum 1/\eta_{i:m})^2}{m^2\sum \nu_{j:m}(\lambda)/\eta^2_{j:m}}$$

$$\frac{MSE(\hat{\lambda}_{RSS})}{MSE(\lambda^*_N)} = \frac{\sum(\nu_{i:m(\lambda)})(\sum 1/\eta_{j:m})^2}{m^2(\sum \nu_{i:m}(\lambda)/\eta^2_{i:m} - (\sum \alpha_{i:m}(\lambda)/\eta_{i:m})^2)}$$

Оцена λ_{RSS} има већу дисперзију и посматрани количник расте са порастом параметра Пуасонове расподеле λ. Пораст количника средњеквадратних грешака није уочљив са порастом λ за све величнине узорка.

Може се закључити да се нормална модификована оцена из узорка рангиранух скупова издваја као најквалитетнија за параметар Пуасонове расподеле. Њена пристрасност је мала, а ефикаснија је и од оцене максималне веродостојности из простог случајног узорка и од класичне оцене из узорка рангираних скупова.

Поглавље 4

Степена расподела

Степена расподела је специјалан случај четвoропараметарске бета расподеле. Такође, може се посматрати и као инверзна Паретова расподела. Погодна је за моделирање неких процеса неуспеха. Менкони и Бари ([11]) су упоредили експоненцијалну, логнормалну, Вејбулову и степену расподелу, и доказали да је степена расподела најбоља расподела за проверу поузданости електричних компоненти. Степена расподела се такође користи и у проучавању редова чекања ([4]), затим у економији за одређивање цена, као и времена протеклог од поруџбине до испоруке производа ([28]).

Функција расподеле степене расподеле је

$$F(x) = \left(\frac{x}{d}\right)^a, 0 \leq x \leq d$$

где за параметре важи $a, d > 0$.

Параметри степене расподеле до сада нису оцењивани из узорка рангираних скупова. Циљ овог поглавља је да оцењивање параметара степене расподеле и то у случају када су оба параметра непозната и када је параметар d познат.

4.1 Оцена параметра a када је параметар d познат

У овој секцији биће испитане оцене параметра a, када је параметар d познат. Без губитка општости, претпоставља се да је $d = 1$. Оваква расподела назива се стандардна или једнопараметарска степена расподела, и обележава се са $S(a)$. Функција расподеле је:

$$F(x) = x^a, 0 \leq x \leq 1$$

где важи $a > 0$. Из претходног, густина стандардне степене расподеле је

$$g(x) = ax^{a-1}, 0 \leq x \leq 1$$

Очекивана вредност постоји

$$EX = \frac{a}{a+1},$$

као и дисперзија

$$DX = \frac{a}{(a+1)^2(a+2)}.$$

4.1.1 Прост случајан узорак

Нека X има стандардну степену расподелу са непознатим параметром a, $S(a)$. Изабран је прост случајан узорак обима n, $X_1, X_2, ..., X_n$.

Метода момената

Изједначавањем теоријског момента првог реда EX и узорачког момента првог реда $\overline{X}$ добија се једначина:

$$\frac{a}{a+1} = \overline{X}$$

Решавањем по непознатом параметру a добија се оцена методом момената[1]:

$$\widehat{a}_{mom,SRS} = \frac{\overline{X}}{1-\overline{X}}$$

где је $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$. Према централној граничној теореми, $\sqrt{n}(\overline{X} - \frac{a}{a+1})$ има нормалну расподелу

$$N(0, \frac{a}{(a+1)^2(a+2)}).$$

Подсетимо се делта метода.

[1] енгл. method of moment (MOM)

Теорема 4.1 *Ако $f'(d)$ постоји и $n^b(X_n - d) \xrightarrow{d} X$ за $b > 0$, тада*

$$n^b(f(X_n) - f(d)) \xrightarrow{d} g'(d)X$$

Сада посматрамо функцију $f(y) = \frac{y}{1-y}$. Важи

$$f'(\frac{a}{a+1}) = (a+1)^2 \neq 0,$$

јер је $a > 0$. Према делта методу, $\sqrt{n}(\widehat{a}_{mom,SRS} - a)$ има расподелу

$$N(0, \frac{a(a+1)^2}{(a+2)}).$$

Према претходном, $\widehat{a}_{mom,SRS}$ је асимптотски непристрасна оцена.

Метод максималне веродостојности

Оцена добијена методом максималне веродостојности[2] је

$$\widehat{a}_{mle,SRS} = -\frac{n}{\sum_{i=1}^{n} \ln X_i}.$$

Ако случајна променљива X има степену расподелу $S(a)$, за случајну променљиву $-\ln X$ важи

$$P(-\ln X \leq x) = 1 - P(\ln X \leq -x) = 1 - P(X \leq e^{-x}) = 1 - e^{-ax}$$

Дакле, $-\ln X$ има $\varepsilon(a)$ расподелу. Нека је

$$T = -\sum_{i=1}^{n} \ln X_i$$

Због претходног T има $\Gamma(n, \frac{1}{a})$ расподелу. Сада се може израчунати очекивање $\widehat{a}_{mle,SRS} = \frac{n}{T}$.

$$E(\widehat{a}_{mle,SRS}) = \int_0^\infty \frac{n}{x} \frac{a}{\Gamma(n)} (xa)^{n-1} e^{-ax} dx = \frac{n}{n-1} a$$

Даље,

$$E(\widehat{a}^2_{mle,SRS}) = \int_0^\infty \left(\frac{n}{x}\right)^2 \frac{a}{\Gamma(n)} (xa)^{n-1} e^{-ax} dx = \frac{n^2}{(n-1)(n-2)} a^2$$

[2]енгл. maximum likelihood estimator (MLE)

Из претходног,

$$D(\widehat{a}_{mle,SRS}) = \frac{n^2a^2}{n-1}\left(\frac{1}{n-2} - \frac{1}{n-1}\right) = \frac{n^2}{(n-1)^2(n-2)}a^2.$$

Оцена добијена методом максималне веродостојности је пристрасна, али је асимптотски непристрасна.

Модификована оцена методом максималне веродостојности

Како је оцена добијена методом максималне веродостојности пристрасна, уводи се њена модификација

$$\widehat{a}^2_{mmle,SRS} = \frac{n-1}{n}\widehat{a}_{mle,SRS} = \frac{n-1}{T}$$

која је непристрасна. Дисперзија нове оцене је

$$D(\widehat{a}_{mmle,SRS}) = D(\frac{n-1}{n}\widehat{a}_{mle,SRS}) = \frac{1}{n-2}a^2$$

Како је

$$D(\widehat{a}_{mmle,SRS}) \leq D(\widehat{a}_{mle,SRS}),$$

оцена $\widehat{a}_{mmle,SRS}$ је боља од оцене $\widehat{a}_{mle,SRS}$.

Непристрасна оцена са минималном дисперзијом

Ако X има степену расподелу са параметром a, заједничка густина за $X_1, X_2, ..., X_n$ је

$$f(X_1, X_2, ..., X_n) = \prod_{i=1}^{n} aX_i^{(a-1)} = a^n e^{\ln \prod\limits_{i=1}^{n} X_i^{a-1}} = a^n e^{(a-1)\sum_{i=1}^{n} \ln X_i} = a^n e^{(1-a)T}$$

Према теореми факторизације T је довољна статистика ([7]). Дакле, ниједна друга статистика не садржи више информација о параметру. За сваку функцију $g(T)$ важи да је она непристрасна оцена са минималном дисперзијом за параметар $E(g(T))$ ([19],стр 320). Нека је $g(T) = (n-1)/T$. Тада је $\frac{n-1}{T}$ непристрасна оцена са минималном дисперзијом за

$E(\frac{n-1}{T}) = a$; односно $\hat{a}_{mmle,SRS}$ је непристрасна оцена са минималном дисперзијом за a. Ова оцена ће надаље бити обележена са $\hat{a}_{mvue,SRS}$[3]

Како $-\ln X$ има експоненцијалну расподелу са очекивањем $1/a$ и дисперзијом $1/a^2$, на основу централне граничне теореме $\sqrt{n}(\frac{T}{n-1} - \frac{1}{a})$ има нормалну расподелу

$$N(0, \frac{1}{a^2}).$$

Посматрамо функција $f(y) = \frac{1}{y}$. Како је

$$f'(\frac{1}{a}) = -a^2 \neq 0,$$

применом делта метода добија се да $\sqrt{n}(\widehat{a}_{mvue,SRS} - a)$ има $N(0, a^2)$ расподелу.

Метода медијане

Принцип методе медијане је сличан методи момената. Случајна величина X има степену расподелу са параметром a. Ако је M медијана, тада важи

$$F(M) = M^a = \frac{1}{2}$$

Одатле,

$$M = 2^{-\frac{1}{a}}$$

Ако је $X_{M,SRS}$ медијана из узорка, оцена параметра a добијена методом медијане је

$$\widehat{a}_{M,SRS} = -\frac{\ln 2}{\ln X_{M,SRS}}$$

По теореми о асимптотској расподели узорачких квантила ([24]), важи

$$\sqrt{n}(X_{M,SRS} - M) \xrightarrow{d} N(0, \frac{1}{4f^2(M)})$$

где је $f(x)$ функција густине за једнопараметарску степену расподелу. Према делта методу

$$\sqrt{n}(g(X_{M,SRS}) - g(M)) \xrightarrow{d} N(0, \frac{1}{4f^2(M)}(g'(M))^2)$$

[3]енгл. minimum variance unbiased estimator (MVUE)

где је $g(M)$ произвољна функција таква да $g'(M)$ постоји и различито је од нуле. Уколико се функција $g(x)$ дефинише као $-\frac{\ln 2}{\ln x}$, добија се

$$\sqrt{n}(-\frac{\ln 2}{\ln X_{M,SRS}}+\frac{\ln 2}{\ln M})\xrightarrow{d}N(0,\frac{ln^2 2}{4f^2(M)(M)^2\ln^4(M)})$$

Одатле,

$$\sqrt{n}(\widehat{a}_{M,SRS}-a)\xrightarrow{d}N(0,\frac{\ln^2 2}{4a^2(M)^{2a}\ln^4(M)})$$

Како је $E(\sqrt{n}(\widehat{a}_{M,SRS}-a))=0$, важи да је $E(\widehat{a}_{M,SRS})=a$, тј оцена $\widehat{a}_{M,SRS}$ је асимптотски непристрасна оцена за a.

