УНИВЕРЗИТЕТ У БЕОГРАДУ

МАТЕМАТИЧКИ ФАКУЛТЕТ

Јелена М. Јоцковић

СТОХАСТИЧКИ МОДЕЛИ ПРЕКОРАЧЕЊА ВИСОКОГ НИВОА И ПРОБЛЕМИ ЧЕКАЊА

докторска дисертација

Београд, 2012.

UNIVERSITY OF BELGRADE

FACULTY OF MATHEMATICS

Jelena M. Jocković

STOCHASTIC MODELS FOR EXCEEDANCES OVER HIGH THRESHOLDS AND WAITING TIME PROBLEMS

Doctoral Dissertation

Belgrade, 2012

Ментор:
др Павле Младеновић, редовни професор
Математичког факултета Универзитета у Београду

Чланови комисије:
др Слободанка Јанковић, ванредни професор
Математичког факултета Универзитета у Београду,

др Љиљана Петровић, редовни професор
Економског факултета Универзитета у Београду

Датум одбране:

Рад на овој докторској дисертацији био ми је велико задовољство. Захвалност за то дугујем, пре свега, мом ментору, професору Павлу Младеновићу, који ми је, кроз наш заједнички рад, пренео део свог великог искуства у овој научној области. Такође, његове драгоцене сугестије и наше бројне дискусије о комбинаторним проблемима омогућиле су да ова докторска дисертација добије свој коначни облик.

Велику захвалност дугујем и колегама из Севиље, професору Емилију Каризоси и др Пепи Рамирез Кобо, од којих сам, током неколико научно-истраживачких посета, научила много тога из области Вероватноће и статистике, али и о научном раду уопште.

Такође, велику захвалност дугујем и професорки Стани Никчевић са Фармацеутског факултета, за дугогодишњу подршку и многобројне корисне савете.

Хвала мојој породици на љубави и разумевању, као и драгим пријатељима, који су у току пута "држали палчеве".

Овај рад посвећујем свом оцу, који није доживео да га види, али би му се сигурно највише радовао.

Београд, септембар 2012. Јелена Јоцковић

Наслов докторске дисертације: *Стохастички модели прекорачења високог нивоа и проблеми чекања*

Резиме: У анализи екстремалних догађаја постоје два основна приступа, који полазе од различитих начина издвајања екстрема из датог узорка $x_1, x_2, \ldots, x_N$. Један метод анализира максимуме (минимуме). Други метод се бави величинама облика $y_1 = x_1 - u, y_2 = x_2 - u, \ldots, y_n = x_n - u$, где су $y_1, y_2, \ldots, y_n$ чланови почетног узорка већи (мањи) од унапред дефинисаног прага u.

Циљеви ове докторске дисертације су примена техника за рад са екстремним вредностима у неким комбинаторним проблемима и допринос области статистичког моделовања екстремалних догађаја. Дисертација се састоји из три дела.

У првој глави дефинишу се расподеле екстремних вредности и генералисане Паретове расподеле. Ове две фамилије расподела имају кључну улогу у два приступа теорије екстремних вредности. Даље су дате теоријске основе оба приступа.

Друга глава бави се применом екстремалних техника у комбинаторним проблемима чекања. Разматра се проблем скупљања купона, који је задат на следећи начин: бирају се елементи са враћањем из скупа $\mathbb{N}_n = \{1, 2, \ldots, n\}$, под претпоставком да сваки елемент има вероватноћу $1/n$ да буде изабран. Величина од интереса је време чекања M_n док сви елементи скупа $\mathbb{N}_n$ или неке друге шеме не буду изабрани. У овом раду се посебно анализирају следећа два случаја:
1. M_n је време чекања до појаве свих елемената скупа $\mathbb{N}_n$ бар r пута, где је r дати природан број;
2. M_n је време чекања до појаве свих парова jj, $j \in \mathbb{N}_n$.
Представљени су нови резултати, у којима је одређена асимптотска расподела времена чекања M_n до краја експеримента, ако је познато да је изведен велики број избора, а експеримент није завршен, тј. гранична расподела прекорачења високог нивоа за ове случајне величине. Такође, одговорено је и на следећа питања: колики треба да буде високи ниво (одређен бројем избора) при коме се добија та гранична расподела; у којим случајевима гранична расподела зависи од високог нивоа; да ли, осим генералисаних Паретових расподела, за граничне расподеле постоје и неке друге могућности. Осим тога, за сваки од случајева оцењена је

брзина конвергенције ка граничној расподели.

Трећа глава ове дисертације посвећена је статистичком задатку оцене параметара и квантила генералисаних Паретових расподела. Модел који се разматра је двопараметарска верзија ових расподела, задата са:

$$W_{\gamma,\sigma}(x) = \begin{cases} 1 - e^{-\frac{x}{\sigma}}, & x \geq 0, \gamma = 0 \\ 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-\frac{1}{\gamma}}, & x \geq 0, \gamma > 0 \\ 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-\frac{1}{\gamma}}, & x \in \left[0, -\frac{\sigma}{\gamma}\right], \gamma < 0. \end{cases}$$

При оцењивању непознатих параметара, у случају кад је $\gamma < 0$, често се појављује проблем несагласности оцена са узорком на основу кога се врши оцењивање. То значи да постоје чланови узорка који су већи од оцењеног десног краја носача расподеле. У овом раду предложена је нова, општа техника за поправку несагласних оцена, која, у многим случајевима, повећава и ефикасност оцене. Техника је овде примењена на оцене методом момената и методом пондерисаних момената. Особине поправљених оцена анализиране су уз помоћ компјутерских симулација, а наведени су и примери. Последњи део овог рада посвећен је оцењивању високих квантила генералисаних Паретових расподела. Дати су резултати који се односе на робусност појединих оцена квантила, као и пример њихове примене у финансијама (оцена параметра вредности при ризику), са освртом на неке практичне проблеме који се при томе појављују.

Кључне речи: екстремне вредности, генералисане Паретове расподеле, прекорачења високог нивоа, проблем скупљања купона, несагласност оцена параметра са узорком, метод момената, метод пондерисаних момената, квантили расподеле, вредност при ризику

Научна област: Математика

Ужа научна област: Вероватноћа и статистика

УДК: 517.51, 519.21, 519.1:519.2 (043.3)

AMS Classification: 60G70, 60G32

Doctoral Dissertation Title: *Stochastic Models for Exceedances over High Thresholds and Waiting Time Problems*

Summary: Statistical methodology for dealing with extremes depend on how extreme values are defined. One way to extract extremes from a given sample $x_1, x_2, ..., x_N$ is to consider maxima (minima). The other way is to consider values $y_1 = x_1 - u, y_2 = x_2 - u, \ldots, y_n = x_n - u$, where $y_1, y_2, \ldots, y_n$ are sample members above (below) a given predetermined threshold u. These two methods lead to two different approaches in extreme value theory.

This doctoral dissertation has two main goals. One of them is to apply the techniques from extreme value framework to certain type of combinatorial problems. The other goal is to contribute to the field of statistical modeling of extremes. The dissertation consists of three chapters.

In the first chapter, we introduce generalized extreme value distributions and generalized Pareto distributions (GPD). These two families play key roles in the two approaches to modeling extremes. We set out the theoretical background for both approaches.

In the second chapter, we apply the extremal techniques to combinatorial waiting time problems. Precisely, we consider Coupon collector's problem, defined as follows: elements are sampled with replacement from the set $\mathbb{N}_n = \{1, 2, \ldots, n\}$ under assumption that each element has probability $1/n$ of being drawn. The subject of interest is the waiting time M_n until all elements of $\mathbb{N}_n$ or some other pattern are sampled. We focus our attention to the following two cases:

1. M_n is the waiting time until all elements of $\mathbb{N}_n$ are sampled at least r times, where r is a positive integer;
2. M_n is the waiting time until all pairs of elements jj, $j \in \mathbb{N}_n$ are sampled.

We present new results related to the asymptotic behavior of the waiting time M_n, if it is known that a large number of trials was performed and the experiment is not over. For both cases, we determine the limiting distribution of exceedances of M_n over high thresholds, and answer some related questions: how to choose a suitable high threshold (depending on n) in order to obtain a limiting distribution; under what conditions the limit does not depend on the threshold; are the generalized Pareto distributions the only possible limits. We also estimate the speed of convergence in both cases.

The third chapter of the dissertation is devoted to estimation of parameters and quantiles of the generalized Pareto distributions. We restrict the attention to the two-parameter version of GPD, defined as:

$$W_{\gamma,\sigma}(x) = \begin{cases} 1 - e^{-\frac{x}{\sigma}}, & x \geq 0, \gamma = 0 \\ 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-\frac{1}{\gamma}}, & x \geq 0, \gamma > 0 \\ 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-\frac{1}{\gamma}}, & x \in \left[0, -\frac{\sigma}{\gamma}\right], \gamma < 0. \end{cases}$$

Well known problem with this model is inconsistency with the sample data, which is that one or more sample observations exceed the estimated upper bound in case when $\gamma < 0$. We propose a new, general technique to overcome the inconsistency problem and improve performance of the existing GPD estimation methods. We apply the proposed technique to method-of-moments and method-of-probability-weighted-moments estimates, investigate its performance through computer simulation and provide some real data examples. Finally, we address the problem of estimating high GPD quantiles. We evaluate the robustness of some estimation methods through simulation study and present a case study from finance (value-at-risk estimation), with special emphasis to certain difficulties related to this field of application.

Key Words: extreme values, generalized Pareto distribution, exceedances over high threshold, coupon collector's problem, infeasible parameter estimates, method of moments, method of probability-weighted moments, quantiles of the distribution, value-at-risk

Scientific Area: Mathematics

Scientific Sub-area: Probability and Statistics

UDC: 517.51, 519.21, 519.1:519.2 (043.3)

AMS Classification: 60G70, 60G32

Садржај

Увод

Теорија екстремних вредности бави се особинама расподела екстремних, односно, највећих и најмањих вредности случајних величина. Те величине могу бити чланови случајног низа, реализације случајних процеса, итд. И саме екстремне вредности могу бити биране на различите начине: то могу бити максимуми, односно минимуми неких скупова, али и величине које су веће или мање од неког унапред дефинисаног нивоа (прага).

Из ова два начина издвајања екстрема настају два различита приступа у вероватносној теорији екстремних вредности. Теоријске основе првог приступа постављене су у радовима Fisher and Tippett (1928) [14] и Gnedenko (1943) [16]. Други приступ, који се базира на прекорачењима високог нивоа, формално је заснован у радовима Balkema and de Haan (1974) [2] и Pickands (1975) [35].

Ова теорија се интензивно развијала у последњих 20-30 година, а разлог за то су и велике могућности њене примене у хидрологији, метеорологији, финансијској математици, теорији ризика, комбинаторици и многим другим областима.

Предмет ове докторске дисертације су примене теорије екстремних вредности, прецизније, оне гране ове теорије која се бави прекорачењима високог нивоа, у неким теоријским проблемима (комбинаторни проблеми чекања) и неким практичнијим проблемима везаним за статистичко моделовање екстремалних догађаја. Дисертација се састоји из три дела.

У првој глави дати су познати резултати класичне теорије екстремних вредности и оне гране ове теорије која се бави прекорачењима високог нивоа. Уведене су фамилије расподела екстремних вредности и генералисаних Паретових расподела, дате њихове особине и наведене граничне теореме у којима ове две фамилије расподела имају кључну улогу. Такође, објашњено је проширење теорије екстремних вредности на стационарне случајне низове. При томе је коришћена литература [2], [11], [28], [31] и [36].

У другој глави објашњене су вероватносне технике које се примењују у комбинаторним проблемима. Наведени су постојећи резултати који се односе на граничну расподелу максимума неких случајних величина које се појављују у комбинаторном проблему скупљања купона. Они су садржани у радовима Erdös and Rényi (1961) [12], Holst (1986) [19] и Mladenović (1999, 2002, 2008) [30, 32, 33]. Изведени су нови резултати везани за граничне расподеле прекорачења високог нивоа за исте случајне величине. Неки од ових нових резултата су публиковани у раду [23].

Трећа глава се односи на оцењивање параметара и квантила двопараметарске фамилије генералисаних Паретових расподела. Наведене су дефиниције и особине различитих оцена. Предложен је нови метод за решавање проблема несагласности појединих оцена са узорком на основу кога се врши оцењивање, тј. нова оцена параметара. Наведени су резултати који су добијени поређењем нове оцене са постојећим, као и примери. Овај резултат је прихваћен за публиковање у [22]. Последње поглавље посвећено је оцењивању квантила генералисаних Паретових расподела. Дати су нови резултати који се односе на робусност неких оцена квантила, као и пример њихове примене у финансијама, са освртом на неке практичне проблеме који се ту појављују. Овај резултат објављен је у раду [21].

Глава 1

Вероватносни модели у теорији екстремних вредности

Теорија екстремних вредности, у ширем смислу, бави се тачним или приближним расподелама екстрема у неком скупу података. При томе, ти екстреми могу бити изабрани на различите начине.