4.1.2 Узорак рангираних скупова

Нека је изабран узорак рангираних скупова обима $n=km$, где је m величина циклуса, а k број циклуса.

$$\begin{array}{cccc} X_{[1]1} & X_{[2]1} & \dots & X_{[m]1} \\ X_{[1]2} & X_{[2]2} & \dots & X_{[m]2} \\ \dots & \dots & \dots & \dots \\ X_{[1]k} & X_{[2]k} & \dots & X_{[m]k} \end{array}$$

Величина $X_{[i]j}$ представља i-ту по реду изабрану јединицу у j-том циклусу. Нека је, на даље, ради једноставности, $X_{[i]j}$ обележено са X_{ij}. Функција густине за X_{ij} је

$$f_{RSS}(X_{ij})=\frac{m!}{(i-1)!(m-i)!}F(x)^{(i-1)}[1-F(x)]^{(m-i)}f(x) \tag{4.1}$$

где су $f(x)$ и $F(x)$ функција густине и функција расподеле случајне променљиве X ([27]). У случају да случајна променљива X има степену расподелу са познатим параметром $d=1$, функција густине за X_{ij} је

$$f_{RSS}(X_{ij})=\frac{m!}{(i-1)!(m-i)!}X_{ij}^{a(i-1)}[1-X_{ij}^a]^{(m-i)}aX_{ij}^{(a-1)}$$

Метод максималне веродостојности

Функција веродостојности је

$$\begin{aligned} L_{RSS} &= \prod_{i=1}^{m}\prod_{j=1}^{k} f_{RSS}(X_{ij}) \\ &= \prod_{i=1}^{m}\prod_{j=1}^{k} \tfrac{m!}{(i-1)!(m-i)!} X_{ij}^{a(i-1)}[1-X_{ij}^{a}]^{(m-i)} a X_{ij}^{a-1} \\ &= ca^{mk}\prod_{i=1}^{m}\prod_{j=1}^{k} X_{ij}^{a(i-1)}[1-X_{ij}^{a}]^{(m-i)} X_{ij}^{a-1} \end{aligned}$$

где је c константа. Даље,

$$\ln L_{RSS} = \ln c + mk\ln a + \sum_{i=1}^{m}\sum_{j=1}^{k}\big((m-i)\ln(1-X_{ij}^{a}) + (ai-1)\ln X_{ij}\big)$$

Да би се добила оцена методом максималне веродостојности посматрамо израз

$$\frac{\partial \ln L_{RSS}}{\partial a} = \frac{mk}{a} + \sum_{i=1}^{m}\sum_{j=1}^{k}(\frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}} + i\ln X_{ij})$$

Кад $a \to 0$, $\frac{\partial \ln L_{RSS}}{\partial a} \to \infty$, а кад $a \to \infty$, $\frac{\partial \ln L_{RSS}}{\partial a} < 0$. Потребно је још доказати да је $\frac{\partial \ln L_{RSS}}{\partial a}$ монотоно опадајућа функција. Како је $0 \leq x \leq 1$

$$\begin{aligned} \frac{\partial^2 \ln L_{RSS}}{\partial a^2} &= -\tfrac{mk}{a^2} + \sum_{i=1}^{m}\sum_{j=1}^{k} \tfrac{-(m-i)\ln X_{ij}[X_{ij}^{a}\ln X_{ij}(1-X_{ij}^{a})+X_{ij}^{2a}\ln X_{ij}]}{(1-X_{ij}^{a})^2} \\ &= -\tfrac{mk}{a^2} + \sum_{i=1}^{m}\sum_{j=1}^{k} \tfrac{-(m-i)\ln X_{ij}[X_{ij}^{a}\ln X_{ij}]}{(1-X_{ij}^{a})^2} < 0 \end{aligned}$$

Дакле, $\frac{\partial \ln L_{RSS}}{\partial a} = 0$ има јединствено решење и оно представља оцену параметра a код узорка рангираних скупова методом максималне веродостојности.

Модификована метода максималне веродостојности

Јединствено решење методом максималне веродостојности постоји али га није могуће издвојити ненумеричким методама. Зато се користи

модификована метода максималне веродостојности. У раду [12] ова метода је коришћена за оцењивање параметара нормалне расподеле, а у раду [1] за оцењивање параметара Паретове расподеле.
Једначину

$$\begin{aligned} \frac{\partial \ln L_{RSS}}{\partial a} &= \frac{mk}{a} + \sum_{i=1}^{m}\sum_{j=1}^{k}\left(\frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}} + i\ln X_{ij}\right) \\ &= \frac{mk}{a} + \sum_{i=1}^{m-1}\sum_{j=1}^{k}\left(\frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}} + i\ln X_{ij}\right) + m\sum_{j=1}^{k}\ln X_{mj} \\ &= 0 \end{aligned}$$

није могуће решити због израза

$$\sum_{i=1}^{m}\sum_{j=1}^{k}\frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}}$$

Из тог разлога је израз $\frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}}$ замењен је својим очекивањем.

$$\begin{aligned} E\left(\frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}}\right) &= \int_0^1 \frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}} f_{RSS}(X_{ij})dx \\ &= \int_0^1 \frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}} \frac{m!}{(i-1)!(m-i)!} X_{ij}^{a(i-1)}[1-X_{ij}^{a}]^{(m-i)} a X_{ij}^{(a-1)}dx \\ &= -(m-i)\frac{m!}{(i-1)!(m-i)!} a \int_0^1 X_{ij}^{a+ai-1}\ln X_{ij}(1-X_{ij}^{a})^{m-i-1}dx \end{aligned}$$

Уколико се искористи смена $\ln X_{ij} = t$, добија се

$$\begin{aligned} E\left(\frac{-(m-i)X_{ij}^{a}\ln X_{ij}}{1-X_{ij}^{a}}\right) &= \frac{m!}{(i-1)!(m-i-1)!} a \int_0^\infty e^{ta(i+1)} t(1-e^{at})^{m-i-1}dt \\ &= \frac{m!}{(i-1)!(m-i-1)!} a \int_0^\infty e^{ta(i+1)} t \sum_{s=0}^{m-i-1}\binom{m-i-1}{s}(-e^{at})^{s}dt \\ &= \frac{m!}{(i-1)!(m-i-1)!a}\sum_{s=0}^{m-i-1}\binom{m-i-1}{s}\frac{(-1)^s}{(1+i+s)^2} \end{aligned}$$

Из једначине

$$\frac{mk}{a}+\sum_{i=1}^{m-1}\sum_{j=1}^{k}\left(i\ln X_{ij}+\frac{m!}{(i-1)!(m-i-1)!a}\sum_{s=0}^{m-i-1}\binom{m-i-1}{s}\frac{(-1)^s}{(1+i+s)^2}\right)+m\sum_{j=1}^{k}\ln X_{mj}=0$$

добија се оцена за a

$$\widehat{a}_{mmle,RSS} = \frac{k\left(m + \sum_{i=1}^{m-1}\frac{m!}{(i-1)!(m-i-1)!}\sum_{s=0}^{m-i-1}\binom{m-i-1}{s}\frac{(-1)^s}{(1+i+s)^2}\right)}{-\sum_{i=1}^{m}\sum_{j=1}^{k} i\ln X_{ij}}$$

Оцене по узору

Оцене по узору[4] добијене су модификацијом оцена из простог случајног узорка. Наиме, статистике израчунате из простог случајног узорка, замењене су статистикама из узорка рангираних скупова.

Оцена по узору на оцену $\widehat{a}_{mom,SRS}$ добијена је тако што је статистика из простог случајног узорка $\overline{X}_{SRS} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ замењена статистиком из узорка рангираних скупова $\overline{X}_{RSS} = \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} X_{[i]j}/km$. Добијена оцена је

$$\widehat{a}_{mom,ah} = \frac{\overline{X}_{RSS}}{1-\overline{X}_{RSS}}$$

На сличан начин је добијена и оцена по узору на оцену добијену методом медијане

$$\widehat{a}_{M,ah} = -\frac{\ln 2}{\ln M_{RSS}}$$

где је M_{RSS} медијана узорка рангираних скупова.

Оцена по узору може бити изведена и из модификоване оцене максималне веродостојности, $\widehat{a}_{mmle,SRS}$. Добијена оцена је

$$\widehat{a}_{mmle,ah} = \frac{p}{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} \ln X_{ij}}.$$

где је p константа таква да је $\widehat{a}_{mmle,ah}$ непристрасна оцена параметра a.

4.1.3 Симулација и упоређивање оцена

Оцене параметра a из простог случајног узорка су: $\hat{a}_{mom,SRS}$, $\hat{a}_{mle,SRS}$, $\hat{a}_{mvue,SRS}$ и $\hat{a}_{M,SRS}$; а оцене параметра a добијене из узорка рангираних скупова су $\hat{a}_{mmle,RSS}$, $\hat{a}_{mom,ah}$, $\hat{a}_{M,ah}$ и $\hat{a}_{mmle,ah}$. У овом поглављу посматране су особине ових оцена као што су пристрасност и средњеквадратна грешка. Вредности ових параметара су добијене помоћу Монте Карло симулације са 10000 итерација. Симулација је вршена у програмском језику R.

[4] Индекс ah у називу оцене потиче од израза ad hoc

Најпре је одређена пристрасност свих добијених оцена. Ради упоређивања пристрасности за различите вредности параметра a, посматране су релативне, а не апсолутне вредности. Тачније, посматрана је величина

$$\frac{|E(\hat{a}) - a|}{a}.$$

Релативна пристрасност је посматрана за три вредности параметра $a = 1, 3, 5$. При том су за сваку вредност параметра посматране две величине узорка $n = 12$ и $n = 24$. За сваку величину узорка су посматране различите комбинације параметара m и k. За $n = 12$, коришћене су комбинације $(m, k) = ((2, 6), (3, 4), (4, 3))$, а за $n = 24$, коришћене су вредности параметара $(m, k) = ((2, 12), (3, 8), (4, 6))$. Добијене вредности су дате у табели (4.1)

Према теоријским разматрањима, а везано за оцену из простог случајног узорка, једино је оцена $\hat{a}_{mvue,SRS}$ непристрасна. Остале оцене: $\hat{a}_{mom,SRS}$, $\hat{a}_{mle,SRS}$ и $\hat{a}_{M,SRS}$ су асимптотски непристрасне, али нису непристрасне. Сходно томе, може се видети да ја њихова пристрасност мања за $n = 24$, него за $n = 12$ за све вредности параметра a. Пристрасност је релативно мала за све оцене из простог случајног узорка, осим за $\hat{a}_{M,SRS}$ где је пристрасност за $n = 12$ преко 10%.

Што се тиче оцена из узорка рангираних скупова њихова пристрасност је углавном мања од пристрасности оцена из простог случајног узорка осим за $\hat{a}_{mvue,SRS}$. Пристрасност је мања када је узорак већи и износи $1-3\%$. Свеукупно, из узорка рангираних скупова, најмању пристрасност има $\hat{a}_{mom,ah}$. Интересантно је да се може уочити благо смањење пристрасности како расте величина циклуса. Ово је неочекивано с обзиром да је генерална препорука да се пре бира мањи циклус са већим бројем понављања него већи циклус са мањим бројем понављања због грешке у рангирању. Пристрасност се углавном не мења у односу на промену параметра a.

a	n	m	k	$\hat{a}_{mom,SRS}$	$\hat{a}_{mle,SRS}$	$\hat{a}_{mvue,SRS}$	$\hat{a}_{M,SRS}$	$\hat{a}_{mmle,RSS}$	$\hat{a}_{mom,ah}$	$\hat{a}_{M,ah}$	$\hat{a}_{mmle,ah}$
1	12	2	6	6,0305	8,8809	0,1926	11,4259	6,5903	4,2794	8,9118	6,7189
1	12	3	4	6,6887	9,4480	0,3273	12,0660	4,8929	2,8905	6,7761	5,1530
1	12	4	3	6,6158	9,4747	0,3518	11,7028	3,9592	2,3800	5,5886	4,4515
1	24	2	12	2,6198	3,9537	0,3777	5,4730	3,1024	1,8950	4,2699	3,1981
1	24	3	8	2,4208	3,9549	0,3766	5,0189	2,4640	1,4674	3,4889	2,4265
1	24	4	6	3,0585	4,3987	0,0487	5,8548	2,2139	1,3285	3,2708	2,3716
3	12	2	6	6,9166	8,7132	0,3462	9,7023	7,0362	5,5401	7,9292	7,0364
3	12	3	4	7,8867	9,5369	0,4088	11,5814	4,8244	3,5601	5,2673	5,1301
3	12	4	3	7,9295	9,6817	0,5416	11,0738	3,9423	3,0297	4,7902	4,4604
3	24	2	12	3,7466	4,5389	0,1831	5,7663	2,8944	2,3083	3,7213	2,9581
3	24	3	8	3,2648	4,1880	0,1531	4,9224	2,2764	1,7444	3,0167	2,3823
3	24	4	6	3,3078	4,1030	0,2346	5,3642	1,8842	1,4243	2,7468	2,0037
5	12	2	6	7,9413	9,1202	0,0269	10,5463	6,2018	5,2220	6,8774	6,3895
5	12	3	4	7,6501	8,9165	0,1598	9,9764	4,7311	3,9892	5,2666	5,0215
5	12	4	3	7,8043	9,0024	0,0812	10,5976	3,7971	3,3056	4,5674	4,2856
5	24	2	12	3,9506	4,5759	0,2186	5,1441	3,3323	2,9155	4,2399	3,3574
5	24	3	8	3,8740	4,5099	0,1553	5,6068	2,1881	1,8854	3,1642	2,3178
5	24	4	6	4,0160	4,5986	0,2403	5,5082	1,7985	1,4781	2,6804	1,8681

Табела 4.1: Пристрасност оцена параметра a

Следеће су израчунате средњеквадратне грешке за добијене оцене. Међусобна упоређивања оцена су вршена помоћу ефикасности која је дефинисана као количник средњеквадратних грешака

$$efikasnost = \frac{MSE(\hat{a}_1)}{MSE(\hat{a}_2)}.$$

Симулација је вршена за исте вредности параметара a, n, m и k као и код пристрасности. У табели 4.2 дата је ефикасност оцене $\hat{a}_{mom,SRS}$ у односу на остале оцене.

Из дате табеле може се видети да је ефикасност оцене углавном већа од 1, тј. да оцена $\hat{a}_{mom,SRS}$ има релативно велику средњеквадратну грешку у односу на друге оцене. Једино има мању средњеквадратну грешку у односу на $\hat{a}_{M,SRS}$. С обзиром да оцена $\hat{a}_{M,SRS}$ има и највећу пристрасност, дефинитивно није добар избор за оцену параметра a. Битно је приметити да је вредност посматраног количника већа када се посматрају оцене из узорка рангираних скупова, у односу кад се посматрају оцене из простог случајног узорка. Другим речима, оцене из узорка рангираних скупова имају мање средњеквадратне грешке него оцене из простог случајног узорка. Од оцена из узорка рангираних скупова највећу средњеквадратну грешку има оцена $\hat{a}_{M,ah}$, док су средњеквадратне грешке осталих оцена приближно исте. Код узорка рангираних скупова посматрани количник расте са порастом величине циклуса. Не примећује се утицај промене вредности параметра a на ефикасност.

С обзиром да оцена $\hat{a}_{mvue,SRS}$ има најмању пристрасност, у табели 4.3 приказана је ефикасност ове оцене у односу на остале оцене.

У табели 4.3 може се уочити да оцена $\hat{a}_{mvue,SRS}$ има мању средњеквадратну грешку у односу на остале оцене из простог случајног узорка. Међутим, када се ова оцена упореди са оценама из узорка рангираних скупова, она има мању средњеквадратну грешку само у односу на $\hat{a}_{M,ah}$. У односу на остале оцене из узорка рангираних скупова:$\hat{a}_{mmle,RSS}$, $\hat{a}_{mom,ah}$ и $\hat{a}_{mmle,ah}$ оцена $\hat{a}_{mvue,SRS}$ има већу средњеквадратну грешку и до приближно 3 пута. Посматрани количник је велик нарочито када је величина циклуса велика.

a	n	m	k	$\hat{a}_{mle,SRS}$	$\hat{a}_{mvue,SRS}$	$\hat{a}_{M,SRS}$	$\hat{a}_{mmle,RSS}$	$\hat{a}_{mom,ah}$	$\hat{a}_{M,ah}$	$\hat{a}_{mmle,ah}$
1	12	2	6	1,1355	1,4459	0,5439	1,8604	1,6781	0,9427	1,7902
1	12	3	4	1,1551	1,4665	0,5255	2,4878	2,2766	1,1875	2,3349
1	12	4	3	1,1475	1,4633	0,5492	3,1033	2,9148	1,4059	2,7983
1	24	2	12	1,2292	1,3905	0,5903	1,7300	1,4785	0,7915	1,6903
1	24	3	8	1,2375	1,3983	0,5882	2,4845	2,1936	1,1254	2,2940
1	24	4	6	1,2352	1,3947	0,5840	3,0186	2,7193	1,2852	2,7184
3	12	2	6	1,0033	1,2723	0,5122	1,5981	1,6229	0,7844	1,5509
3	12	3	4	1,0021	1,2751	0,4953	2,1358	2,1615	1,0142	1,9726
3	12	4	3	0,9992	1,2715	0,5013	2,7472	2,7922	1,2930	2,4848
3	24	2	12	1,0311	1,1727	0,4847	1,4582	1,4533	0,7005	1,4120
3	24	3	8	1,0378	1,1655	0,4905	1,9389	1,9628	0,8793	1,8133
3	24	4	6	1,0228	1,1561	0,4978	2,4904	2,5174	1,0737	2,2576
5	12	2	6	0,9868	1,2605	0,4981	1,5737	1,5814	0,7649	1,5285
5	12	3	4	0,9935	1,2635	0,4783	2,2587	2,2637	1,0934	2,1159
5	12	4	3	0,9934	1,2618	0,4810	2,6741	2,6065	1,2272	2,3827
5	24	2	12	1,0033	1,1284	0,4852	1,5057	1,4933	0,6960	1,4467
5	24	3	8	1,0045	1,1342	0,4854	1,9375	1,8914	0,8797	1,7820
5	24	4	6	1,0067	1,1305	0,4935	2,3801	2,3692	1,0208	2,1843

Табела 4.2: Ефикасност оцене $\hat{a}_{mom,SRS}$ у односу на остале оцене

a	n	m	k	$\hat{a}_{mom,SRS}$	$\hat{a}_{mle,SRS}$	$\hat{a}_{M,SRS}$	$\hat{a}_{mmle,RSS}$	$\hat{a}_{mom,ah}$	$\hat{a}_{M,ah}$	$\hat{a}_{mmle,ah}$
1	12	2	6	0,6794	0,7868	0,3682	1,2352	1,1144	0,5970	1,1859
1	12	3	4	0,6754	0,7886	0,3704	1,7694	1,6308	0,8134	1,6222
1	12	4	3	0,6792	0,7873	0,3453	2,2599	2,1004	1,0145	2,0392
1	24	2	12	0,7248	0,8822	0,4342	1,3619	1,1494	0,6301	1,3098
1	24	3	8	0,7214	0,8840	0,4328	1,7566	1,5661	0,8079	1,6475
1	24	4	6	0,7194	0,8870	0,4268	2,1348	1,9282	0,9240	1,9170
3	12	2	6	0,7788	0,7802	0,3850	1,2454	1,2524	0,6230	1,1933
3	12	3	4	0,7853	0,7876	0,3912	1,7486	1,7668	0,8047	1,6269
3	12	4	3	0,7867	0,7893	0,3732	2,0164	2,0411	0,9510	1,8209
3	24	2	12	0,8619	0,8899	0,4257	1,3277	1,3299	0,6189	1,2895
3	24	3	8	0,8633	0,8886	0,4334	1,7598	1,7581	0,7809	1,6273
3	24	4	6	0,8556	0,8803	0,4101	2,2168	2,2515	0,9229	2,0344
5	12	2	6	0,7934	0,7869	0,3878	1,2424	1,2521	0,5954	1,2058
5	12	3	4	0,7872	0,7815	0,3783	1,6165	1,6143	0,8090	1,5075
5	12	4	3	0,7930	0,7853	0,3967	2,1810	2,1757	0,9867	1,9988
5	24	2	12	0,8770	0,8803	0,4439	1,3070	1,3016	0,6027	1,2595
5	24	3	8	0,8732	0,8803	0,4047	1,6774	1,6698	0,7312	1,5800
5	24	4	6	0,8771	0,8810	0,4186	2,1441	2,1190	0,8935	1,9462

Табела 4.3: Ефикасност оцене $\hat{a}_{mvue,SRS}$ у односу на остале оцене

4.2 Оцена параметара када су оба параметра непозната

Густина степене расподеле $S(a,d)$ је

$$g(x) = \frac{a}{d^a} x^{a-1}$$

где је $0 \leq x \leq d$ и $a, d > 0$. Очекивање и дисперзија степене расподеле су

$$E(X) = \frac{da}{a+1},$$

и

$$D(X) = \frac{d^2 a}{(a+2)(a+1)^2}.$$

4.2.1 Прост случајан узорак

Нека је $X_1, ..., X_n$ прост случајан узорак из двопараметарске степене расподеле $S(a,d)$. Параметри a и d су оцењени различитим методама.