Приступ преко максимума бави се случајним величинама облика

$$M_n = \max\{X_1, X_2, \ldots, X_n\} \quad \text{и} \quad m_n = \min\{X_1, X_2, \ldots, X_n\}, \quad n \in \mathbb{N},$$

као и њиховим својствима кад $n \to \infty$, при чему су $X_1, X_2, \ldots, X_n$ случајне величине са датим расподелама вероватноћа.

Приступ преко прекорачења високог нивоа бави се случајним величинама облика

$$Y_1 = X_1 - u, Y_2 = X_2 - u, \ldots, Y_n = X_n - u,$$

где су $X_1, X_2, \ldots, X_n$ случајне величине из датог скупа за које се зна да су веће од задатог високог нивоа (прага) u. Аналогно, могу се разматрати случајне величине мање од, унапред дефинисаног, ниског прага.

Из ова два начина посматрања (издвајања) екстремних вредности настају два различита вероватносна модела, са паралелним резултатима. У поглављу 1.1 дата је основа првог приступа, а у 1.2 основа другог. Проширење теорије екстремних вредности на стационарне низове објашњено је у поглављу 1.3.

1.1 Екстремне вредности у низу независних случајних величина са истом расподелом

1.1.1 Основни појмови

Ако за случајну величину $M_n = \max\{X_1, X_2, \ldots, X_n\}$ важи

$$\lim_{n\to\infty} P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = G(x) \tag{1.1}$$

за сваку тачку непрекидности x неке недегенерисане функције расподеле G и реалне бројеве $a_n > 0$ и b_n, онда се каже да је G **гранична расподела линеарно нормираног максимума** M_n, а a_n и b_n су **нормирајуће константе**.

У случају када су $X_1, X_2, \ldots, X_n$ независне случајне величине са заједничком функцијом расподеле F важи

$$\begin{aligned} P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} &= P\{X_1 \leq a_n x + b_n, \ldots, X_n \leq a_n x + b_n\} \\ &= (P\{X_1 \leq a_n x + b_n\})^n = F^n(a_n x + b_n). \end{aligned}$$

Тада се једнакост (1.1) своди на

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x)$$

и каже се још и да функција F припада **области привлачења за максимуме** функције расподеле G ($F \in D(G)$).

За функције расподела G_1 и G_2 се каже да су **истог типа** ако постоје реални бројеви $a > 0$ и b, такви да за свако $x \in \mathbb{R}$ важи

$$G_2(x) = G_1(ax + b).$$

Недегенерисана функција расподеле G је **максимум стабилна** ако за свако $n \geq 2$, $n \in \mathbb{N}$ постоје константе $a_n > 0$ и $b_n \in \mathbb{R}$, такве да за свако $x \in \mathbb{R}$ важи

$$G^n(a_n x + b_n) = G(x).$$

У даљем тексту биће наведени само резултати који се односе на максимум низа независних случајних величина. Одговарајући резултати за минимум добијају се коришћењем једнакости

$$\min\{X_1, X_2, \ldots, X_n\} = -\max\{-X_1, -X_2, \ldots, -X_n\}.$$

1.1.2 Гумбелова, Фрешеова и Вејбулова расподела

При овом приступу, важну улогу имају **расподеле екстремних вредности**, чији су стандардни представници:

- *Гумбелова расподела*:

$$G_0(x) = e^{-e^{-x}}, \text{ за } x \in \mathbb{R},$$

- *Фрешеова расподела с параметром* $\alpha > 0$:

$$G_{1,\alpha}(x) = \begin{cases} 0, & \text{за } x < 0; \\ e^{-x^{-\alpha}}, & \text{за } x \geq 0, \end{cases}$$

- *Вејбулова расподела с параметром* $\alpha > 0$:

$$G_{2,\alpha}(x) = \begin{cases} e^{-(-x)^{\alpha}}, & \text{за } x < 0; \\ 1, & \text{за } x \geq 0. \end{cases}$$

Параметар α код Фрешеове и Вејбулове расподеле се назива **параметар облика**, а за функције расподеле се каже да су дате у **α - параметризацији**.

Овако дефинисана једнопараметарска фамилија расподела може се додатно обогатити увођењем **параметра размере**, σ, и (или) **параметра положаја**, μ. У том случају, расподеле имају следећи облик:

- *Гумбелова расподела*:

$$G_{0,\mu,\sigma}(x) = e^{-e^{-\frac{x-\mu}{\sigma}}}, \text{ за } x \in \mathbb{R},$$

- *Фрешеова расподела*:

$$G_{1,\alpha,\mu,\sigma}(x) = \begin{cases} 0, & \text{за } x < \mu; \\ e^{-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}}, & \text{за } x \geq \mu, \end{cases}$$

- *Вејбулова расподела*:

$$G_{2,\alpha,\mu,\sigma}(x) = \begin{cases} e^{-\left(-\frac{x-\mu}{\sigma}\right)^{\alpha}}, & \text{за } x < \mu; \\ 1, & \text{за } x \geq \mu. \end{cases}$$

γ-параметризација

Расподеле екстремних вредности, иако су на први поглед различитог типа, имају заједничку особину

$$\lim_{\alpha\to\infty} G_{1,\alpha}(x) = G_0(x) \quad \text{и} \quad \lim_{\alpha\to\infty} G_{2,\alpha}(x) = G_0(x).$$

Ако се уведе смена $\gamma = 1/\alpha$ код Фрешеове расподеле и $\gamma = -1/\alpha$ код Вејбулове расподеле, нови модел, $G_\gamma(x)$, за који важи

$$\lim_{\gamma\to 0} G_\gamma(x) = G_0(x),$$

тј. непрекидан је по γ. Сада се све три функције расподеле екстремних вредности могу приказати као фамилија расподела која зависи од истог параметра облика, γ:

$$G_\gamma(x) = \begin{cases} e^{-e^{-x}}, & x \in \mathbb{R}, \gamma = 0 \quad \text{(Гумбелова расподела)} \\ e^{-(1+\gamma x)^{-\frac{1}{\gamma}}}, & x \geq -\frac{1}{\gamma}, \gamma > 0 \quad \text{(Фрешеова расподела)} \\ e^{-(1+\gamma x)^{-\frac{1}{\gamma}}}, & x < -\frac{1}{\gamma}, \gamma < 0 \quad \text{(Вејбулова расподела)} \end{cases}$$

Ова параметризацију увео је математичар фон Мизес [39].

Веза између α и γ - параметризације дата је са

$$G_\gamma(x) = \begin{cases} G_{1,\alpha,-\alpha,\alpha}(x), & \text{за } \gamma = \frac{1}{\alpha} > 0; \\ G_{2,\alpha,\alpha,\alpha}(x), & \text{за } \gamma = -\frac{1}{\alpha} < 0. \end{cases}$$

Густине расподела екстремних вредности, датих у γ-параметризацији, су:

$$g_\gamma(x) = \begin{cases} e^{-e^{-x}} e^{-x}, & x \in \mathbb{R}, \gamma = 0; \\ e^{-(1+\gamma x)^{-\frac{1}{\gamma}}} (1+\gamma x)^{-\left(1+\frac{1}{\gamma}\right)}, & x \geq -\frac{1}{\gamma}, \gamma > 0; \\ e^{-(1+\gamma x)^{-\frac{1}{\gamma}}} (1+\gamma x)^{-\left(1+\frac{1}{\gamma}\right)}, & x < -\frac{1}{\gamma}, \gamma < 0. \end{cases}$$

Моменти расподела екстремних вредности

Моменти Гумбелове, Фрешеове и Вејбулове расподеле могу се изразити у терминима Ојлерове Г-функције,

$$\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx, \quad t > 0. \tag{1.2}$$

Ако је X случајна величина са G_γ расподелом, онда су математичко очекивање и дисперзија случајне величине X задати изразима

$$\mathrm{E}(X) = m_{G_\gamma} = \frac{\Gamma(1-\gamma)-1}{\gamma}, \quad \gamma < 1, \tag{1.3}$$

$$\mathrm{var}_{G_\gamma} = \frac{\Gamma(1-2\gamma)-\Gamma^2(1-\gamma)}{\gamma^2}, \quad \gamma < \frac{1}{2}. \tag{1.4}$$

Моменти Гумбелове расподеле се добијају одређивањем граничне вредности десних страна једнакости (1.3) и (1.4) кад $\gamma \to 0$. Важе једнакости

$$m_{G_0} = \lambda, \quad \text{и} \quad \mathrm{var}_{G_0} = \frac{\pi^2}{6},$$

где је $\lambda = 0.577216\ldots$ Ојлерова константа.

Максимум стабилност расподела екстремних вредности

Важна карактеристика расподела екстремних вредности је да су оне максимум стабилне. Важе једнакости

$$G_0^n(x+\log n) = G_0(x), \quad \text{и} \quad G_\gamma^n\left(n^\gamma x - \frac{1}{\gamma}\right) = G_\gamma(x), \quad \gamma \neq 0, \tag{1.5}$$

тј. нормирајуће константе су $a_n = 1, b_n = \log n$ за Гумбелову расподелу, а $a_n = n^\gamma, b_n = \frac{n^\gamma - 1}{\gamma}$ за Фрешеову и Вејбулову расподелу, ако су оне дате у γ-параметризацији. Може се показати да су расподеле екстремних вредности једине непрекидне функције расподела са овом особином.

1.1.3 Теорема о екстремалним типовима и области привлачења за максимуме

Најважнији резултат оне гране теорије екстремних вредности која се бави максимумом је теорема о екстремалним типовима. Она даје карактеризацију граничне расподеле максимума низа независних и једнако расподељених случајних величина.

Теорема 1.1.1. [Gnedenko (1943), de Haan (1976)] *Нека је* (X_n), $n \in \mathbb{N}$, *низ независних случајних величина са заједничком функцијом расподеле* F *и* $M_n = \max\{X_1, X_2, \dots, X_n\}$. *Ако постоје нормирајуће константе* $a_n > 0$ *и* $b_n \in \mathbb{R}$, $n \in \mathbb{N}$, *такве да*

$$\lim_{n\to\infty} P\left\{\frac{M_n - b_n}{a_n} \le x\right\} = \lim_{n\to\infty} F^n(a_n x + b_n) = G(x),$$

где је G *недегенерисана функција расподеле, тада је функција* G *истог типа као једна од функција расподеле екстремних вредности.*

Избор нормирајућих константи одређује брзину конвергенције расподеле максимума ка граничној расподели. Нормирајуће константе нису јединствено одређене. Веза између граничних расподела максимума одређених различитим низовима нормирајућих константи добија се из следеће теореме.

Теорема 1.1.2. [Хинчин] *Нека је* (F_n), $n \in \mathbb{N}$, *низ функција расподеле за који постоји недегенерисана функција расподеле* G *и низови константи* $a_n > 0$ *и* $b_n \in \mathbb{R}$, $n \in \mathbb{N}$, *тако да важи*

$$\lim_{n\to\infty} F_n(a_n x + b_n) = G(x). \tag{1.6}$$

Нека је G^* *недегенерисана функција расподеле, а* $\alpha_n > 0$ *и* $\beta_n \in \mathbb{R}$, $n \in \mathbb{N}$, *низови константи.*
(а) *Ако важи*

$$\lim_{n\to\infty} F_n(\alpha_n x + \beta_n) = G^*(x), \tag{1.7}$$

онда постоје a *и* b, *тако да је*

$$\lim_{n\to\infty} \frac{\alpha_n}{a_n} = a > 0, \quad \lim_{n\to\infty} \frac{\beta_n - b_n}{a_n} = b \in \mathbb{R}, \tag{1.8}$$

$$G^*(x) = G(ax + b) \quad \text{за све} \quad x \in \mathbb{R}. \tag{1.9}$$

(б) *Ако важи* (1.8), *онда важи* (1.7) *и* (1.9).

Следећа теорема даје везу између понашања максимума низа и понашања репа расподеле појединачног члана низа.

Теорема 1.1.3. [Leadbetter, Lindgren, Rootzén (1983)] *Нека је* (X_n), $n \in \mathbb{N}$, *низ независних случајних величина са заједничком функцијом расподеле* F, $M_n = \max\{X_1, X_2, \ldots, X_n\}$, (u_n), $n \in \mathbb{N}$, *низ реалних бројева и* $0 \leq \tau \leq \infty$. *Тада једнакост*

$$\lim_{n\to\infty} P\{M_n \leq u_n\} = e^{-\tau}, \tag{1.10}$$

важи ако и само ако је

$$\lim_{n\to\infty} n(1 - F(u_n)) = \tau. \tag{1.11}$$

Теорема 1.1.4. [Leadbetter, Lindgren, Rootzén (1983)] *Нека је* (X_n), $n \in \mathbb{N}$, *низ независних случајних величина са заједничком функцијом расподеле* F, $M_n = \max\{X_1, X_2, \ldots, X_n\}$ *и* $0 < \tau < \infty$. *Тада низ* (u_n) *за који важи* (1.11) *постоји ако и само ако је*

$$\lim_{x\to x_0} \frac{F(x) - F(x-0)}{1 - F(x-0)} = 0. \tag{1.12}$$

За сваку од расподела екстремних вредности одређени су потребни и довољни услови да нека функција расподеле F припада њеној области привлачења за максимуме. Познато је да је све функције из области привлачења $D(G_{1,\alpha})$ десни крај носача расподеле једнак $+\infty$, за функције из области привлачења $D(G_{2,\alpha})$ је десни крај носача расподеле коначан број, а области привлачења $D(G_0)$ припадају функције и једног и другог типа. Карактеризације области привлачења сваке од расподела екстремних вредности изражене су у терминима Π - променљивих и Γ - променљивих функција, а одговарајуће нормирајуће константе у терминима уопштеног инверза функције $1/(1-F)$, где је F функција расподеле.