Метода момената

Узорачки моменат k-тог реда је

$$m_k = \frac{\sum\limits_{i=1}^{n} X_i^k}{n}.$$

Узорачки моменат провог реда је заправо $\bar{X}_{SRS}$, а узорачки моменат другог реда је једнак $S_{SRS}^2 + \bar{X}_{SRS}^2$. Теоријски моменти првог и другог реда степене расподеле су:

$$EX = \int\limits_0^d x \frac{a}{d^a} x^{a-1} dx = \frac{da}{a+1}$$

и

$$EX^2 = \int\limits_0^d x^2 \frac{a}{d^a} x^{a-1} dx = \frac{d^2 a}{a+2}$$

Изједначавањем теоријских и узорачких момената добијају се оцене параметара a и d.

$$\hat{a}_{mom,SRS} = -1 + \sqrt{1 + \frac{\bar{X}^2_{SRS}}{S^2_{SRS}}}$$

$$\hat{d}_{mom,SRS} = \frac{\bar{X}_{SRS}(\hat{a}_{mom,SRS} + 1)}{\hat{a}_{mom,SRS}}$$

Комбинована оцена

Оцене параметара се могу наћи тако што се параметар a оцени методом момената, а параметар d помоћу статистике поретка. Како је код степене расподеле $x < d$, оцену параметра d налазимо из максималне статистике поретка. Функција расподеле максимума $X_{(n)}$ је

$$\begin{aligned} F_{X_{(n)}}(x) &= P(X_{(n)} \leq x) = P(X_{(1)} \leq x, X_{(2)} \leq x, ..., X_{(n)} \leq x) \\ &= (P(X \leq x))^n = \left(\tfrac{x}{d}\right)^{an} \end{aligned}$$

Одатле, густина је

$$f_{X_{(n)}}(x) = \frac{anx^{an-1}}{d^{an}},$$

а очекивање

$$EX_{(n)} = \int\limits_0^d \frac{anx^{an-1}}{d^{an}} x dx = \frac{dan}{an+1}.$$

Када се добијени израз изједначи са максималном вредношћу из узорка $X_{(n)}$, добија се

$$\hat{d}_{ko,SRS} = \frac{X_{(n)}(n\hat{a}_{ko,SRS} + 1)}{n\hat{a}_{ko,SRS}} \tag{4.2}$$

Оцена $\hat{a}_{ko,SRS}$ добија се када се изједначе теоријски моменат првог реда $EX = \frac{da}{a+1}$ и узорачки моменат првог реда $\overline{X}_{SRS} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$

$$\hat{a}_{ko,SRS} = \frac{\overline{X}_{SRS}}{\hat{d}_{ko,SRS} - \overline{X}_{SRS}}$$

. Према томе, из чињенице да је $a > 0$ и једнакости

$$\hat{a}_{ko,SRS} = \frac{\overline{X}_{SRS}}{\frac{X_{(n)}(n\hat{a}_{ko,SRS}+1)}{n\hat{a}_{ko,SRS}} - \overline{X}_{SRS}},$$

добија се оцена параметра a:

$$\hat{a}_{ko,SRS} = \frac{n\overline{X}_{SRS} - X_{(n)}}{n(X_{(n)} - \overline{X}_{SRS})} \tag{4.3}$$

Дакле, оцене параметара d и a дате су са (4.2) и (4.3).

Метод максималне веродостојности

Функција веродостојности је

$$\begin{aligned} L(x,a,d) &= f(x_1)\cdot f(x_2)\cdot \ldots \cdot f(x_n) \\ &= \tfrac{ax_1^{a-1}}{d^a}\cdot \tfrac{ax_2^{a-1}}{d^a}\cdot \ldots \cdot \tfrac{ax_n^{a-1}}{d^a}\cdot I\{x_1 \le d\}\cdot I\{x_2 \le d\}\cdot \ldots \cdot I\{x_n \le d\} \\ &= \tfrac{ax_1^{a-1}}{d^a}\cdot \tfrac{ax_2^{a-1}}{d^a}\cdot \ldots \cdot \tfrac{ax_n^{a-1}}{d^a}\cdot I\{x_{(n)} \le d\} \end{aligned}$$

Како је $L(x,a,d)$ опадајућа функција по d, њена максимална вредност је за минималну вредност параметра d. Даље, функција $L(x,a,d)$ је позитивна уколико је индикатор једнак јединици, односно ако је $d \ge x_{(n)}$. Према томе, оцена параметра d методом максималне веродостојности је

$$\hat{d}_{mle,SRS} = x_{(n)}$$

Функција веродостојности сада има облик

$$L(x,a) = \frac{ax_1^{a-1}}{x_{(n)}^a}\cdot \frac{ax_2^{a-1}}{x_{(n)}^a}\cdot \ldots \cdot \frac{ax_n^{a-1}}{x_{(n)}^a}$$

Одатле,

$$\ln L(X,a) = n\ln a - an\ln x_{(n)} + (a-1)\sum_{i=1}^{n}\ln x_i.$$

Из израза

$$\frac{\partial \ln L}{\partial a} = \frac{n}{a} - n\ln x_{(n)} + \sum_{i=1}^{n}\ln x_i = 0$$

добија се оцена параметра a методом максималне веродостојности

$$\hat{a}_{mle,SRS} = \frac{n}{n\ln x_{(n)} - \sum\limits_{i=1}^{n}\ln x_i}$$

Оцена методом квантила

Оцене параметара степене расподеле методом квантила дате су у ([6]). Као што је познато, функција расподеле је

$$F(X_i) = \left(\frac{X_i}{d}\right)^a$$

Одавде,

$$X_i = dF(X_i)^{\frac{1}{a}}$$

Нека су са X_{p_i} и p_i обележени узорачки квантил и одговарајућа вероватноћа. За два изабрана квантила важе следеће једнакости:

$$X_{p_1} = d \cdot p_1^{\frac{1}{a}}$$

и

$$X_{p_2} = d \cdot p_2^{\frac{1}{a}}$$

Дељењем ове две једнакости добија се

$$\frac{X_{p_1}}{X_{p_2}} = \left(\frac{p_1}{p_2}\right)^{\frac{1}{a}}.$$

Одавде,

$$\hat{a}_{qt,SRS} = \frac{\ln\left(\frac{p_1}{p_2}\right)}{\ln\frac{X_{p_1}}{X_{p_2}}}$$

и

$$\hat{d}_{qt,SRS} = \frac{X_{p_1}}{p_1^{\frac{1}{\hat{a}_{qt,SRS}}}}$$

У раду ([6]) коришћене су вероватноће: $(0,1;0,9)$,$(0,25;0,75)$, $(0,05;0,95)$ и $(0,08;0,92)$. Најбоље оцене су добијене са вероватноћама $(0,05;0,95)$.

Оцена методом најмањих квадрата

Зака и Актар ([5]) дали су оцене параметара степене расподеле методом најмањих квадрата. Као и код методе квантила из функције расподеле се добија

$$X_i = dF(X_i)^{\frac{1}{a}}$$

Да би се добила линеарна веза између параметара, горњи израз се логаритмује.

$$\ln X_i = \ln d + \frac{1}{a}\ln F(X_i)$$

Ради поједностављивања, уведене су следеће ознаке

$$Z_i = \ln X_i$$

$$\alpha = \ln d$$
$$\beta = \frac{1}{a}$$
$$Y_i = \ln F(X_i)$$

Горања једнакост се сада може записати на следећи начин:

$$Z_i = \alpha + \beta Y_i$$

Оцене параметара се добијају помоћу једнакости

$$S(\alpha,\beta) = \sum_{i=1}^{n}(Z_i - (\alpha + \beta Y_i))^2$$

Диференцирањем $S(\alpha,\beta)$ по α и β и изједначвањем са нулом, добија се

$$\sum_{i=1}^{n} Z_i = n\alpha + \beta\sum_{i=1}^{n} Y_i$$

$$\sum_{i=1}^{n} Z_i Y_i = \alpha\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} Y_i^2$$

Решавањем претходних једначина добијају се оцене за α и β

$$\beta = \frac{\sum_{i=1}^{n}\ln F(X_i)\ln X_i - \frac{(\sum_{i=1}^{n}\ln F(X_i))(\sum_{i=1}^{n}\ln X_i)}{n}}{\sum_{i=1}^{n}(\ln F(X_i))^2 - \frac{(\sum_{i=1}^{n}\ln F(X_i))^2}{n}}$$

$$\alpha = \frac{\sum_{i=1}^{n}\ln X_i}{n} - \beta\frac{\sum_{i=1}^{n}\ln F(X_i)}{n}$$

Вредност функција расподеле може се апроксимирати помоћу Бенардове апроксимације ([10]):

$$F(X_i) = F(X_{(j)}) = \frac{j - 0,3}{n + 0,4}$$

где је X_i, j-ти по реду члан у сортираном узорку. Како је $\alpha = \ln d$ и $\beta = 1/a$, оцене параметара a и d су

$$\hat{a}_{lsm,SRS}{}^5 = \frac{\sum_{i=1}^{n} (\ln F(X_i))^2 - \frac{(\sum_{i=1}^{n} \ln F(X_i))^2}{n}}{\sum_{i=1}^{n} \ln F(X_i) \ln X_i - \frac{(\sum_{i=1}^{n} \ln F(X_i))(\sum_{i=1}^{n} \ln X_i)}{n}}$$

$$\hat{d}_{lsm,SRS} = e^{\frac{\sum_{i=1}^{n} \ln X_i}{n} - \frac{\sum_{i=1}^{n} \ln F(X_i)}{n \hat{a}_{lsm,SRS}}}$$

Модификована непристрасна оцена са најмањом дисперзијом

Уколико X има степену расподелу, за случајну промењливу $Y = -\ln X$ важи

$$\begin{aligned} P(Y \leq y) &= P(-\ln X \leq y) \\ &= P(\ln X \geq -y) \\ &= 1 - P(\ln X \leq -y) \\ &= 1 - P(X \leq e^{-y}) \\ &= 1 - \left(\frac{e^{-y}}{d}\right)^a \\ &= 1 - e^{-a(y - \ln \frac{1}{d})} \end{aligned},$$