Дефиниција 1.1.1. (правилна променљивост у бесконачности) *Мерљива функција* $F : (0, +\infty) \to (0, +\infty)$ *је* **правилно променљива у бесконачности, са индексом** $\boldsymbol{\rho}$ $(F \in \Pi\Pi_\rho)$ *ако постоји* $\rho \in \mathbb{R}$, *такво да за свако* $x > 0$ *важи*

$$\lim_{t\to+\infty} \frac{F(tx)}{F(t)} = x^\rho.$$

Дефиниција 1.1.2. (правилна промењивост у коначној тачки) *Нека је* $F : (0, +\infty) \to (0, +\infty)$ *мерљива функција. Фукција* F *је* **правилно промењива у нули**, *ако је функција* $F\left(\frac{1}{x}\right)$ *правилно промењива у бесконачности. Функција* F *је* **правилно промењива у тачки** $\mathbf{x_0 > 0}$ *(при* $x \uparrow x_0$*) ако је функција* $F\left(x_0 - \frac{1}{x}\right)$ *правилно промењива у бесконачности.*

Појам правилне промењивости увео је Карамата у раду [26].

Дефиниција 1.1.3. (Π - промењивост) *Растућа функција* $F : (c, +\infty) \to \mathbb{R}$ *је* **Π - промењива** $(F \in \Pi)$ *ако постоје помоћне функције*

$$a : (c, +\infty) \to (0, +\infty), \quad b : (c, +\infty) \to \mathbb{R},$$

такве да за свако $x > 0$ *важи*

$$\lim_{t \to +\infty} \frac{F(tx) - b(t)}{a(t)} = \log x.$$

Дефиниција 1.1.4. (Γ - промењивост) *Растућа функција* $F : (c, x_0) \to \mathbb{R}$ *је* **Γ - промењива** $(F \in \Gamma)$ *ако важе услови:*
(а) $\lim\limits_{x \uparrow x_0} F(x) = +\infty$*;*
(б) *постоји функција* $g : (c, x_0) \to (0, +\infty)$*, таква да за свако* $x \in \mathbb{R}$ *важи*

$$\lim_{t \uparrow x_0} \frac{F(t + xg(t))}{F(t)} = e^x.$$

Класе Γ и Π увео је математичар де Хан [18].

Област привлачења Фрешеове расподеле

Следеће две теореме дају карактеризацију области привлачења Фрешеове расподеле.

Теорема 1.1.5. [von Mises (1936)] *Нека је* F *апсолутно непрекидна функција расподеле са густином расподеле* f*. Ако важе услови:*
(а) $f(x) > 0$ *за* $x \geq x_0$*;*
(б) $\lim\limits_{t \to \infty} \frac{tf(t)}{1-F(t)} = \alpha > 0$,
тада $F \in D(G_{1,\alpha})$.

Теорема 1.1.6. [Gnedenko (1943)] *Функција расподеле F припада области привлачења $D(G_{1,\alpha})$ ако и само ако важе услови:*
(а) $x_0 = \sup\{t : F(t) < 1\} = +\infty$;
(б) $1 - F \in \Pi\Pi_{-\alpha}$.
Нормирајуће константе су $a_n = \left(\frac{1}{1-F}\right)^{-1}(n)$ *и* $b_n = 0$.

Примери расподела које припадају Фрешеовој области привлачења су Фрешеова, Паретова и Кошијева расподела.

Област привлачења Вејбулове расподеле

Следеће две теореме дају карактеризацију области привлачења Вејбулове расподеле.

Теорема 1.1.7. [von Mises (1936)] *Нека је F апсолутно непрекидна функција расподеле са густином расподеле f. Ако важе услови:*
(а) $x_0 = \sup\{t : F(t) < 1\} < +\infty$;
(б) $f(x) > 0$ *за* $x \in (a, x_0)$;
(в) $\lim\limits_{t \uparrow x_0} \frac{(x_0 - t) f(t)}{1 - F(t)} = \alpha > 0$,
тада $F \in D(G_{2,\alpha})$.

Теорема 1.1.8. [Gnedenko (1943)] *Функција расподеле F припада области привлачења $D(G_{2,\alpha})$ ако и само ако важе услови:*
(а) $x_0 = \sup\{t : F(t) < 1\} < +\infty$;
(б) $1 - F\left(x_0 - \frac{1}{x}\right) \in \Pi\Pi_{-\alpha}$, $x \to +\infty$.
Нормирајуће константе су $a_n = x_0 - \left(\frac{1}{1-F}\right)^{-1}(n)$ *и* $b_n = x_0$.

Примери расподела из Вејбулове области привлачења су Вејбулова и равномерна расподела.

Област привлачења Гумбелове расподеле

Следеће теореме дају карактеризацију области привлачења Гумбелове расподеле.

Теорема 1.1.9. [von Mises (1936)] *Нека је F апсолутно непрекидна функција расподеле и $x_0 = \sup\{t : F(t) < 1\}$. Ако важе услови:*
(а) $F''(x) < 0$ *за све* $x \in (a, x_0)$, $x_0 \leq +\infty$;
(б) $F'(x) = 0$ *за* $x \geq x_0$;

(в) $\lim\limits_{t\to x_0}\frac{F''(t)(1-F(t))}{(F'(t))^2}=-1$,
тада важи $F\in D(G_0)$.

Теорема 1.1.10. [Gnedenko (1943), Mejzler (1949), de Haan (1970)] *Нека је* F *функција расподеле,* $x_0=\sup\{t:F(t)<1\}$ *и* $H(x)=\frac{1}{1-F(x)}$ *за* $x<x_0$. *Тада су еквивалентни следећи услови:*
(а) $F\in D(G_0)$;
(б) $H\in\Gamma$;
(в) $H^{-1}\in\Pi$.

Теорема 1.1.11. [de Haan (1970)] $F\in D(G_0)$ *ако и само ако важи*

$$\lim_{x\uparrow x_0}\frac{(1-F(x))\int_x^{x_0}\int_y^{x_0}(1-F(t))dtdy}{(\int_x^{x_0}(1-F(t))dt)^2}=1,$$

и сви записани интеграли су коначни. Тада важи $\frac{1}{1-F}\in\Gamma$, *а помоћна функција* g *се може одредити тако да важи*

$$g(t)=\frac{\int_x^{x_0}\int_y^{x_0}(1-F(t))dtdy}{\int_x^{x_0}(1-F(t))dt}$$

или

$$g(t)=\frac{\int_x^{x_0}(1-F(t))dt}{1-F(x)}.$$

Нормирајуће константе се могу тако одабрати да важи

$$a_n=F(b_n)\ \text{и}\ b_n=\left(\frac{1}{1-F}\right)^{-1}(n).$$

Теорема 1.1.12. [de Haan (1970)] $F\in D(G_0)$ *ако и само ако важи*

$$\lim_{x\uparrow x_0}\frac{\int_x^{x_0}(1-F(t))^\alpha dt}{(1-F(x))\int_x^{x_0}(1-F(t))^{\alpha-1}dt}=\frac{\alpha-1}{\alpha}$$

за неко $\alpha>1$.

Примери расподела из Гумбелове области привлачења су Гумбелова, нормална и експоненцијална расподела.

Из Теореме 1.1.4 следи да дискретна функција расподеле F не припада ниједној од области привлачења расподела екстремних вредности ако важи да низ

$$\frac{F(n) - F(n-0)}{1 - F(n-0)}, \quad n \in \mathbb{N}, \tag{1.13}$$

не конвергира ка нули кад $n \to \infty$. Примери таквих расподела су геометријска, негативна биномна и Пуасонова расподела са фиксираним параметрима.

Докази теорема из овог поглавља се могу наћи у књизи [31].

1.2 Прекорачења високог нивоа

1.2.1 Расподела прекорачења датог високог нивоа

Нека је дат случајни узорак $(X_1, X_2, \ldots, X_N)$, где су $X_1, X_2, \ldots, X_N$ независне копије случајне величине X са функцијом расподеле F и нека је одређен високи ниво, u. По једној од дефиниција, екстремне вредности у узорку $(X_1, X_2, \ldots, X_N)$ су прекорачења нивоа u, тј. чланови узорка $X_1, X_2, \ldots, X_n$ већи од u. Аналогно, могу се посматрати исте те величине умањене за вредност u, тј.

$$Y_1 = X_1 - u, Y_2 = X_2 - u, \ldots, Y_n = X_n - u.$$

Променом нивоа (прага) u, мењају се број и величина прекорачења тог прага. Број прекорачења, $N_u = \sum_{i \leq N} I\{X_i > u\}$, је случајна величина која има биномну расподелу са параметрима n и p, где је $p = 1 - F(u)$. Средња вредност броја прекорачења датог прага u је онда $E(N_u) = n(1 - F(u))$.

Појму расподеле максимума низа независних и једнако расподељених случајних величина у овом приступу одговара **функција расподеле прекорачења нивоа u**, $F^{[u]}(x)$, задата са

$$F^{[u]}(x) = P\{X \leq x \mid X > u\} = \frac{P\{X \leq x, X > u\}}{P\{X > u\}} = \frac{F(x) - F(u)}{1 - F(u)}, \quad x \geq u, \tag{1.14}$$

или, у мало другачијем облику,

$$F^{(u)}(x) = P\{X \leq x + u \mid X > u\} = \frac{F(x + u) - F(u)}{1 - F(u)}, \quad x \geq 0. \tag{1.15}$$

Математичко очекивање величине прекорачења датог нивоа u, за дату функцију расподеле F, се назива **средње прекорачење**:

$$e_F(u) = E(X - u \mid X > u).$$

Појмовима максимум стабилности и граничне расподеле максимума сада одговарају појмови стабилности у односу на прекорачења датог нивоа и граничне расподеле прекорачења високог нивоа.

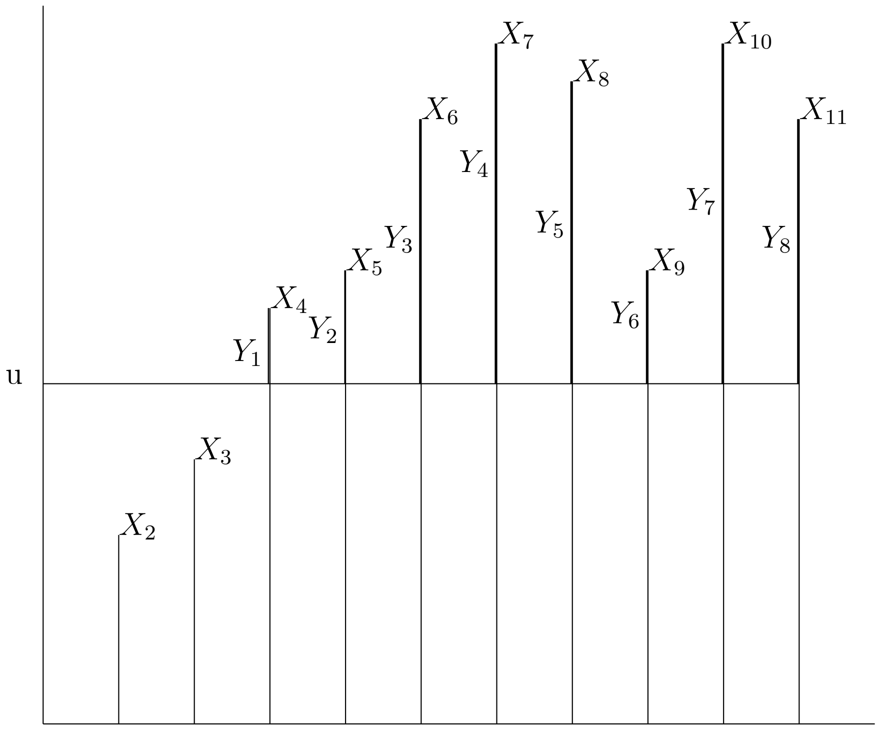
За недегенерисану функцију расподеле F се каже да је **стабилна у односу на прекорачења датог нивоа** ако за сваки задати праг u, $u < \varpi(F)$, где је $\varpi(F)$ десни крај носача расподеле, постоје нормирајуће константе $a_u > 0$ и $b_u \in \mathbb{R}$, такве да за свако $x \in \mathbb{R}$ важи

$$F^{[u]}(a_u x + b_u) = F(x).$$

За функцију расподеле F се каже да припада **области привлачења за прекорачења високог нивоа** недегенерисане функције расподеле G ($F \in D_p(G)$) ако постоје нормирајуће константе $a_u > 0$ и $b_u \in \mathbb{R}$, такве да важи

$$\lim_{u \to \varpi(F)} F^{[u]}(a_u x + b_u) = G(x).$$

У том случају се функција G назива **гранична расподела прекорачења високог нивоа** за функцију F.