Дакле, $Y = -\ln X$ има експоненцијалну расподелу са параметрима $a' = \frac{1}{a}$ и $b' = \ln \frac{1}{d}$. Непристрасне оцене са најмањом дисперзијом параметра a' и b' експоненцијалне расподеле су

$$\hat{a}' = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - Y_{(1)})$$

и

$$\hat{b}' = Y_{(1)} + \frac{1}{n(n-1)} \sum_{i=1}^{n} (Y_i - Y_{(1)}),$$

[5] енгл. least squares method (LSM)

где је $(Y_1, ..., Y_n) = (-\ln X_1, ..., -\ln X_n)$ и $Y_{(1)}$ је минимум скупа $(Y_1, ..., Y_n)$([8]). Одатле, оцене за параметре a и d из степене расподеле су

$$\hat{a}_{mmvue,SRS}{}^{6} = \frac{n-1}{\sum_{i=1}^{n}(Y_i - Y_{(1)})}$$

и

$$\hat{d}_{mmvue,SRS} = e^{-(Y_{(1)} + \frac{1}{n(n-1)}\sum_{i=1}^{n}(Y_i - Y_{(1)}))}$$

4.2.2 Узорак рангираних скупова

Нека је изабран узорак методом рангираних скупова из двопараметарске степене расподеле $S(a, d)$ обима $n = km$, и нека важе исте ознаке као и до сада. Из формуле (4.1), добија се функција густине за X_{ij}

$$g_{RSS}(X_{ij}) = \frac{m!}{(i-1)!(m-i)!}\left(\frac{X_{ij}}{d}\right)^{a(i-1)}\left[1 - \left(\frac{X_{ij}}{d}\right)^{a}\right]^{(m-i)}\frac{aX_{ij}^{a-1}}{d^a}$$

Метод максималне веродостојности

Функција веродостојности је

$$\begin{aligned} L_{RSS}(X, a, d) &= \prod_{i=1}^{m}\prod_{j=1}^{k} g_{RSS}(X_{ij}) \\ &= ca^{km}\prod_{i=1}^{m}\prod_{j=1}^{k}\frac{X_{ij}^{ai-1}}{d^{ai}}\left(1 - \left(\frac{X_{ij}}{d}\right)^{a}\right)^{(m-i)} I\{X_{11} \le d\}I\{X_{12} \le d\}...I\{X_{mk} \le d\} \\ &= ca^{km}\prod_{i=1}^{m}\prod_{j=1}^{k}\frac{X_{ij}^{ai-1}}{d^{ai}}\left(1 - \left(\frac{X_{ij}}{d}\right)^{a}\right)^{(m-i)} I\{\max_{i,j}\{X_{ij}\} \le d\} \end{aligned}$$

где је c константа. Слично методи максималне веродостојности код простог случајног узорка, $L_{RSS}(X, a, d)$ је опадајућа функција по d. Стога, функција веродостојности се максимизира за минималну вредност параметра d. Због индикатора, функција $L_{RSS}(X, a, d)$ је позитивна уколико је $d \ge \max_{i,j}\{X_{ij}\}$. Дакле, оцена параметра d методом максималне веродостојности код узорка рангираних скупова је

$$\hat{d}_{mle,RSS} = \max_{i,j}\{X_{ij}\}.$$

[6]енгл. modified minimum variance unbiased estimator (MMVUE)

Сада функција веродостојности има облик

$$L_{RSS}(X,a) = ca^{km}\prod_{i=1}^{m}\prod_{j=1}^{k}\frac{X_{ij}^{ai-1}}{\max\limits_{i,j}\{X_{ij}\}^{ai}}\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\right]^{(m-i)}$$

Оцена параметра a методом максималне веродостојности је решење једначине

$$\begin{aligned}\frac{\partial \ln L_{RSS}(X,a)}{\partial a} &= \frac{km}{a}+\sum_{i=1}^{m}\sum_{j=1}^{k} i\ln X_{ij}-\sum_{i=1}^{m}\sum_{j=1}^{k} i\ln \max_{i,j}\{X_{ij}\}\\ &-\sum_{i=1}^{m-1}\sum_{j=1}^{k}(m-i)\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\ln\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\right]^{-1}\\ &=0\end{aligned}$$

Међутим, ову једначину није могуће решити нумеричким методама.

Модификована метода максималне веродостојности

Како није могуће наћи оцену параметара a из функције веродостојности користи се модификована метода максималне веродостојности. Једначину $\frac{\partial \ln L_{RSS}(X,a)}{\partial a}=0$ није могуће решити због израза

$$-(m-i)\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\ln\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\right]^{-1}$$

Зато је он замењен својим очекивањем.

$$\begin{aligned}&E\left(-(m-i)\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\ln\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\right]^{-1}\right)\\ &=\int_{0}^{\max\limits_{i,j}\{X_{ij}\}}-(m-i)\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\ln\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\right]^{-1} g_{RSS}(X_{ij})dx\\ &=\int_{0}^{\max\limits_{i,j}\{X_{ij}\}}-(m-i)\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\ln\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\right]^{-1}\\ &\frac{m!}{(i-1)!(m-i)!}\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a(i-1)}\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\right]^{(m-i)}\frac{aX_{ij}^{a-1}}{\left(\max\limits_{i,j}\{X_{ij}\}\right)^{a}}dx\\ &=\frac{(m-i)m!}{(i-1)!(m-i)!}a\int_{0}^{\max\limits_{i,j}\{X_{ij}\}}-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a(i+1)}\ln\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\\ &\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^{a}\right]^{m-i-1}\frac{1}{X_{ij}}dx\end{aligned}$$

Сменом $t = \ln \frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}$, добија се

$$E\left(-(m-i)\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^a \ln\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)\left[1-\left(\frac{X_{ij}}{\max\limits_{i,j}\{X_{ij}\}}\right)^a\right]^{-1}\right)$$
$$= \frac{m!}{(i-1)!(m-i-1)!} a \int\limits_0^\infty e^{ta(i+1)} t (1-e^{at})^{m-i-1} dt$$
$$= \frac{m!}{(i-1)!(m-i-1)!} a \int\limits_0^\infty e^{ta(i+1)} t \sum_{s=0}^{m-i-1} \binom{m-i-1}{s} (-e^{at})^s dt$$
$$= \frac{m!}{(i-1)!(m-i-1)!a} \sum_{s=0}^{m-i-1} \binom{m-i-1}{s} \frac{(-1)^s}{(1+i+s)^2}$$

Заменом добијене вредности у једнакости $\frac{\partial \ln L_{RSS}(X,a)}{\partial a} = 0$ добија се оцена параметра a

$$\hat{a}_{mmle,RSS} = \frac{k(m + \sum\limits_{i=1}^{m-1} \frac{m!}{(i-1)!(m-i-1)!} \sum\limits_{s=0}^{m-i-1} \binom{m-i-1}{s} \frac{(-1)^s}{(1+i+s)^2})}{\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{k} i \ln \frac{\max\limits_{i,j} \hat{\{X_{ij}\}}_{mmle,RSS}}{X_{ij}}}$$

Оцене по узору

Оцене по узору добијене су модификацијом већ изведених оцена тако што су статистике из простог случајног узорка замењене статистикама из узорка рангираних скупова.

Оцене по узору на оцене добијене методом момената, $\hat{a}_{mom,SRS}$ и $\hat{d}_{mom,SRS}$, су

$$\hat{a}_{mom,ah} = -1 + \sqrt{1 + \frac{\bar{X}^2_{RSS}}{S^2_{RSS}}}$$

$$\hat{d}_{mom,ah} = \frac{\bar{X}_{RSS}(\hat{a}_{mom,ah} + 1)}{\hat{a}_{mom,ah}}$$

Оцене по узору на оцене добијене комбинованом методом, $\hat{a}_{ko,SRS}$ и $\hat{d}_{ko,SRS}$, су

$$\hat{a}_{ko,ah} = \frac{n\overline{X}_{RSS} - X_{(n)}}{n(X_{(n)} - \overline{X}_{RSS})}$$

$$\hat{d}_{ko,ah} = \frac{(n\hat{a}_{ko,ah} + 1)X_{(n)}}{n\hat{a}_{ko,ah}}$$

где је $\overline{X}_{RSS}$ аритметичка средина у узорку рангираних скупова, а $X_{(n)}$ је максимална вредност у узорку.

Оцена параметра a по узору на оцену добијену методом максималне веродостојности је

$$\hat{a}_{mle,ah} = \frac{n}{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} \ln \frac{\hat{d}_{mle,ah}}{X_{ij}}}$$

где је $\hat{d}_{mle,ah}$ оцена параметра d и једнака је максималној вредности у узорку рангираних скупова, $\hat{d}_{mle,ah} = X_{(n)}$.

Оцене по узору на оцене добијене методом квантила су

$$\hat{a}_{qt,ah} = \frac{\ln\left(\frac{p_1}{p_2}\right)}{\ln \frac{X_{p_1}}{X_{p_2}}}$$

и

$$\hat{d}_{qt,ah} = \frac{X_{p_1}}{p_1^{\frac{1}{\hat{a}_{qt,ah}}}}$$

где се X_{p_1} и X_{p_2} бирају из узорка рангираних скупова.

Оцене по узору се могу добити и из оцене најмањих квадрата:

$$\hat{a}_{lsm,ah} = \frac{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} (\ln F(X_{ij}))^2 - \frac{(\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} \ln F(X_{ij}))^2}{mk}}{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} \ln F(X_{ij}) \ln X_{ij} - \frac{(\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} \ln F(X_{ij}))(\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} \ln X_{ij})}{mk}}$$

$$\hat{d}_{lsm,ah} = e^{\frac{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} \ln X_{ij}}{mk} - \frac{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} \ln F(X_{ij})}{mk\hat{a}_{lsm,ah}}}$$

Оцена по узору на модифковану непристрасну оцену са најмањом дисперзијом је

$$\hat{a}_{mmvue,ah} = \frac{mk - 1}{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} ((\ln X)_{(1)} - \ln X_{ij})}$$

и

$$\hat{d}_{mmvue,ah} = e^{(\ln X)_{(1)} - \frac{1}{mk(mk-1)} \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{k} (\ln X_{ij} - (\ln X)_{(1)})}$$

где је $(\ln X)_{(1)}$ најмања вредност из скупа $(\ln X_{11}, \ln X_{12}, ..., \ln X_{1k}, ..., \ln X_{mk})$.