Слика 1.1: Прекорачења високог нивоа

1.2.2 Експоненцијална, Паретова и бета расподела

У грани теорије екстремних вредности која се бави прекорачењима високог нивоа важну улогу игра фамилија **генералисаних Паретових расподела**, чији су стандардни представници (у α - параметризацији):

- *експоненцијална расподела*:

$$W_0(x) = \begin{cases} 0, & \text{за } x < 0; \\ 1 - e^{-x}, & \text{за } x \geq 0, \end{cases}$$

- *Паретова расподела с параметром* $\alpha > 0$:

$$W_{1,\alpha}(x) = \begin{cases} 0, & \text{за } x < 1; \\ 1 - x^{-\alpha}, & \text{за } x \geq 1, \end{cases}$$

- *бета расподела с параметром* $\alpha > 0$:

$$W_{2,\alpha}(x) = \begin{cases} 0, & \text{за } x < -1; \\ 1 - (-x)^{\alpha}, & \text{за } -1 \leq x \leq 0; \\ 1, & \text{за } x > 0. \end{cases}$$

Код генералисаних Паретових расподела се могу увести параметри положаја и размере на исти начин као код расподела екстремних вредности.

Слично као у случају расподела екстремних вредности, сменом $\gamma = 1/\alpha$ код Паретове расподеле и $\gamma = -1/\alpha$ код бета расподеле, добија се фамилија генералисаних Паретових расподела задата у γ-параметризацији:

$$W_\gamma(x) = \begin{cases} 1 - e^{-x}, & x \geq 0, \gamma = 0 \qquad \text{(експоненцијална расп.)} \\ 1 - (1 + \gamma x)^{-\frac{1}{\gamma}}, & x \geq 0, \gamma > 0 \qquad \text{(Паретова расп.)} \\ 1 - (1 + \gamma x)^{-\frac{1}{\gamma}}, & x \in \left[0, -\frac{1}{\gamma}\right], \gamma < 0 \qquad \text{(бета расп.)} \end{cases}$$

За овај модел важи

$$\lim_{\gamma \to 0} W_\gamma(x) = W_0(x),$$

тј. непрекидан је по γ, а веза између једне и друге параметризације дата је са

$$W_\gamma(x) = \begin{cases} W_{1,\alpha,-\alpha,\alpha}(x), & \text{за } \gamma = \frac{1}{\alpha} > 0; \\ W_{2,\alpha,\alpha,\alpha}(x), & \text{за } \gamma = -\frac{1}{\alpha} < 0. \end{cases}$$

Густине генералисаних Паретових расподела, датих у γ-параметризацији, су:

$$w_\gamma(x) = \begin{cases} e^{-x}, & x \geq 0, \gamma = 0; \\ (1+\gamma x)^{-\left(1+\frac{1}{\gamma}\right)}, & x \geq 0, \gamma > 0; \\ (1+\gamma x)^{-\left(1+\frac{1}{\gamma}\right)}, & x \in \left[0, -\frac{1}{\gamma}\right], \gamma < 0. \end{cases}$$

Веза између фамилије расподела екстремних вредности и фамилије генералисаних Паретових расподела је

$$W_\gamma(x) = 1 + \log G_\gamma(x), \text{ у случају кад је } \log G_\gamma(x) > -1.$$

Моменти генералисаних Паретових расподела

Моменти случајне величине X са W_γ расподелом дати су са

$$\mathrm{E}\left[(1+\gamma X)^r\right] = \frac{1}{1-\gamma r}, \quad 1 - \gamma r > 0, \tag{1.16}$$

одакле следи

$$\mathrm{E}(X^r) = \frac{r!}{\gamma^{r+1}} \frac{\Gamma\left(\frac{1}{\gamma} - r\right)}{\Gamma\left(1 + \frac{1}{\gamma}\right)}, \quad \gamma < \frac{1}{r}. \tag{1.17}$$

Математичко очекивање и дисперзија случајне величине X добијају се из (1.17) за $r = 1$ и $r = 2$ и задати су изразима

$$\mathrm{E}(X) = m_{W_\gamma} = \frac{1}{1-\gamma}, \quad \gamma < 1, \tag{1.18}$$

$$\mathrm{var}_{W_\gamma} = \frac{1}{(1-\gamma)^2(1-2\gamma)}, \quad \gamma < \frac{1}{2}. \tag{1.19}$$

Математичко очекивање и дисперзија случајне величине са експоненцијалном расподелом се добијају одређивањем граничне вредности десних страна једнакости (1.18) и (1.19) кад $\gamma \to 0$. У овом случају се добијају познати резултати:

$$m_{W_0} = 1 \quad \text{и } \mathrm{var}_{W_0} = 1.$$

Средње прекорачење за генералисане Паретове расподеле

Једна од карактеристика фамилије генералисаних Паретових расподела је да им је средње прекорачење линеарна функција. Ова особина игра важну улогу при моделовању конкретних узорака овим типом расподела, када је потребно одредити вредност високог нивоа изнад кога је ова апроксимација адекватна.

- *средње прекорачење у γ-параметризацији:*

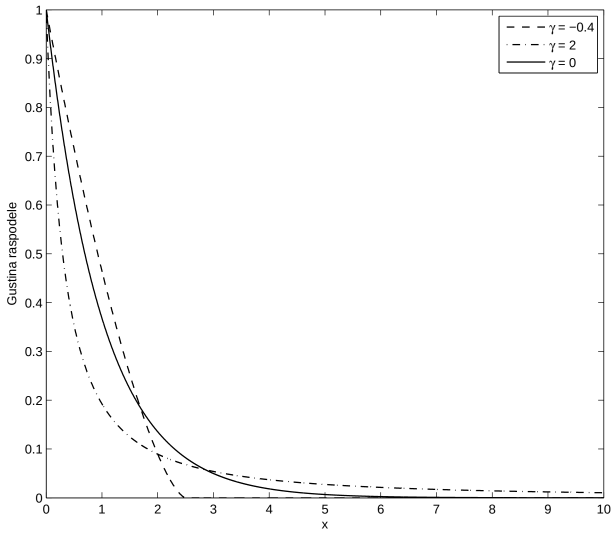
$$e_{W_\gamma}(u) = \begin{cases} 1, & \gamma = 0, u > 0; \\ \frac{1}{1-\gamma} + \frac{\gamma}{1-\gamma}u, & 0 < \gamma < 1, u > 0; \\ \frac{1}{1-\gamma} + \frac{\gamma}{1-\gamma}u, & \gamma < 0, 0 < u < -\frac{1}{\gamma}. \end{cases}$$

Стабилност у односу на прекорачења високог нивоа

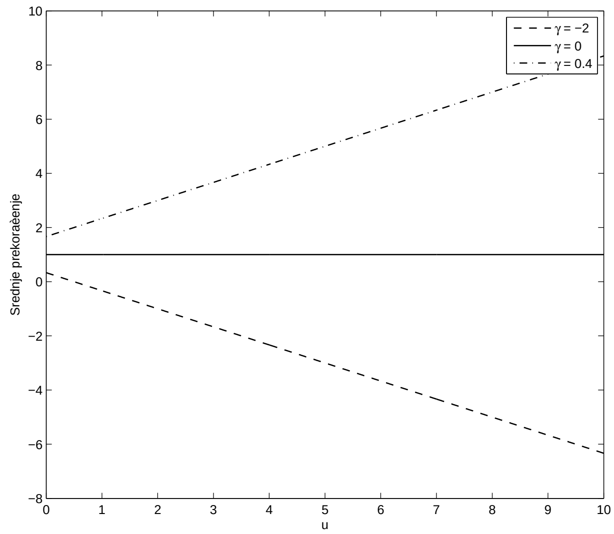
Још једна важна карактеристика генералисаних Паретових расподела је да су оне стабилне у односу на прекорачења високог нивоа. Важе једнакости

$$W_0^{(u)}(x) = W_0(x) \quad \text{и} \quad W_\gamma^{(u)}((1+\gamma u)x) = W_\gamma(x), \quad \gamma \neq 0, \tag{1.20}$$

тј. нормирајуће константе су $a_u = 1, b_u = 0$ за експоненцијалну расподелу, а $a_u = 1 + \gamma u, b_u = 0$ за Паретову и бета расподелу, ако су оне дате у γ-параметризацији. Може се показати да су генералисане Паретове расподеле једине непрекидне функције расподела са овом особином.



Слика 1.2: Густине генералисаних Паретових расподела са различитим вредностима параметра облика



Слика 1.3: Средње прекорачење за генералисане Паретове расподеле са различитим вредностима параметра облика

1.2.3 Граничне расподеле прекорачења високог нивоа и области привлачења

Нека је X случајна величина са функцијом расподеле F и $\varpi(F)$ десни крај носача те расподеле. Следеће теореме дају типове граничних расподела прекорачења високог нивоа за дату функцију расподеле F.

Теорема 1.2.1. [Balkema, de Haan (1974)] *Нека је F функција расподеле. Ако постоје нормирајуће константе $a_u > 0$ и $b_u \in \mathbb{R}$, такве да важи*

$$\lim_{u \to \varpi(F)} F^{[u]}(a_u x + b_u) = G(x),$$

за недегенерисану функцију расподеле G, тада је функција G истог типа као једна од следећих функција расподеле:

$$\begin{aligned} W_0(x) &= 1 - e^{-x}, \quad x \geq 0, \\ W_{1,\alpha}(x) &= 1 - x^{-\alpha}, \quad x \geq 1, \\ T_\gamma(x) &= 1 - e^{-\gamma[1+x]}, \quad x \geq 0, \\ Q_{\gamma,\alpha}(x) &= 1 - e^{-\gamma[1+\alpha \log(1+x)]}, \quad x \geq 0, \end{aligned}$$

за $\alpha, \gamma > 0$.

Под одређеним условима, функција расподеле прекорачења нивоа u, за дату функцију расподеле F, може се произвољно тачно апроксимирати једном од генералисаних Паретових расподела. Ту се као једна могућност појављује и бета расподела, иако не спада у граничне расподеле из претходне теореме.

Теорема 1.2.2. [Balkema, de Haan (1974)] *Нека је F апсолутно непрекидна функција расподеле, а γ и δ реални бројеви. Ако важе услови:*
(а) *$f = F'$ је позитивна, диференцијабилна функција на интервалу $[u_0, +\infty)$;*
(б) *$\gamma \leq \frac{d}{du}\left(\frac{1-F(u)}{f(u)}\right) \leq \delta$ за $u \geq u_0$,*
тада важи

$$W_\gamma(x) \leq F^{[u]}(a(u)x) \leq W_\delta(x), \quad \text{где је} \quad a(u) = \frac{1-F(u)}{f(u)}.$$

Теорема 1.2.3. [Balkema, de Haan (1974)] *Нека је F функција расподеле. Ако важи*

$$\lim_{u \to \varpi(F)} F^{[u]}(a_u x + b_u) = G(x),$$

за непрекидну функцију расподеле G и неке нормирајуће константе $a_u > 0$ и $b_u \in \mathbb{R}$, тада је функција G истог типа као једна од функција расподеле W_γ, $\gamma \in \mathbb{R}$.

Области привлачења граничних расподела прекорачења високог нивоа

Следеће две теореме дају карактеризације области привлачења сваке од граничних расподела прекорачења високог нивоа. Показује се да се области привлачења за максимуме сваке од расподела екстремних вредности поклапају са областима привлачења за прекорачења оне генералисане Паретове расподеле која има исти параметар облика.

Теорема 1.2.4. [Balkema, de Haan (1974)] *Ако је D_0 скуп функција расподеле F, таквих да $F(x) < 1$ за свако x, онда важи*

$$D_p(W_0) = D(G_0) \cap D_0.$$

Теорема 1.2.5. [Balkema, de Haan (1974)] *$D_p(W_{1,\alpha}) = D(G_{1,\alpha})$, за свако $\alpha > 0$.*

Докази теорема из овог поглавља дати су у раду [2].

1.3 Екстремне вредности у стационарним низовима

У многим ситуацијама, нарочито у применама, није реално претпоставити да су чланови посматраног низа независне случајне величине. Уместо тога, услов независности се замењује слабијим, условом стационарности, тј. претпоставља се да се функција расподеле чланова низа не мења идући кроз низ. Прецизније:

Дефиниција 1.3.1. *Случајан низ* (X_n), $n \in \mathbb{N}$, *је* **строго стационаран**, *ако за све природне бројеве* n *и* k *случајни вектори*

$$(X_1, X_2, \ldots, X_n) \text{ и } (X_{1+k}, X_{2+k}, \ldots, X_{n+k})$$

имају исту расподелу.

1.3.1 Услови слабе зависности

Услови слабе зависности су начин да се "измери" зависност чланова посматраног низа. Уколико је та зависност мала ако су чланови низа довољно удаљени, могу се користити исте технике као за случај независности. Постоји више начина да се формулишу услови слабе зависности. Неки од њих су:

Дефиниција 1.3.2. [Rosenblat (1956)] *Нека је* $\mathcal{M}_a^b$ σ*-алгебра генерисана случајним величинама* X_t, $a \leq t \leq b$. *Строго стационаран низ* $(X_n), n \in \mathbb{N}$, *задовољава* **услов јаке помешаности** *ако постоји низ* $\alpha(k)$ **(коефицијент помешаности)**, *такав да* $\alpha(k) \to 0$ *при* $k \to \infty$ *и да неједнакост*

$$\sup_{A \in \mathcal{M}_1^t, B \in \mathcal{M}_{t+k}^{+\infty}} |P(AB) - P(A)P(B)| \leq \alpha(k),$$

важи за све природне бројеве t *и* k.

Дефиниција 1.3.3. [Ибрагимов (1962)] *Строго стационаран низ* $(X_n), n \in \mathbb{N}$, *задовољава* **услов равномерне јаке помешаности** *ако постоји низ* $\beta(k)$ **(коефицијент равномерне јаке помешаности)**, *такав да* $\beta(k) \to 0$ *при* $k \to \infty$ *и да неједнакост*

$$\sup_{A \in \mathcal{M}_1^t, P(A) > 0, B \in \mathcal{M}_{t+k}^{+\infty}} |P(B \mid A) - P(B)| \leq \beta(k),$$

важи за све природне бројеве t и k.

При раду са екстремним вредностима важни су догађаји облика $P\{X \leq x\}$, где је X нека случајна величина. Зато се често користе услови слабе зависности формулисани помоћу догађаја овог типа.

Дефиниција 1.3.4. *Строго стационаран низ $(X_n), n \in \mathbb{N}$, задовољава* **услов D** *ако за све природне бројеве*

$$1 \leq j_1 < \cdots < j_k < j_{k+1} < \cdots < j_{k+l} \leq n,$$

где је $j_{k+1} - j_k \geq l$, и сваки реалан број и важи неједнакост

$$\left| P\left(\bigcap_{s=1}^{k+l}\{X_{j_s} \leq u\}\right) - P\left(\bigcap_{s=1}^{k}\{X_{j_s} \leq u\}\right) P\left(\bigcap_{s=k+1}^{k+l}\{X_{j_s} \leq u\}\right)\right| \leq d(l),$$

и $d(l) \to 0$ при $l \to \infty$.

Услов D омогућава да се теорема о екстремалним типовима, која важи за низове независних случајних величина пренесе и на стационарне низове. Међутим, овај услов се може и ослабити, на следећи начин.

Дефиниција 1.3.5. [Leadbetter (1974)] *Нека је $(X_n), n \in \mathbb{N}$, строго стационаран случајан низ, а $(u_n), n \in \mathbb{N}$, низ реалних бројева.* **Услов $\mathbf{D(u_n)}$** *је задовољен ако за све природне бројеве*

$$1 \leq j_1 < \cdots < j_k < j_{k+1} < \cdots < j_{k+l} \leq n,$$

где је $j_{k+1} - j_k \geq l$, важи неједнакост

$$\left| P\left(\bigcap_{s=1}^{k+l}\{X_{j_s} \leq u_n\}\right) - P\left(\bigcap_{s=1}^{k}\{X_{j_s} \leq u_n\}\right) P\left(\bigcap_{s=k+1}^{k+l}\{X_{j_s} \leq u_n\}\right)\right| \leq \alpha_{n,l}$$

и $\alpha_{n,l_n} \to 0$ при $n \to \infty$ за неки низ $l_n = o(n)$.

Да би се описало понашање репа расподеле појединачног члана стационарног низа потребан је следећи, мало јачи, услов.

Дефиниција 1.3.6. [Loynes (1965)] *Нека је $(X_n), n \in \mathbb{N}$, строго стационаран случајан низ, а $(u_n), n \in \mathbb{N}$, низ реалних бројева.* **Услов $\mathbf{D'(u_n)}$** *је задовољен ако је*

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \cdot \sum_{j=2}^{[n/k]} P\{X_1 > u_n, X_j > u_n\} = 0.$$

1.3.2 Екстремални типови за стационарне низове

Следећа теорема даје могуће типове граничних расподела максимума за стационарне случајне низове.

Теорема 1.3.1. [Leadbetter (1974)] *Нека је* $(X_n), n \in \mathbb{N}$, *строго стационаран случајан низ,* $a_n > 0$ *и* $b_n \in \mathbb{R}$, *где је* $n \in \mathbb{N}$, *низови константи и* G *недегенерисана функција расподеле, тако да за сваку тачку непрекидности* x *функције* G *важи*

$$\lim_{n \to \infty} P\{M_n \leq a_n x + b_n\} = G(x).$$

Ако за сваки реалан број x *и* $u_n = a_n x + b_n$ *важи услов* $D(u_n)$, *онда је* G *функција расподеле екстремних вредности.*

Веза између граничне расподеле максимума стационарног низа и асимптотског понашања репа расподеле појединачног члана тог низа дата је следећом теоремом.

Теорема 1.3.2. [Leadbetter, Lindgren, Rootzén (1983)] *Нека је* (X_n), $n \in \mathbb{N}$, *строго стационаран случајан низ,* $M_n = \max\{X_1, X_2, \ldots, X_n\}$, $F(x) = P\{X_1 \leq x\}$, $(u_n), n \in \mathbb{N}$, *низ реалних бројева такав да важе услови* $D(u_n)$ *и* $D'(u_n)$ *и* $0 \leq \tau \leq \infty$. *Тада једнакост*

$$\lim_{n \to \infty} P\{M_n \leq u_n\} = e^{-\tau},$$

важи ако и само ако је

$$\lim_{n \to \infty} n(1 - F(u_n)) = \tau.$$

1.3.3 Пратећи низ независних случајних величина

Сваком строго стационарном случајном низу може се доделити један низ независних случајних величина на следећи начин:

Дефиниција 1.3.7. *Нека је* $(X_n), n \in \mathbb{N}$, *строго стационаран случајан низ, и* $F(x) = P\{X_n \leq x\}$ *заједничка функција расподеле чланова тог низа. Низ* $(X_n^*), n \in \mathbb{N}$, *независних случајних величина са заједничком функцијом расподеле* F *зове се* **пратећи низ независних случајних величина**.

Следећа теорема даје везу између граничне расподеле максимума основног низа и граничне расподеле максимума одговарајућег пратећег низа независних случајних величина.

Теорема 1.3.3. [Leadbetter (1974)] *Нека је* $(X_n), n \in \mathbb{N}$, *строго стационаран случајан низ,* $a_n > 0$, $b_n \in \mathbb{R}$, $u_n = a_n x + b_n$ *низови константи и* $M_n = \max\{X_1, X_2, \dots, X_n\}$. *Претпоставимо да за сваки реалан број* x *важе услови* $D(u_n)$ *и* $D'(u_n)$. *Нека је* $(X_n^*), n \in \mathbb{N}$, *пратећи низ независних случајних величина и* $M_n^* = \max\{X_1^*, X_2^*, \dots, X_n^*\}$. *Тада једнакост*

$$\lim_{n\to\infty} P\{M_n \leq a_n x + b_n\} = G(x),$$

важи за неку недегенерисану функцију расподеле G *и сваку тачку непрекидности* x *функције* G *ако и само ако је*

$$\lim_{n\to\infty} P\{M_n^* \leq a_n x + b_n\} = G(x).$$

Докази теорема из овог поглавља дати су у књигама [28] и [31].

Глава 2

Комбинаторни проблеми чекања

2.1 Вероватносни методи у комбинаторним проблемима

Многи комбинаторни проблеми у вероватноћи могу се најбоље описати преко такозваних **модела урне** (*urn models*). Проблеми тог типа су актуелни већ неколико деценија, а неке од класичних референци из те области су Feller (1971) [13] и Johnson and Kotz (1977) [24].

Основна поставка проблема је следећа:
У кутији (урни) се налази r различитих куглица. Куглице се извлаче са враћањем, док се не оствари неки унапред дефинисани догађај (*stopping rule*).

Постоје разне варијанте овог проблема, које се разликују по томе како се дефинисан крај експеримента и по величини која се анализира. Нека од могућих питања која се постављају су:

- Колико различитих куглица ће бити извучено после n извлачења? Ово је у литератури познато као *classical occupancy problem*.
- Колико најмање извлачења је потребно да би се извукло k различитих куглица? Ово је у литератури познато као *coupon collector's problem*.

- Колико најмање извлачења је потребно да свих r куглица буде извучено бар по $m > 1$ пута? Ово је у литератури познато као *dixie cup problem*.

- Колико најмање извлачења је потребно да било која куглица буде извучена m пута? Ово је у литератури познато као *birthday problem*.

Оним проблемима овог типа који се односе на остваривање свих догађаја $E_1, E_2, \dots, E_n$ из неке колекције може се приступити на више начина. Неке од техника су:

1. Време чекања (број изведених експеримената) S_n до остварења свих догађаја $E_1, E_2, \dots, E_n$ се представља у облику $S_n = V_1 + V_2 + \cdots + V_n$, где су $V_i, i \in \mathbb{N}$, времена чекања између остварења узастопних догађаја из датог скупа. Ако је од интереса гранична расподела случајне величине S_n, користи се теорија сумирања случајних величина.

2. Ако се посматра број догађаја из дате колекције остварених у неком интервалу, што је исто тако једна случајна величина, може се користити теорија тачкастих случајних процеса. Ако се посматрани догађаји остварују релативно ретко, користи се метод утапања у Пуасонове процесе.

3. Време чекања до остварења свих догађаја $E_1, E_2, \dots, E_n$ се представља у облику $S_n = \max\{X_1, X_2, \dots X_n\}$, где је X_i време чекања до остварења догађаја E_i. У овом случају користе се методе теорије екстремних вредности које се односе на расподелу максимума низова независних или стационарних случајних величина.

Даље излагање односи се само на трећи приступ.

Такође, у литератури се разматрају уопштења проблема урне, као и бројне примене. Неки од радова на ту тему су Samuel-Cahn (1974) [38], Dawkins (1991) [6], Boneh and Papanicolau (1996) [4] и Ross (2007) [37].

2.2 Проблем скупљања купона и уопштења

У овом поглављу разматра се проблем скупљања купона који је задат на следећи начин: *бирају се елементи са враћањем из скупа* $\mathbb{N}_n = \{1, 2, \ldots, n\}$, *под претпоставком да сваки елемент има вероватноћу* $1/n$ *да буде изабран и посматра се време чекања* M_n *док сви елементи скупа* $\mathbb{N}_n$ *или неке друге шеме не буду изабрани.* Овај основни проблем има више варијанти и уопштења.

Случај 1а. У радовима Erdös, Rényi (1961) [12] и Mladenović (1999) [30] разматран је проблем скупљања купона у оригиналној формулацији. Ту је M_n време чекања до појаве свих елемената из скупа $\mathbb{N}_n$ и важи $M_n = \max\{X_{n1}, X_{n2}, \ldots, X_{nn}\}$, где је за свако $j \in \mathbb{N}_n$, X_{nj} време чекања до појаве елемента j у низу бирања из $\mathbb{N}_n$. X_{nj} је случајна величина са геометријском расподелом с параметром $1/n$,

$$P\{X_{nj} = k\} = \frac{1}{n}\left(1 - \frac{1}{n}\right)^{k-1}, \quad k \in \{1, 2, \ldots\}. \tag{2.1}$$

Гранична расподела максимума M_n, као и брзина конвергенције, одређене су у следећој теореми.

Теорема 2.2.1. (а)[Erdös, Rényi (1961)] *За сваки реалан број* x *важи једнакост*

$$\lim_{n\to\infty} P\{M_n \le n(x + \log n)\} = \exp(-e^{-x}). \tag{2.2}$$

(б)[Mladenović (1999)] *За сваки реалан број* x *при* $n \to \infty$ *важи*

$$|P\{M_n \le n(x + \log n)\} - \exp(-e^{-x})| = O\left(\frac{\log n}{n}\right). \tag{2.3}$$

Једно уопштење овог проблема састоји се у томе да се, уместо времена чекања до појаве свих елемената, посматра време чекања до појаве свих елемената неког подскупа $A \subset \mathbb{N}_n$. Овај проблем је решен у раду Baum, Billingsley (1965) [3].

Случај 1б. Такође, проблем се може уопштити у другом правцу, тако што се захтева да се сваки елемент скупа $\mathbb{N}_n$ појави одређен број пута, тј. M_n је време чекања до појаве свих елемената из

$\mathbb{N}_n$ бар r пута, где је r природан број. Ова верзија проблема је решена у радовима Holst (1986) [19] и Mladenović (2002) [32]. У овом случају је $M_n = \max\{X_{n1}, X_{n2}, \dots, X_{nn}\}$, где зависне случајне величине $X_{n1}, X_{n2}, \dots, X_{nn}$ имају исту негативну биномну расподелу с параметрима r и $1/n$, тј. за сваки $j \in \mathbb{N}_n$,

$$P\{X_{nj} = k\} = \binom{k-1}{r-1}\left(1 - \frac{1}{n}\right)^{k-r} \frac{1}{n^r}, \quad k \in \{r, r+1, \dots\}. \tag{2.4}$$

У раду [32] добијена је функција расподеле случајне величине M_n:

$$P\{M_n \leq k\} = \sum_{m=0}^{n} (-1)^m \binom{n}{m} \sum_{s_m=0}^{r-1} \cdots \sum_{s_2=0}^{r-1} \sum_{s_1=0}^{r-1} p_{nm}(k, s_1, s_2, \dots, s_m), \tag{2.5}$$

где је

$$\begin{aligned} p_{nm}(k, s_1, s_2, \dots, s_m) \quad = \quad & \binom{k}{s_1}\binom{k-s_1}{s_2}\dots\binom{k-s_s-s_2-\dots-s_{m-1}}{s_m} \times \\ & \times \left(\frac{n-m}{n}\right)^{k-s_1-s_2-\dots-s_m} \left(\frac{1}{n}\right)^{s_1+s_2+\dots+s_m}. \end{aligned} \tag{2.6}$$

Једнакост (2.5) важи за $k \geq nr$, а (2.6) за $m \in \{0, 1, \dots, n\}$. Гранична расподела максимума M_n, као и брзина конвергенције, одређене су у следећој теореми.