4.2.3 Симулација и упоређивање оцена

Да резимирамо, из простог случајног узорка добијене су оцене параметра a: $\hat{a}_{mom,SRS}$, $\hat{a}_{ko,SRS}$, $\hat{a}_{mle,SRS}$, $\hat{a}_{qt,SRS}$, $\hat{a}_{lsm,SRS}$ и $\hat{a}_{mmvue,SRS}$. Из узорка рангираних скупова добијена је оцена модификованом методом максималне веродостојности $\hat{a}_{mmle,RSS}$, као и оцене по узору $\hat{a}_{mom,ah}$,$\hat{a}_{ko,ah}$, $\hat{a}_{mle,ah}$, $\hat{a}_{qt,ah}$, $\hat{a}_{lsm,ah}$ и $\hat{a}_{mmvue,ah}$. Истим методама добијене су и оцене параметра d.

Симулација је вршена за исте вредности параметара n, m и k као код једнопараметарске степене расподеле. Вредности параметара a и d које су посматране су $(a, d) = ((1, 1), (3, 2)(5, 4))$. За одређивање квалитета оцена коришћени су пристрасност и средњеквадратна грешка.

У табели 4.4 дата је релативна пристрасност параметра a оцењена из простог случајног узорка.

a	d	n	$\hat{a}_{mom,SRS}$	$\hat{a}_{ko,SRS}$	$\hat{a}_{mle,SRS}$	$\hat{a}_{qt,SRS}$	$\hat{a}_{lsm,SRS}$	$\hat{a}_{mmvue,SRS}$
1	1	12	6,4843	6,8540	19,5956	14,4704	7,9012	9,6293
1	1	24	2,2144	2,6269	9,0158	18,3653	2,6046	4,4735
3	2	12	11,8553	7,6331	20,0166	13,9321	7,7985	10,0152
3	2	24	5,4855	3,4119	9,0487	18,6290	2,7461	4,5050
5	4	12	15,7624	8,5955	19,6379	14,2825	8,0213	9,6680
5	4	24	7,0757	3,9843	9,0182	18,7944	2,7434	4,4758

Табела 4.4: Пристрасност оцена параметра a из простог случајног узорка

Пристрасност ових оцена је генерално већа него што је то био случај код једнопараметарске степене расподеле. Оцене $\hat{a}_{mle,SRS}$ и $\hat{a}_{qt,SRS}$ имају јако велику пристрасност, па нису добар избор за оцену параметра a. Оцене и $\hat{a}_{mom,SRS}$ имају умерено велику пристрасност, док $\hat{a}_{ko,SRS}$ и $\hat{a}_{lsm,SRS}$ имају најмању пристрасност међу оценама параметра a из простог случајног узорка. Код свих оцена се уочава да је пристрасност доста мања за $n = 24$, у односу када је $n = 12$. Не уочава се разлика у пристрасности за различите вредности параметара a и d.

У табели 4.5 дата је пристрасност оцена параметра d из простог случајног узорка. Пристрасност је мања него код оцена параметра a за све оцене. Оцена $\hat{d}_{ko,SRS}$ опет има најмању пристрасност, и то јако малу.

a	d	n	$\hat{d}_{mom,SRS}$	$\hat{d}_{ko,SRS}$	$\hat{d}_{mle,SRS}$	$\hat{d}_{qt,SRS}$	$\hat{d}_{lsm,SRS}$	$\hat{d}_{mmvue,SRS}$
1	1	12	0,6204	0,5389	7,7334	2,7286	2,4809	15,1058
1	1	24	0,3241	0,1151	4,0116	3,6298	2,3707	7,9243
3	2	12	0,1923	0,0872	2,7103	0,9798	0,3214	5,3753
3	2	24	0,1420	0,0461	1,3644	1,2611	0,4458	2,7259
5	4	12	0,3842	0,0432	1,6494	0,6045	0,0940	3,2781
5	4	24	0,1897	0,0082	0,8308	0,7543	0,2209	1,6537

Табела 4.5: Пристрасност оцена параметра d из простог случајног узорка

Код параметра d и оцена $\hat{d}_{mom,SRS}$ има малу пристраснст. Смањење пристрасности са порастом обима узорка није лако уочљиво као што се то случај са оценама параметра a. Штавише, код оцене $\hat{d}_{qt,SRS}$ пристрасност је већа када је $n = 12$, у односу на ситуацију када је $n = 24$. Пристрасност се мења са променом параметара a и d. Највећа пристрасност је за вредности параметара $(a, d) = (1, 1)$. Са порастом вредности параметара a и d, пристрасност је мања.

Добијене вредности за пристрасност оцена параметра a из узорка рангираних скупова дата је у табели 4.6. Генерално, пристрасност је већа него што је била код оцена из узорка рангираних скупова за једнопараметарску степену расподелу. Оцене по узору $\hat{a}_{mle,ah}$ и $\hat{a}_{qt,ah}$ имају велику пристрасност, што није неочекивано с обзиром да су имале велику пристрасност и код простог случајног узорка. Иако велика, пристрасност ових оцена је мања него што је код простог случајног узорка. Пристрасност оцене $\hat{a}_{mmle,RSS}$ је велика, иако је код једнопараметарске степене расподеле оцена добијена овом методом дала задовољавајуће резултате. Оцене $\hat{a}_{mom,ah}$, $\hat{a}_{ko,ah}$, $\hat{a}_{lsm,ah}$ и $\hat{a}_{mmvue,ah}$ имају мању пристрасност него оцене добијене истим методама код простог случајног узорка. Може се уочити смањивање пристрасности са повећањем обима узорка за све оцене. Промена пристрасности са променом вредности параметара a и d се не може уочити. Пристрасност благо опада са порастом величине циклуса

У табели 4.7 дата је добијена пристрасност оцена параметра d из узорка рангираних скупова. Резултати су слични као код оцена из простог случајног узорка. Пристрасност је углавном мала, осим код оцене $\hat{d}_{mmvue,ah}$. Пристрасност је мања за већи обим узорка за све оцене осим $\hat{d}_{qt,ah}$. Јасно се уочава да је пристрасност мања када су параметри a и d већи. Пристрасност благо опада са порастом величине циклуса.

a	d	n	m	k	$\hat{a}_{mmle,RSS}$	$\hat{a}_{mom,ah}$	$\hat{a}_{ko,ah}$	$\hat{a}_{mle,ah}$	$\hat{a}_{qt,ah}$	$\hat{a}_{lsm,ah}$	$\hat{a}_{mmvue,ah}$
1	1	12	2	6	18,4261	3,6664	3,7842	16,8706	12,3254	5,7415	7,1314
1	1	12	3	4	16,8898	1,5594	2,5172	14,1918	9,1205	2,3698	4,6759
1	1	12	4	3	16,9210	0,0856	0,8172	13,5612	8,3758	1,6900	4,0978
1	1	24	2	12	8,7206	1,1490	1,6535	8,0193	17,1243	1,6071	3,5185
1	1	24	3	8	8,2197	0,6509	1,3838	6,9878	16,0356	0,4890	2,5300
1	1	24	4	6	7,8536	0,2639	1,1556	6,2631	15,0179	0,5496	1,8354
3	2	12	2	6	17,7564	9,0277	4,6537	16,2756	11,3041	4,9896	6,5860
3	2	12	3	4	17,8183	7,1736	3,3705	15,3339	10,415	3,8004	5,7227
3	2	12	4	3	17,0787	5,3955	2,2183	13,7812	8,5959	1,7657	4,2994
3	2	24	2	12	8,5486	3,8004	1,9983	7,8055	16,8454	1,3745	3,3136
3	2	24	3	8	8,0083	2,9603	1,6669	6,7256	15,5999	0,1503	2,2787
3	2	24	4	6	8,0005	2,6365	1,2747	6,4965	15,7580	0,1407	2,0592
5	4	12	2	6	18,1311	12,7798	5,9991	16,4065	11,3661	4,7704	6,7059
5	4	12	3	4	17,1099	10,0956	4,1445	14,5376	9,6852	3,1785	4,9928
5	4	12	4	3	17,1683	8,2076	2,3684	13,6351	7,8452	1,1164	4,1655
5	4	24	2	12	8,6079	6,0158	2,6920	7,8431	16,9478	1,4111	3,3496
5	4	24	3	8	8,1422	5,0514	2,1051	6,8879	15,6325	0,2933	2,4342
5	4	24	4	6	8,2441	3,5533	1,0334	6,6229	14,9480	0,2715	2,1803

Табела 4.6: Пристрасност оцена параметра a из узорка рангираних скупова

a	d	n	m	k	$\hat{d}_{mmle,RSS}$	$\hat{d}_{mom,ah}$	$\hat{d}_{ko,ah}$	$\hat{d}_{mle,ah}$	$\hat{d}_{qt,ah}$	$\hat{d}_{lsm,ah}$	$\hat{d}_{mmvue,ah}$
1	1	12	2	6	7,2755	1,2998	0,9262	7,2755	2,2055	3,6803	14,7234
1	1	12	3	4	6,8895	2,0478	1,2137	6,8895	1,7057	5,3397	14,4337
1	1	12	4	3	6,6227	2,1438	1,4289	6,6227	1,3822	6,3262	14,1852
1	1	24	2	12	3,9314	0,7762	0,2939	3,9314	3,3724	2,9805	7,8501
1	1	24	3	8	3,7250	0,9854	0,3447	3,7250	3,0664	3,7384	7,6681
1	1	24	4	6	3,6432	1,0740	0,3726	3,6432	2,9561	3,8992	7,5981
3	2	12	2	6	2,5320	0,1146	0,2531	2,5320	0,7821	0,7109	5,2237
3	2	12	3	4	2,4355	0,4014	0,3386	2,4355	0,6684	1,1216	5,1296
3	2	12	4	3	2,3508	0,6013	0,4707	2,3508	0,5832	1,2095	5,0500
3	2	24	2	12	1,3278	0,1009	0,0733	1,3278	1,1426	0,7444	2,6921
3	2	24	3	8	1,2879	0,2516	0,1079	1,2879	1,0529	0,8888	2,6578
3	2	24	4	6	1,2380	0,2556	0,1329	1,2380	1,0190	0,9610	2,6054
5	4	12	2	6	1,5191	0,1463	0,1103	1,5191	0,4601	0,4263	3,1601
5	4	12	3	4	1,4569	0,0674	0,1962	1,4569	0,3880	0,5679	3,1062
5	4	12	4	3	1,4291	0,1222	0,2616	1,4291	0,3498	0,8033	3,0793
5	4	24	2	12	0,8025	0,0618	0,0573	0,8025	0,6977	0,3726	1,6255
5	4	24	3	8	0,7723	0,0081	0,0545	0,7723	0,6499	0,4686	1,5987
5	4	24	4	6	0,7559	0,0769	0,0875	0,7559	0,6095	0,5779	1,5817

Табела 4.7: Пристрасност оцена параметра d из узорка рангираних скупова

Након пристрасности испитана је и средњеквадратна грешка оцена. Као и код једнопараметарске степене расподеле посматран је количник средњеквадратних грешака. Посматране су остале оцене у односу на оцену добијену из простог случајног узорка комбинованом методом.

$$efikasnost = \frac{MSE(\hat{a}_{ko,SRS})}{MSE(\hat{a})}.$$

У табели 4.8 дата је ефикасност оцене $\hat{a}_{ko,SRS}$ у односу на остале оцене из простог случајног узорка. Посматрани количник је углавном мањи од 1 што говори да оцена $\hat{a}_{ko,SRS}$ има мању средњеквадратну грешку у односу на остале оцене из простог случајног узорка. Због мале пристрасности и мале средњеквадратне грешке $\hat{a}_{ko,SRS}$ је добра оцена из простог случајног узорка. Оцена $\hat{a}_{ko,SRS}$ има већу средњеквадратну грешку само од $\hat{a}_{mumvue,SRS}$. Како ни оцена $\hat{a}_{mumvue,SRS}$ нема превелику пристрасност, и она може бити добар избор за оцену параметра a.

a	d	n	$\hat{a}_{mom,SRS}$	$\hat{a}_{mle,SRS}$	$\hat{a}_{qt,SRS}$	$\hat{a}_{lsm,SRS}$	$\hat{a}_{mmvue,SRS}$
1	1	12	1,0152	0,8034	0,6265	0,7897	1,1140
1	1	24	1,0883	1,0256	0,4428	0,8194	1,2299
3	2	12	0,6794	0,7230	0,5735	0,7359	0,9992
3	2	24	0,7555	0,8556	0,3654	0,6900	1,0286
5	4	12	0,5860	0,7187	0,5604	0,7033	0,9927
5	4	24	0,6311	0,8390	0,3520	0,6462	1,0068

Табела 4.8: Ефикасност оцене $\hat{a}_{ko,SRS}$ у односу на друге оцене из простог случајног узорка

Ефикасност оцене $\hat{a}_{ko,SRS}$ у односу на оцене из узорка рангираних скупова дата је у табели 4.9. Најпре је битно приметити да је за све оцене осим $\hat{a}_{qt,SRS}$ посматрана ефикасност већа од 1 за бар неке вредности параметара a и d. Како оцена $\hat{a}_{qt,SRS}$ има и велику пристрасност, метода квантила није добра метода за добијање оцене ни код простог случајног узорка, ни код узорка рангираних скупова. Средњеквадратна грешка оцене $\hat{a}_{ko,SRS}$ је већа од средњеквадратне грешке оцена $\hat{a}_{mmle,RSS}$, $\hat{a}_{ko,ah}$, $\hat{a}_{mle,ah}$ и $\hat{a}_{mmvue,ah}$ за скоро све вредности параметара a и d и обиме узорка. Средњеквадратна грешка ових оцена је и до два и по пута мања. Како је посматрана величина мања од један или блиска јединици у табели 4.8, закључујемо да су ове оцене боље од оцене добијене било којом методом из простог случајног узорка. Везано за ове четири оцене, може се приметити и да њихова ефикасност расте у односу на $\hat{a}_{ko,SRS}$ како расте величина циклуса. За мање вредности параметара a

и d оцена $\hat{a}_{mom,ah}$ је ефикаснија од $\hat{a}_{ko,SRS}$, а веће вредности параметара ефикаснија је оцена $\hat{a}_{mom,ah}$. Оцена $\hat{a}_{lsm,ah}$ има екстремно понашање. За неке вредности параметара, њена средњеквадратна грешка је много мања од средњеквадратне грешке $\hat{a}_{ko,SRS}$, а за неке је много већа. Посматрајући пристрасност и ефикасност, најбоља оцена је $\hat{a}_{ko,ah}$ јер има најмању пристрасност и најмању средњеквадратну грешку, а затим оцена $\hat{a}_{mmvue,ah}$.

a	d	n	m	k	$\hat{a}_{mmle,RSS}$	$\hat{a}_{mom,ah}$	$\hat{a}_{ko,ah}$	$\hat{a}_{mle,ah}$	$\hat{a}_{qt,ah}$	$\hat{a}_{lsm,ah}$	$\hat{a}_{mmvue,ah}$
1	1	12	2	6	1,0900	1,3232	1,4031	1,0867	0,7495	1,0688	1,5302
1	1	12	3	4	1,3172	1,7816	1,853	1,3760	0,8517	32,2100	1,9725
1	1	12	4	3	1,7058	2,3988	2,5786	1,8770	1,0097	1,8251	2,7199
1	1	24	2	12	1,4450	1,4899	1,5712	1,4530	0,5078	3,2199	1,7652
1	1	24	3	8	1,7282	1,7984	1,9747	1,7358	0,5450	1,2044	2,1261
1	1	24	4	6	2,0611	2,2015	2,4314	2,0963	0,6107	1,1531	2,5811
3	2	12	2	6	0,9921	0,8611	1,4553	0,9990	0,6744	1,3233	1,4086
3	2	12	3	4	1,3175	1,0805	2,0949	1,3844	0,8023	2,0388	1,9786
3	2	12	4	3	1,4414	1,2552	2,5115	1,5948	0,8825	8,3396	2,3322
3	2	24	2	12	1,1962	0,8888	1,4846	1,1927	0,4087	0,1145	1,4522
3	2	24	3	8	1,4810	1,0538	1,9168	1,4871	0,4546	1,0818	1,8179
3	2	24	4	6	1,6543	1,1640	2,2672	1,6821	0,4812	8,0470	2,0790
5	4	12	2	6	1,0589	0,8127	1,5717	1,0787	0,7414	0,7416	1,5284
5	4	12	3	4	1,2837	0,9257	2,0648	1,3647	0,8005	0,0997	1,9681
5	4	12	4	3	1,3851	1,0086	2,3556	1,5277	0,8255	0,7879	2,2200
5	4	24	2	12	1,1037	0,7295	1,3761	1,1063	0,4008	91,0584	1,3393
5	4	24	3	8	1,4119	0,8764	1,8632	1,4377	0,4395	0,5417	1,7720
5	4	24	4	6	1,6613	0,9561	2,2423	1,6849	0,4868	13,3852	2,0738

Табела 4.9: Ефикасност оцене $\hat{a}_{ko,SRS}$ у односу на оцене из узорка рангираних скупова

У табели 4.10 дата је ефикасност оцене $\hat{d}_{ko,SRS}$ у односу на друге оцене из простог случајног узорка. Ова оцена има углавном вишеструко мању средњеквадратну грешку у односу на остале оцене из простог случајног узорка. С обзиром на јако малу пристрасност, ово је најбоља оцена за параметар d из простог случајног узорка.

a	d	n	$\hat{d}_{mom,SRS}$	$\hat{d}_{mle,SRS}$	$\hat{d}_{qt,SRS}$	$\hat{d}_{lsm,SRS}$	$\hat{d}_{mmvue,SRS}$
1	1	12	0,4919	0,6261	1,0229	0,0905	0,2474
1	1	24	0,2700	0,5569	0,3732	0,0474	0,2199
3	2	12	0,2493	0,5707	0,9364	0,1229	0,2281
3	2	24	0,1306	0,5282	0,3546	0,0571	0,2120
5	4	12	0,1965	0,5651	0,9180	0,1286	0,2265
5	4	24	0,0951	0,5225	0,3452	0,0566	0,2075

Табела 4.10: Ефикасност оцене $\hat{d}_{ko,SRS}$ у односу на друге оцене из простог случајног узорка

У табели 4.11 дата је ефикасност оцене $\hat{d}_{ko,SRS}$ у односу на оцене добијене из узорка рангираних скупова. Супериорност оцена из узорка рангираних скупова није толико уочљива као што је то био случај код оцена за параметар a. Оцена $\hat{d}_{ko,SRS}$ има мању средњеквадратну грешку у односу на оцене $\hat{d}_{mmle,RSS}$, $\hat{d}_{mom,ah}$, $\hat{d}_{mle,ah}$, $\hat{d}_{lsm,ah}$ и $\hat{d}_{mmvue,ah}$. Оцена $\hat{d}_{ko,SRS}$ има већу средњеквадратну грешку у односу на оцене $\hat{d}_{qt,RSS}$ када је обим узорка $n = 12$, а мању када је $n = 24$. Када се посматра однос $\hat{d}_{ko,SRS}$ и $\hat{d}_{ko,ah}$, може се приметити да $\hat{d}_{ko,SRS}$ има већу средњеквадратну грешку за све вредности параметара a, d и n.

a	d	n	m	k	$\hat{d}_{mmle,RSS}$	$\hat{d}_{mom,ah}$	$\hat{d}_{ko,ah}$	$\hat{d}_{mle,ah}$	$\hat{d}_{qt,ah}$	$\hat{d}_{lsm,ah}$	$\hat{d}_{mmvue,ah}$
1	1	12	2	6	0,6822	0,4979	1,4031	0,6822	1,1209	0,0889	0,2600
1	1	12	3	4	0,8290	0,6029	1,8530	0,8290	1,4021	0,0936	0,2962
1	1	12	4	3	0,9008	0,6920	2,5786	0,9008	1,5839	0,0995	0,3067
1	1	24	2	12	0,5999	0,2833	1,5712	0,5999	0,4131	0,0495	0,2292
1	1	24	3	8	0,6437	0,3326	1,9747	0,6437	0,4566	0,0494	0,2424
1	1	24	4	6	0,6590	0,3578	2,4314	0,6590	0,4900	0,0504	0,2394
3	2	12	2	6	0,6815	0,2769	1,4553	0,6815	1,1514	0,1344	0,2537
3	2	12	3	4	0,7491	0,2920	2,0949	0,7491	1,3074	0,1319	0,2626
3	2	12	4	3	0,7576	0,2956	2,5115	0,7576	1,3525	0,1276	0,2560
3	2	24	2	12	0,5917	0,1445	1,4846	0,5917	0,4098	0,0607	0,2273
3	2	24	3	8	0,6094	0,1533	1,9168	0,6094	0,4339	0,0602	0,2283
3	2	24	4	6	0,6770	0,1699	2,2672	0,6770	0,4995	0,0622	0,2426
5	4	12	2	6	0,6575	0,2184	1,5717	0,6575	1,1091	0,1382	0,2468
5	4	12	3	4	0,7488	0,2375	2,0648	0,7488	1,3317	0,1443	0,2611
5	4	12	4	3	0,8107	0,2425	2,3556	0,8107	1,4927	0,1399	0,2693
5	4	24	2	12	0,5823	0,1041	1,3761	0,5823	0,4038	0,0600	0,2215
5	4	24	3	8	0,6218	0,1129	1,8632	0,6218	0,4298	0,0632	0,2288
5	4	24	4	6	0,6395	0,1145	2,2423	0,6395	0,4683	0,0595	0,2290

Табела 4.11: Ефикасност оцене $\hat{d}_{ko,SRS}$ у односу на оцене из узорка рангираних скупова

Поглавље 5

Закључак

Узорковање методом рангираних скупова се показује као добра алтернатива простом случајном узорку. У случају да постоји помоћно обележје које је лако и јефтино мерити, узорковање методом рангираних скупова даће боље резултате у односу на прост случајан узорак. Уколико се узорак бира овом методом смањује се вероватноћа да узорак обухвати превише екстрема.