Теорема 2.2.2. *Нека је*

$$a_n = \log n + (r-1)\log\log n - \log(r-1)!, \quad n \geq 1.$$

(а)[Holst (1986)] *Гранична расподела случајне величине M_n дата је са*

$$\lim_{n\to\infty} P\{M_n \leq n(x + a_n)\} = \exp(-e^{-x}). \tag{2.7}$$

(б)[Mladenović (2002)] *За $r \geq 2$, при $n \to \infty$ важи*

$$P\{M_n \leq n(x+a_n)\} - \exp(-e^{-x}) \sim -(r-1)^2 e^{-x} \exp(-e^{-x}) \frac{\log\log n}{\log n}. \tag{2.8}$$

Случај 2. Проблем скупљања купона се може уопштити на још један начин. Уместо појаве појединачних елемената скупа $\mathbb{N}_n$, може се захтевати да се појаве сви парови (тројке, четворке, m-торке...) елемената. Проблем чекања за парове решен је у раду Mladenović (2008) [33].

Нека је M_n време чекања до појаве свих парова jj, $j \in \mathbb{N}_n$. Нека је $(Z_n), n \in \mathbb{N}$, низ независних случајних величина са равномерном расподелом на скупу $\mathbb{N}_n$ и

$$X_{nj} = \min\{k : Z_{k-1} = Z_k = j\}, \quad j \in \mathbb{N}_n \text{ је фиксиран број.} \tag{2.9}$$

У овом случају $M_n = \max\{X_{n1}, X_{n2}, \ldots, X_{nn}\}$.

Нека су случајне величине $Y_{nn}, Y_{n,n-1}, \ldots, Y_{n1}$ дефинисане на следећи начин: Y_{nn} је време чекања до појаве првог пара $j_1 j_1$, где је $j_1 \in \mathbb{N}_n$; $Y_{n,n-1}$ је време чекања до појаве другог пара $j_2 j_2$, $j_2 \in \mathbb{N}_n \setminus \{j_1\}$, после појаве првог пара $j_1 j_1$; и тако даље. Тада за свако $j \in \mathbb{N}_n$ и за сваки природан број k важи једнакост

$$P\{X_{n1} > k, X_{n2} > k, \ldots, X_{nj} > k\} = P\{Y_{nj} > k\}. \tag{2.10}$$

Расподеле вероватноћа случајних величина Y_{nj}, $j \in \mathbb{N}_n$, одређене су у Теореми 2 у раду Mladenović (2008). Коришћењем принципа укључивања и искључивања, добијено је

$$P\{M_n \leq k\} = \sum_{m=0}^{n} (-1)^m \binom{n}{m} P\{Y_{nm} > k\}, \tag{2.11}$$

где је

$$\begin{aligned} P\{Y_{nm} > k\} &= \frac{m}{n\sqrt{(n+1)^2 - 4m}} \times \\ &\quad \times \left\{ \left(\frac{t_1(m)}{n}\right)^k \frac{n}{n - t_1(m)} - \left(\frac{t_2(m)}{n}\right)^k \frac{n}{n - t_2(m)} \right\}, \end{aligned} \tag{2.12}$$

$$t_1(m) = \frac{n - 1 + \sqrt{(n+1)^2 - 4m}}{2}, \tag{2.13}$$

$$t_2(m) = \frac{n - 1 - \sqrt{(n+1)^2 - 4m}}{2}. \tag{2.14}$$

Зависне случајне величине $X_{n1}, X_{n2}, \ldots, X_{nn}$ имају исту расподелу, која се добија из следеће теореме.

Теорема 2.2.3. [Mladenović (2008)] (а) *Расподела случајне величине* X_{nj} *дата је са*

$$P\{X_{nj}=k\}=\sum_{s=0}^{[k/2]-1}\binom{k-s-2}{s}\left(1-\frac{1}{n}\right)^{k-s-2}\frac{1}{n^{s+2}},\quad k\in\{2,3,\dots\}. \tag{2.15}$$

(б) *Ако је* $u_n=n^2(x+\log n)$, *онда важи једнакост*

$$\lim_{n\to\infty} n(1-F_n(u_n))=\lim_{n\to\infty} nP\{X_{nj}>u_n\}=e^{-x}. \tag{2.16}$$

Важи и следећи резултат.

Лема 2.2.1. [Mladenović (2008)] *Нека су* X_{nj}^*, $j\in\{1,2,\dots,n\}$ *независне случајне величине са расподелом*

$$P\{X_{nj}^*=k\}=\sum_{s=0}^{[k/2]-1}\binom{k-s-2}{s}\left(1-\frac{1}{n}\right)^{k-s-2}\frac{1}{n^{s+2}},\quad k\in\{2,3,\dots\}, \tag{2.17}$$

и нека је $M_n^*=\max\{X_{n1}^*,X_{n2}^*,\dots,X_{nn}^*\}$. *Тада за свако реално x важи једнакост:*

$$\lim_{n\to\infty} P\{M_n^*\le n^2(x+\log n)\}=\exp(-e^{-x}). \tag{2.18}$$

У истом раду изведена је гранична расподела k-тог максимума случајне величине M_n, која је дата у следећој теореми.

Теорема 2.2.4. [Mladenović (2008)] *Ако је*

$$S_{n,n-k+1}=\max\{X_{n1},X_{n2},\dots,X_{n,n-k+1}\},\quad k\in\mathbb{N},$$

онда је гранична расподела случајне величине $S_{n,n-k+1}=M_n^{(k)}$ *дата са*

$$\lim_{n\to\infty} P\{M_n^{(k)}\le n^2(x+\log n)\}=\exp(-e^{-x})\sum_{s=0}^{k-1}\frac{e^{-sx}}{s!}. \tag{2.19}$$

Специјално, гранична расподела максимума $M_n=S_{nn}$ *је Гумбелова расподела.*

Следећа теорема даје нови резултат везан за брзину конвергенције максимума M_n ка граничној расподели.

Теорема 2.2.5. *За свако реално x при $n \to \infty$ важи*

$$\left|P\{M_n \le n^2(x+\log n)\} - \exp(-e^{-x})\right| = O\left(\frac{\log n}{n}\right). \tag{2.20}$$

Доказ Теореме 2.2.5 Нека је $u_n = n^2(x+\log n)$, $[u_n]$ цео део тог броја и $r_n = u_n - [u_n]$. Тада важи

$$\begin{aligned}|P\{M_n \le [u_n]\} - \exp(-e^{-x})| &= |(P\{M_n \le [u_n]\} - P\{M_n^* \le [u_n]\}) \\ &+(P\{M_n^* \le [u_n]\} - \exp(-\tau_n)) + (\exp(-\tau_n) - \exp(-e^{-x}))| \\ &\le |P\{M_n \le [u_n]\} - P\{M_n^* \le [u_n]\}| + |P\{M_n^* \le [u_n]\} - \exp(-\tau_n)| \\ &+|\exp(-\tau_n) - \exp(-e^{-x})|\end{aligned} \tag{2.21}$$

У раду [33] је доказано да у овом случају важе Лидбетерови услови $D(u_n)$ и $D'(u_n)$ (Лема 5.9, Лема 5.10).
Из једнакости (2.15) следи

$$F_n([u_n]) = P\{X_{nj} \le [u_n]\} = \sum_{i=2}^{[u_n]} \sum_{s=0}^{[\frac{i}{2}]-1} \binom{i-s-2}{s} \left(1-\frac{1}{n}\right)^{i-s-2} \frac{1}{n^{s+2}}. \tag{2.22}$$

Даље, нека је $\tau_n = n(1-F_n(u_n)) = n(1-F_n([u_n]))$. У раду [33] добијено је да важи

$$\tau_n = \frac{1}{n}\left(1-\frac{1}{n}\right)^{-\frac{1}{2}}\left(1+\frac{3}{n}\right)^{-\frac{1}{2}}\left(\frac{q_1^{[u_n]}}{1-q_1} - \frac{q_2^{[u_n]}}{1-q_2}\right),$$

$$q_1 = 1 - \frac{1}{n^2} + \frac{5}{32n^3} + o\left(\frac{1}{n^3}\right), \quad q_2 = -\frac{1}{n} + \frac{1}{n^2} - \frac{5}{32n^3} + o\left(\frac{1}{n^3}\right).$$

Из тога следи

$$\begin{aligned}\tau_n = \frac{1}{n}\left(1-\frac{1}{n}\right)^{-\frac{1}{2}}\left(1+\frac{3}{n}\right)^{-\frac{1}{2}}\left(1-\frac{1}{n^2}+\frac{5}{32n^3}+o\left(\frac{1}{n^3}\right)\right)^{[u_n]} \times \\ \times\left(\frac{1}{n^2}-\frac{5}{32n^3}+o\left(\frac{1}{n^3}\right)\right)^{-1}(1+o(1)) \\ = \frac{1}{n}\left(1+\frac{2}{n}-\frac{3}{n^2}\right)^{-\frac{1}{2}}\left(1-\frac{1}{n^2}+o\left(\frac{1}{n^2}\right)\right)^{[u_n]} n^2(1+o(1)) \\ = n\exp\left\{[u_n]\log\left(1-\frac{1}{n^2}+o\left(\frac{1}{n^2}\right)\right)\right\}(1+o(1))\end{aligned}$$

$$= n\exp\left\{(n^2(x+\log n)-r_n)\left(-\frac{1}{n^2}-\frac{1}{2n^4}+o\left(\frac{1}{n^4}\right)\right)\right\}(1+o(1))$$

$$= n\exp\left\{-(x+\log n)+\frac{2r_n-x-\log n}{2n^2}+o\left(\frac{\log n}{n^2}\right)\right\}(1+o(1))$$

$$= e^{-x}\left\{1-\frac{\log n}{2n^2}+o\left(\frac{\log n}{n^2}\right)\right\}(1+o(1)),\ \text{кад } n\to\infty. \tag{2.23}$$

Такође, кад $n \to \infty$ важе низови релација

$$P\{M_n^* \le [u_n]\} = (1-(1-F_n([u_n])))^n =$$
$$= \sum_{m=0}^{n}(-1)^m\binom{n}{m}\times$$
$$\times\exp\left\{-m\left((x+\log n)+\frac{2r_n-x-\log n}{2n^2}+o\left(\frac{\log n}{n^2}\right)\right)\right\}(1+o(1))$$
$$= \sum_{m=0}^{n}(-1)^m\binom{n}{m}\frac{e^{-mx}}{n^m}\left\{1-\frac{m\log n}{2n^2}+o\left(\frac{\log n}{n^2}\right)\right\}(1+o(1)), \tag{2.24}$$

као и

$$P\{M_n \le [u_n]\} = \sum_{m=0}^{n}(-1)^m\binom{n}{m}\frac{m}{n\sqrt{(n+1)^2-4m}}\times$$
$$\times\left(\left(\frac{t_1(m)}{n}\right)^{[u_n]}\frac{n}{n-t_1(m)}-\left(\frac{t_2(m)}{n}\right)^{[u_n]}\frac{n}{n-t_2(m)}\right)$$
$$= \sum_{m=0}^{n}(-1)^m\binom{n}{m}\frac{m}{n\sqrt{(n+1)^2-4m}}\left(\frac{t_1(m)}{n}\right)^{[u_n]}\frac{n}{n-t_1(m)}(1+o(1))$$
$$= \sum_{m=0}^{n}(-1)^m\binom{n}{m}\times$$
$$\times\exp\left\{(u_n-r_n)\log\left(1-\frac{m}{n^2}+\frac{m}{n^3}-\frac{m+m^2}{n^4}+o\left(\frac{1}{n^4}\right)\right)\right\}(1+o(1))$$
$$= \sum_{m=0}^{n}(-1)^m\binom{n}{m}\times$$
$$\times\exp\left\{(u_n-r_n)\left(-\frac{m}{n^2}+\frac{m}{n^3}-\frac{2m+3m^2}{2n^4}+o\left(\frac{1}{n^4}\right)\right)\right\}(1+o(1))$$

$$
\begin{aligned}
&= \sum_{m=0}^{n}(-1)^m \binom{n}{m}\frac{e^{-mx}}{n^m}\times \\
&\quad\times \exp\left\{(n^2(x+\log n)-r_n)\left(\frac{m}{n^3}-\frac{2m+3m^2}{2n^4}+o\left(\frac{1}{n^4}\right)\right)\right\}(1+o(1)) \\
&= \sum_{m=0}^{n}(-1)^m \binom{n}{m}\frac{e^{-mx}}{n^m}\times \\
&\qquad\times\left\{1+\frac{m(x+\log n)}{n}-\frac{(2m+3m^2)\log n}{2n^2}+o\left(\frac{\log n}{n^2}\right)\right\}(1+o(1)).
\end{aligned}
\tag{2.25}
$$

Из (2.24) и (2.25) следи

$$
|P\{M_n \le [u_n]\} - P\{M_n^* \le [u_n]\}| = O\left(\frac{\log n}{n}\right). \tag{2.26}
$$

Даље,

$$
\begin{aligned}
&|\exp(-\tau_n)-\exp(-e^{-x})\,| = \\
&= \left|\exp\left\{-e^{-x}\left\{1-\frac{\log n}{2n^2}+o\left(\frac{\log n}{n^2}\right)\right\}(1+o(1))\right\}-\exp(-e^{-x})\right| \\
&= \left|\exp(-e^{-x})\left\{1-e^{-x}\frac{\log n}{2n^2}+o\left(\frac{\log n}{n^2}\right)\right\}(1+o(1))-\exp(-e^{-x})\right| \\
&= \left|\exp(-e^{-x})\left\{-e^{-x}\frac{\log n}{2n^2}+o\left(\frac{\log n}{n^2}\right)\right\}(1+o(1))\right| \\
&\sim \exp(-e^{-x})e^{-x}\frac{\log n}{2n^2}, \qquad \text{кад } n\to\infty.
\end{aligned}
\tag{2.27}
$$

Из Теореме 2.11.1. у књизи [31] следи:

$$
P\{M_n^* \le [u_n]\} - \exp(-\tau_n) \sim \frac{\exp(-e^{-x})e^{-2x}}{2n}, \text{ кад } n\to\infty. \tag{2.28}
$$

Из (2.26), (2.27), (2.28) следи

$$
\left|P\{M_n \le [u_n]\} - \exp(-e^{-x})\right| = O\left(\frac{\log n}{n}\right).
$$

Овим је доказ комплетиран. ■

2.3 Проблеми чекања и генералисане Паретове расподеле

У овом поглављу приказани су нови резултати везани за граничну расподелу прекорачења високог нивоа случајне величине M_n из формулације проблема скупљања купона. Теореме (2.3.1) и (2.3.2) су објављене у раду [23].

2.3.1 Нови резултати

Теорема 2.3.1. [Jocković, Mladenović (2011)] *Нека је M_n време чекања до појаве свих елемената из $\mathbb{N}_n$ најмање r пута, где је r природан број. Нека је*

$$a_n = \log n + (r-1)\log\log n - \log(r-1)!, \quad n \geq 1.$$

(а) *За свако $x > 0$ и $c \in \mathbb{R}$ важи једнакост*

$$\lim_{n\to\infty} P\{M_n \leq n(x+c+a_n) \mid M_n > n(c+a_n)\} = \frac{\exp\left(-e^{-(x+c)}\right) - \exp\left(-e^{-c}\right)}{1-\exp\left(-e^{-c}\right)}. \tag{2.29}$$

(б) *Нека је $(c_n), n \in \mathbb{N}$, низ реалних бројева такав да $c_n \to +\infty$ кад $n \to \infty$. Тада, за свако $x > 0$ важи једнакост*

$$\lim_{n\to\infty} P\{M_n \leq n(x+c_n+a_n) \mid M_n > n(c_n+a_n)\} = 1-e^{-x}. \tag{2.30}$$

Теорема 2.3.2. [Jocković, Mladenović (2011)] *Нека је M_n време чекања до појаве свих парова jj, $j \in \mathbb{N}_n$.*
(а) *За свако $c \in \mathbb{R}$ и $x > 0$ важи једнакост*

$$\lim_{n\to\infty} P\{M_n \leq n^2(x+c+\log n) \mid M_n > n^2(c+\log n)\} = \frac{\exp\left(-e^{-(x+c)}\right) - \exp\left(-e^{-c}\right)}{1-\exp\left(-e^{-c}\right)}. \tag{2.31}$$

(б) *Нека је $(c_n), n \in \mathbb{N}$, низ реалних бројева такав да $c_n \to +\infty$ кад $n \to \infty$. Тада, за свако $x > 0$ важи једнакост*

$$\lim_{n\to\infty} P\{M_n \leq n^2(x+c_n+\log n) \mid M_n > n^2(c_n+\log n)\} = 1-e^{-x}. \tag{2.32}$$

Брзине конвергенције ка граничним расподелама одређене су у следеће две теореме.

Теорема 2.3.3. *Нека је M_n време чекања до појаве свих елемената из $\mathbb{N}_n$ најмање r пута, где је r природан број. Нека је*

$$a_n = \log n + (r-1)\log\log n - \log(r-1)!, \quad n \geq 1.$$

(а) *Нека је $(c_n), n \in \mathbb{N}$, низ реалних бројева такав да $c_n \to +\infty$ и $c_n = o(n)$ кад $n \to \infty$. Тада, за свако $x > 0$, при $n \to \infty$ важи једнакост*

$$P\{M_n \leq n(x+c_n+a_n) \mid M_n > n(c_n+a_n)\} - (1-e^{-x}) \sim (e^{-x}-1)\frac{e^{-x}}{2e^{c_n}}. \quad (2.33)$$

(б) *Нека је $(c_n), n \in \mathbb{N}$, низ реалних бројева такав да $c_n \to +\infty$ и $c_n = O(n)$ или $c_n/n \to \infty$ кад $n \to \infty$. Тада, за свако $x > 0$, при $n \to \infty$ важи једнакост*

$$P\{M_n \leq n(x+c_n+a_n) \mid M_n > n(c_n+a_n)\} - (1-e^{-x}) \sim (e^{-x}-1)\frac{e^{-x}}{2e^{c_n}e^{\frac{3c_n}{2n}}}. \quad (2.34)$$

Теорема 2.3.4. *Нека је M_n време чекања до појаве свих парова jj, $j \in \mathbb{N}_n$. Нека је $(c_n), n \in \mathbb{N}$, низ реалних бројева такав да $c_n \to +\infty$ кад $n \to \infty$. Тада, за свако $x > 0$, при $n \to \infty$ важи једнакост*

$$P\{M_n \leq n^2(x+c_n+\log n) \mid M_n > n^2(\log n+c_n)\} - (1-e^{-x}) \sim (e^{-x}-1)\frac{e^{\frac{c_n}{n}}e^{-x}}{2e^{c_n}}. \quad (2.35)$$

Познато је да за геометријску и негативну биномну расподелу не постоји гранична расподела максимума, као ни гранична расподела прекорачења високог нивоа. Ипак, у претходно разматраним случајевима добијене су граничне расподеле: Гумбелова расподела као гранична расподела максимума M_n, односно, експоненцијална расподела као гранична расподела прекорачења високог нивоа. Те чињенице нису у супротности, јер овде параметри тих расподела нису фиксирани, већ зависе од броја случајних величина.

При одређивању граничне расподеле прекорачења високог нивоа за неку функцију расподеле F циљ је одредити најнижи могући праг, u_n, при коме се добија та гранична расподела. Онда за све прагове $v_n > u_n$ може да се докаже иста гранична расподела.

Гранична расподела која се појављује у (2.29) и (2.31) не припада ниједној од три класе генералисаних Паретових расподела и зависи од прага. Ако је праг у Теореми 2.3.1 виши од $n(c + a_n)$ (прецизније, ако је једнак $n(c_n + a_n)$, где $c_n \to \infty$ кад $n \to \infty$), онда је гранична расподела прекорачења M_n преко овог прага експоненцијална расподела, која не зависи од прага. Исто твђење важи за праг $n^2(c+\log n)$ у Теореми 2.3.2. То значи да су прагови одређени у Теореми 2.3.1 (б) и Теореми 2.3.2 (б) најбољи (најнижи могући) у разматраним случајевима.

2.3.2 Докази

Доказ Теореме 2.3.1 (а) Из једнакости (2.7) следи да за свако $x > 0$ важи релација

$$\begin{aligned}
&P\{M_n \le n(x + c + a_n) \mid M_n > n(c + a_n)\} \\
&= \frac{P\{n(c + a_n) < M_n \le n(x + c + a_n)\}}{P\{M_n > n(c + a_n)\}} \\
&= \frac{P\{M_n \le n(x + c + a_n)\} - P\{M_n \le n(c + a_n)\}}{1 - P\{M_n \le n(c + a_n)\}} \\
&\to \frac{\exp\left(-e^{-(x+c)}\right) - \exp(-e^{-c})}{1 - \exp(-e^{-c})}, \quad \text{кад } n \to \infty.
\end{aligned}$$

(б) Нека је

$$a_{nk}(m) = \binom{n}{m} \sum_{s_m=0}^{r-1} \cdots \sum_{s_2=0}^{r-1} \sum_{s_1=0}^{r-1} p_{nm}(k, s_1, s_2, \ldots, s_m), \tag{2.36}$$

где је p_{nm} задато са (2.6). Циљ је одредити асимптотско понашање величина $a_{n,[u_n]}(1)$ и $a_{n,[u_n]}(2)$ кад $n \to \infty$, где је $u_n = n(x + c_n + a_n)$, $[u_n]$ је цео део броја u_n и $r_n = u_n - [u_n]$.

Из једнакости (2.6) и (2.36) за $m = 1$ изводи се низ релација

$$
\begin{aligned}
a_{n,[u_n]}(1) &= n \sum_{s_1=0}^{r-1} \binom{[u_n]}{s_1} \left(1 - \frac{1}{n}\right)^{u_n - r_n - s_1} \left(\frac{1}{n}\right)^{s_1} \\
&= n \frac{(n \log n)^{r-1}}{(r-1)!} \left(\frac{1}{n}\right)^{r-1} (1 + o(1)) \exp\left\{ (u_n - r_n - r + 1) \log\left(1 - \frac{1}{n}\right) \right\} \\
&= \frac{n(\log n)^{r-1}}{(r-1)!} \exp\left\{ \big(n(x + \log n + (r-1)\log\log n - \log(r-1)! + c_n) \right. \\
&\qquad \left. - r_n - r + 1\big) \left(-\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) \right) \right\} (1 + o(1)) \\
&= \frac{n(\log n)^{r-1}}{(r-1)!} \cdot \frac{e^{-x}(r-1)!}{n(\log n)^{r-1} e^{c_n}} \exp\left\{ -\frac{\log n + c_n}{2n} + o\left(\frac{\log n + c_n}{n}\right) \right\} (1 + o(1)) \\
&= \frac{e^{-x}}{e^{c_n}} \exp\left\{ -\frac{\log n + c_n}{2n} + o\left(\frac{\log n + c_n}{n}\right) \right\} (1 + o(1)). \qquad (2.37)
\end{aligned}
$$

Слично, за $m = 2$ добија се низ једнакости

$$
\begin{aligned}
a_{n,[u_n]}(2) &= \binom{n}{2} \sum_{s_1=0}^{r-1} \sum_{s_2=0}^{r-1} \binom{[u_n]}{s_1} \binom{[u_n] - s_1}{s_2} \left(1 - \frac{2}{n}\right)^{u_n - r_n - s_1 - s_2} \frac{1}{n^{s_1+s_2}} \\
&= \frac{n(n-1)}{2} \frac{(n \log n)^{2r-2}}{((r-1)!)^2} \left(\frac{1}{n}\right)^{2r-2} (1 + o(1)) \times \\
&\qquad \times \exp\left\{ (u_n - r_n - 2r + 2) \log\left(1 - \frac{2}{n}\right) \right\} \\
&= \frac{n(n-1)(\log n)^{2r-2}}{2((r-1)!)^2} \times \\
&\qquad \times \exp\left\{ \big(n(x + \log n + (r-1)\log\log n - \log(r-1)! + c_n) \right. \\
&\qquad \left. - r_n - 2r + 2\big) \left(-\frac{2}{n} - \frac{2}{n^2} + o\left(\frac{1}{n^2}\right) \right) \right\} (1 + o(1)) \\
&= \frac{n(n-1)(\log n)^{2r-2}}{2((r-1)!)^2} \frac{e^{-2x}((r-1)!)^2}{n^2(\log n)^{2r-2} e^{2c_n}} \times \\
&\qquad \times \exp\left\{ -\frac{2(\log n + c_n)}{n} + o\left(\frac{\log n + c_n}{n}\right) \right\} (1 + o(1)) \\
&= \frac{e^{-2x}}{2\, e^{2c_n}} \exp\left\{ -\frac{2(\log n + c_n)}{n} + o\left(\frac{\log n + c_n}{n}\right) \right\} (1 + o(1)). \qquad (2.38)
\end{aligned}
$$

Ако је $c_n = o(n)$ кад $n \to \infty$, онда се, коришћењем (2.37) и (2.38) добија следећи низ асимптотских једнакости за $a_{n,[u_n]}(1)$ и $a_{n,[u_n]}(2)$:

$$a_{n,[u_n]}(1) = \frac{e^{-x}}{e^{c_n}}(1+o(1)), \quad \text{кад } n \to \infty; \tag{2.39}$$

$$a_{n,[u_n]}(2) = \frac{e^{-2x}}{2\,e^{2c_n}}(1+o(1)), \quad \text{кад } n \to \infty. \tag{2.40}$$