Узорковање методом рангираних скупова је област која има пуно потенцијала, па је стога врло популарна. Поред особина класичне методе рангираних скупова, постоје и многа побољшања у зависности од специфичности испитиваног обележја. Неке од модификација класичне методе су метода рангираних скупова са оптималном алокацијом, метода рангираних скупова помоћу медијане и метода рангираних скупова помоћу екстрема.

Особине оцена добијених из узорка рангираних скупова углавном су боље од особина оцена добијених из простог случајног узорка. Оцена средње вредности из узорка рангираних скупова је непристрасна и има мању дисперзију у односу на оцену средње вредности из простог случајног узорка. Оцена дисперзије из узорка рангираних скупова није непристрасна, али је асимптотски непристрасна. Она не мора бити ефикаснија од оцене добијене из простог случајног узорка (у смислу мање средњеквадратне грешке) нарочито ако је обим узорка мали. Међутим, за велике узорке може се доказати да је оцена из узорка рангираних скупова ефикаснија од оцене добијене из простог случајног узорка.

У литератури се могу наћи оцене параметара из узорка рангираних скупова за броје расподеле. У овом раду су приказане оцене параметара за нормалну и Пуасонову расподелу. Поред класичних метода за добијање оцена, предложена су и бројна побољшања.

Степена расподела има широку примену у електроници, економији и код редова чекања. Стога су изведене оцене параметара из узорка рангираних скупова код једнопараметарске и двопараметарске степене расподеле. Оцене су упоређене симулацијом где су посматране пристрасност и средњеквадратна грешка.

Код једнопараметарске степене расподеле параметар a је оцењен методом момената, модификованом методом максималне веродостојности и методом медијане из простог случајног узорка. Из узорка рангираних скупова параметар a је оцењен модификованом методом максималне веродостојности и оценама по узору на оцене добијене из простог случајног узорка. Међу оценама добијеним из простог случајног узорка најмању пристрасност и средњеквадратну грешку има оцена добијена модификованом методом максималне веродостојности. Премда има мању пристрасност, ова оцена има већу средњеквадратну грешку у односу на скоро све оцене добијене из узорка рангираних скупова. Као најбоља се показује оцена по узору методом момената из узорка рангираних скупова која има релативно малу пристрасност и малу средњеквадратну грешку.

Код двопараметарске степене расподеле оцењени су параметри a и d. Из простог случајног узорка параметри су оцењивани методом момената, комбинованом методом, методом максималне веродостојности, методом квантила, методом најмањих квадрата и као модификована непристрасна оцена са најмањом дисперзијом. Из узорка рангираних скупова параметри су оцењени модификованом методом максималне веродостојности и оценама по узору на оцене из простог случајног узорка. Најбоље особине показују оцена по узору комбинованом методом из узорка рангираних скупова и када се посматра пристрасност и средњеквадратна грешка. Оцена по узору на модификовану непристрасну оцену са најмањом дисперзијом такође има малу пристрасност и малу средњеквадратну грешку када се оцењује параметар a. Међутим, ова оцена је лоша када се оцењује d.

Свеукупно, симулација је допринела закључку да уколико се узорковање врши на прави начин и изабере одговарајућа метода оцењивања, тада су оцене из узорка рангираних скупова прецизније од оцена из простог случајног узорка.

У даљим истраживањима параметри степене расподеле из узорка рангираних скуповима могу се оценити и другим методама, као на пример, Бајесовим оценама. Поред тога, методе оцењивања примењене у овом раду за добијање оцена из простог случајног узорка и из узорка рангираних скупова могу се применити и на друге расподеле. С обзиром на то да комбинована метода није често коришћена, било би интересантно проверити да ли даје добре резултате и када се примени на узорцима рангираних скупова из других расподела.

Поглавље 6

Литература

[1] Abu-Dayyeh, W., Assrhani, A., Ibrahim, K. (2013). Estimation of the shape and scale parameters of Pareto distribution using ranked set sampling. Stat. Papers 54, 207 - 225.

[2] Barnett, V., Baretto, M. C. (2001). Estimators for a Poisson parameter using ranked set sampling. Journal of Applied Statistics. Vol. 28. No. 8. 929-941.

[3] Barnett, V., Moore, K. (1997), Best linear unbiased estimates in ranked set sampling with particular reference to imperfect ordering. Journal of Applied Statistics. 24. 697-710.

[4] Ferreira M. A., Andrade M. (2011): The M/G/1 queue busy period distribution exponentiality. Journ. of Appl. Math. 4, 249 - 260

[5] Zaka, A., Akhtar, A. S. (2013). Methods for estimating the parameters of the Power Function Distribution. Pakistan Journal of Statistics and Operation Research. Vol. 9 No. 2.

[6] Zaka, A., Akhtar, A. S. Estimation of Parameters of the Power Function Distribution: Different Classical Methods http://www.academia.edu/6526953/Estimation_of_Parameters _of_the_Power_Function_Distribution_Different_Classical_Methods

[7] Zheng, S. Suficient Statistics and Exponential Family. http://afhadi.blog.unej.ac.id/wp-content/uploads/2014/07/Sufficient1.pdf

[8] Lehmann. E. L., Casella, G., (1983). Theory of Point Estimation. Springer

[9] Lloyd, E. H. (1952). Least squares estimation of location and scale parameters using order statistics. Biometrika. 39. 88- 95.

[10] Марковић, Д. (2009). Проблем процене параметара у Вејбуловом моделу, Свеучилиште у Загребу, Природословно - математички факултет. Загреб.

[11] Meniconi, M., Barry, D. M., (1996). The power function distribution: A useful and simple distribution to assess electrical component reliability, Microelectronics Reliability, Vol. 36, Issue 9, 1207 - 1212.

[12] Mehrotra. K., Nanda, P. (1974). Unbiased estimation of parameters by order statistics in the case of censored samples. Biometrika 61. 601 - 606.

[13] Миниħ, М. (2014). Узорковање методом рангираних скупова. В симпозијум - математика и примене. Београд

[14] Minić, M., Obradović, M., (2014). Estimating parameters using Ranked Set Sampling. Conference on Applied Statistics 2014. Ribno. Slovenia

[15] Muttalk, A. (1997). Median ranked set sampling. Journal of Applied Statistical Science. 6(4), 245-55.

[16] McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. Aust. J. Agric. Res. 3. 385 - 390.

[17] Patil, P. (2002). Ranked set sampling. Encyclopedia of Environmetrics. Vol. 3, 1684-1690.

[18] Rytgaard M (1990) Estimation in the Pareto distribution. ASTIN Bull J Int Actuar Assoc 20(2):201 - 215

[19] Rao, C.R. (1973). Linear Statitislical Inference and Its Applications. John Wiley & Sons.

[20] Samawi, H.M., Ahmed, M.S., Abu-Dayyeh, W.A. (1996). Estimating the population mean using extreme raked set sampling. The Biometrical Journal. 38(5). 577-86.

[21] Sinha, B. K., Sinha, B. K., Purkayastha, S. (1996). On some aspects of ranked set sampling for estimation of normal and exponential parameters. Statistical Decisions 14. 223- 240.

[22] Stokes, S.L. (1980). Estimation of Variance Using Judgment Ordered Ranked Set Samples. Biometrics 36. 35-42.

[23] Takahasi, K., Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. Annals of the Institute of Statistical Mathematics. Vol. 20. Issue 1. 1-31.

[24] Hsieh, H. K. (2012). A.R.E. of Sample mean to Sample Median. http://people.math.umass.edu/ hsieh/stat725/s725ht1-s12.pdf

[25] Chen, Z. Bai, Z. Sinha B.,(2003). Ranked Set Sampling: Theory and Applications. Springer.

[26] Yu, P., Lam, K., Sinha, B. (1999) . Estimation of normal variance based on balanced and unbalanced ranked set samples. Environmental and Ecological Statistics 6 23-45.

[27] Wolfe, D. A. (2004). Ranked Set Sampling: An Approach to More Efficient Data Collection. Statistical Science. 19. 636-643.

[28] Wu, Z., Kazaz, B., Webster, S., Yang, K.K. (2012). Ordering, Pricing, and Lead-Time Quotation Under Lead-Time and Demand Uncertainty, Production and Operations Management, Vol. 21, No. 3, 576 - 589

Биографија

Марија Минић је рођена 04.02.1990. године у Крушевцу, где је и завршила Гимназију, смер - специјализовано одељење ученика обдарених за математику. Основне студије на Математичком факултету, смер - вероватноћа, статистика и актуарска математика, завршила је 2013. године са просечном оценом $8,95$. Студент је мастера примењене статистике на Универзитетском центру за примењену статистику у Новом Саду. На Математичом факултету у Београду дала је све испите на мастер студијама, смер - вероватноћа, статистика и актуарска математика, са просечном оценом $9,67$. Презентовала је на скупу ,, Conference on Applied Statistics 2014" у Рибну (Словенија) са радом ,,Estimating parameters using Ranked Set Sampling", као и на скупу ,,V симпозијум - математика и примене" у Београду са радом ,,Узорковање методом рангираних скупова". Радила је као наставник математике у Гимназији Михајило Пупин, као студент демонстратор на курсу статистика на Пољопривредном факултету и као аналитичар тржишта. Тренутно ради у основној школи Соња Маринковић у Београду.