У случајевима $c_n = O(n)$ или $c_n/n \to \infty$ кад $n \to \infty$, важи:

$$a_{n,[u_n]}(1) = \frac{e^{-x}}{e^{c_n} e^{c_n/(2n)}}(1+o(1)), \quad \text{кад } n \to \infty; \tag{2.41}$$

$$a_{n,[u_n]}(2) = \frac{e^{-2x}}{2\,e^{2c_n} e^{2c_n/n}}(1+o(1)), \quad \text{кад } n \to \infty. \tag{2.42}$$

Даље, да би се ограничила вероватноћа $P\{M_n > u_n\}$ користи се Бонферонијева неједнакост. Важи

$$\begin{aligned} P\{M_n > u_n\} &= P\left(\bigcup_{j=1}^{n}\{X_{nj} > u_n\}\right) \le \sum_{j=1}^{n} P\{X_{nj} > u_n\} = nP\{X_{n1} > u_n\} \\ &= n\sum_{s_1=0}^{r-1} p_{n1}([u_n], s_1) = a_{n,[u_n]}(1), \end{aligned} \tag{2.43}$$

и

$$\begin{aligned} P\{M_n > u_n\} &\ge \sum_{j=1}^{n} P\{X_{nj} > u_n\} - \sum_{i,j=1, i\neq j}^{n} P\{X_{ni} > u_n, X_{nj} > u_n\} \\ &= n\sum_{s_1=0}^{r-1} p_{n1}([u_n], s_1) - \binom{n}{2}\sum_{s_1=0}^{r-1}\sum_{s_2=0}^{r-1} p_{n2}([u_n], s_1, s_2) \\ &= a_{n,[u_n]}(1) - a_{n,[u_n]}(2). \end{aligned} \tag{2.44}$$

Ако је $c = o(n)$ кад $n \to \infty$, онда се, коришћењем (2.43), (2.44), (2.39) и (2.40), добија

$$P\{M_n > u_n\} = P\{M_n > n(x + c_n + a_n)\} = \frac{e^{-x}}{e^{c_n}}(1+o(1)), \quad \text{кад } n \to \infty. \tag{2.45}$$

Коначно, коришћењем (2.45), добијају се релације

$$\begin{aligned}
&P\{M_n \le n(x+a_n+c_n) \mid M_n > n(a_n+c_n)\}\\
&= \frac{P\{M_n > n(a_n+c_n)\} - P\{M_n > n(x+a_n+c_n)\}}{P\{M_n > n(a_n+c_n)\}}\\
&= \frac{\frac{1}{e^{c_n}}(1+o(1)) - \frac{e^{-x}}{e^{c_n}}(1+o(1))}{\frac{1}{e^{c_n}}(1+o(1))} \to 1-e^{-x} \quad \text{кад } n\to\infty.
\end{aligned}$$

Ако је $c_n = O(n)$ или $c_n/n \to \infty$ кад $n \to \infty$, онда се, коришћењем (2.41)-(2.44) добија

$$P\{M_n > n(x+c_n+a_n)\} = \frac{e^{-x}}{e^{c_n}e^{c_n/(2n)}}(1+o(1)), \quad \text{кад } n\to\infty, \tag{2.46}$$

и, слично као у претходном случају, једнакост (2.30) лако следи. ■

Доказ Теореме 2.3.2 (а) Доказ следи из једнакости (2.19), на исти начин као у доказу Теореме 2.3.1 (а).

(б) Из Бонферонијеве неједнакости следи да

$$1 - nP\{Y_{n1} > k\} \le P\{M_n \le k\} \le 1 - nP\{Y_{n1} > k\} + \binom{n}{2}P\{Y_{n2} > k\}, \tag{2.47}$$

где су расподеле случајних величина M_n и Y_{nm} дате у (2.11)-(2.14). Одатле,

$$nP\{Y_{n1} > k\} - \binom{n}{2}P\{Y_{n2} > k\} \le P\{M_n > k\} \le nP\{Y_{n1} > k\}. \tag{2.48}$$

За свако фиксирано m важе низови једнакости:

$$\begin{aligned}
t_1(m) &= \frac{n-1}{2} + \frac{n+1}{2}\sqrt{1 - \frac{4m}{(n+1)^2}}\\
&= \frac{n-1}{2} + \frac{n+1}{2}\left(1 - \frac{2m}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right)\right)\\
&= n - \frac{m}{n+1} + o\left(\frac{1}{n}\right), \quad \text{кад } n\to\infty,
\end{aligned} \tag{2.49}$$

и

$$
\begin{aligned}
t_2(m) &= \frac{n-1}{2} - \frac{n+1}{2}\sqrt{1-\frac{4m}{(n+1)^2}} \\
&= \frac{n-1}{2} - \frac{n+1}{2}\left(1-\frac{2m}{(n+1)^2}+o\left(\frac{1}{(n+1)^2}\right)\right) \\
&= -1+\frac{m}{n+1}+o\left(\frac{1}{n}\right), \quad \text{кад } n\to\infty. \qquad (2.50)
\end{aligned}
$$

Из (2.49) и (2.50) следи да је члан $\left(\frac{t_2(m)}{n}\right)^k \frac{n}{n-t_2(m)}$ много мањи од $\left(\frac{t_1(m)}{n}\right)^k \frac{n}{n-t_1(m)}$ и може бити занемарен у асимптотском понашању вероватноће $P\{Y_{nm}>k\}$ кад $n\to\infty$.

Даље, нека је $u_n = n^2(x+c_n+\log n)$ и $r_n = u_n - [u_n]$. Треба одредити понашање величине $\binom{n}{m}P\{Y_{nm}>u_n\}$ за свако фиксирано m кад $n\to\infty$.

Тачне су следеће једнакости и асимптотске релације:

$$
\begin{aligned}
&\frac{m}{n\sqrt{(n+1)^2-4m}}\cdot\frac{n}{n-t_1(m)} \\
&= \frac{m}{n\sqrt{(n+1)^2-4m}}\cdot\frac{2n}{n+1-\sqrt{(n+1)^2-4m}} \\
&= \frac{m}{n(n+1)\sqrt{1-\frac{4m}{(n+1)^2}}}\cdot\frac{2n}{(n+1)\left(1-\sqrt{1-\frac{4m}{(n+1)^2}}\right)} \\
&= \frac{2m}{(n+1)^2}\left\{1-\frac{4m}{(n+1)^2}\right\}^{-1/2}\left\{1-\left(1-\frac{4m}{(n+1)^2}\right)^{1/2}\right\}^{-1} \\
&= \frac{2m}{(n+1)^2}\left\{1+\frac{2m}{(n+1)^2}+o\left(\frac{1}{(n+1)^2}\right)\right\}\times \\
&\qquad\qquad \times\left\{1-\left(1-\frac{2m}{(n+1)^2}+o\left(\frac{1}{(n+1)^2}\right)\right)\right\}^{-1} \\
&= \frac{2m}{(n+1)^2}\left\{1+\frac{2m}{(n+1)^2}+o\left(\frac{1}{(n+1)^2}\right)\right\}\times \\
&\qquad\qquad \times\left\{\frac{2m}{(n+1)^2}+o\left(\frac{1}{(n+1)^2}\right)\right\}^{-1} \\
&= 1+o(1), \quad \text{кад } n\to\infty; \qquad (2.51)
\end{aligned}
$$

и

$$\left\{\frac{t_1(m)}{n}\right\}^{[u_n]} = \left\{\frac{n-1+\sqrt{(n+1)^2-4m}}{2n}\right\}^{[u_n]}$$
$$= \left\{\frac{1}{2n}\left(n-1+(n+1)\left(1-\frac{4m}{(n+1)^2}\right)^{1/2}\right)\right\}^{[u_n]}$$
$$= \left\{\frac{1}{2n}\left(n-1+(n+1)\left(1-\frac{2m}{(n+1)^2}-\frac{2m^2}{(n+1)^4}+o\left(\frac{1}{(n+1)^4}\right)\right)\right)\right\}^{[u_n]}$$

$$= \left\{\frac{1}{2n}\left(2n-\frac{2m}{n+1}-\frac{2m^2}{(n+1)^3}+o\left(\frac{1}{(n+1)^3}\right)\right)\right\}^{[u_n]}$$
$$= \left\{1-\frac{m}{n(n+1)}-\frac{m^2}{n(n+1)^3}+o\left(\frac{1}{n^4}\right)\right\}^{[u_n]}$$
$$= \left\{1-\frac{m}{n^2}\left(1+\frac{1}{n}\right)^{-1}-\frac{m^2}{n^4}\left(1+\frac{1}{n}\right)^{-3}+o\left(\frac{1}{n^4}\right)\right\}^{[u_n]}$$
$$= \left\{1-\frac{m}{n^2}\left(1-\frac{1}{n}+\frac{1}{n^2}-\frac{1}{n^3}+o\left(\frac{1}{n^3}\right)\right)-\frac{m^2}{n^4}\left(1+o(1)\right)+o\left(\frac{1}{n^4}\right)\right\}^{[u_n]}$$
$$= \left\{1-\frac{m}{n^2}+\frac{m}{n^3}-\frac{m+m^2}{n^4}+o\left(\frac{1}{n^4}\right)\right\}^{n^2(x+c_n+\log n)-r_n}$$
$$= \exp\left\{\left(n^2(x+c_n+\log n)-r_n\right)\log\left(1-\frac{m}{n^2}+\frac{m}{n^3}-\frac{m+m^2}{n^4}+o\left(\frac{1}{n^4}\right)\right)\right\}$$
$$= \exp\left\{\left(n^2(x+c_n+\log n)-r_n\right)\left(-\frac{m}{n^2}+\frac{m}{n^3}-\frac{m+m^2}{n^4}-\frac{m^2}{2n^4}+o\left(\frac{1}{n^4}\right)\right)\right\}$$
$$= \exp\left\{-m(x+c_n+\log n)+\frac{m}{n}(x+c_n+\log n)+o\left(\frac{\log n}{n}\right)\right\}$$
$$= \frac{e^{-mx}e^{mc_n/n}}{n^m e^{mc_n}}(1+o(1)), \quad \text{кад } n\to\infty. \tag{2.52}$$

Коришћењем (2.12) и (2.51)-(2.52) добија се

$$
\begin{aligned}
\binom{n}{m} P\{Y_{nm} > u_n\} &= \binom{n}{m} P\{Y_{nm} > [u_n]\} \\
&= \binom{n}{m} \frac{m}{n\sqrt{(n+1)^2 - 4m}} \left(\frac{t_1(m)}{n}\right)^{[u_n]} \frac{n}{n - t_1(m)} \\
&= \binom{n}{m} \frac{e^{-mx} e^{mc_n/n}}{n^m e^{mc_n}} (1 + o(1)), \quad \text{кад } n \to \infty. \qquad (2.53)
\end{aligned}
$$

За $m = 1$ и $m = 2$, из (2.53) следи да

$$
nP\{Y_{n1} > u_n\} = \frac{e^{-x} e^{c_n/n}}{e^{c_n}} (1 + o(1)), \qquad (2.54)
$$

$$
\binom{n}{2} P\{Y_{n2} > u_n\} = \frac{e^{-2x} e^{2c_n/n}}{2e^{2c_n}} (1 + o(1)). \qquad (2.55)
$$

Коришћењем (2.48) и (2.54)-(2.55) добија се

$$
\begin{aligned}
P\{M_n > u_n\} &= P\{M_n > n^2(x + c_n + \log n)\} \\
&= \frac{e^{-x} e^{c_n/n}}{e^{c_n}} (1 + o(1)), \quad \text{кад } n \to \infty. \qquad (2.56)
\end{aligned}
$$

Коначно, коришћењем (2.56) изводе се следеће релације

$$
\begin{aligned}
& P\{M_n \le n^2(x + c_n + \log n) \mid M_n > n^2(c_n + \log n)\} \\
&= \frac{P\{M_n > n^2(c_n + \log n)\} - P\{M_n > n^2(x + c_n + \log n)\}}{P\{M_n > n^2(c_n + \log n)\}} \\
&= \frac{\frac{e^{c_n/n}}{e^{c_n}}(1 + o(1)) - \frac{e^{-x} e^{c_n/n}}{e^{c_n}}(1 + o(1))}{\frac{e^{c_n/n}}{e^{c_n}}(1 + o(1))} \to 1 - e^{-x}, \quad \text{кад } n \to \infty.
\end{aligned}
$$

Овим је теорема доказана. ■

Доказ Теореме 2.3.3 (а) Ако је $c_n = o(n)$ кад $n \to \infty$, онда из (2.6) и (2.36), на исти начин као (2.39) и (2.40), за $m = 3$ следи једнакост

$$
a_{n,[u_n]}(3) = \frac{e^{-3x}}{6e^{3c_n}} (1 + o(1)), \qquad \text{кад } n \to \infty. \qquad (2.57)
$$

Из Бонферонијеве неједнакости, слично као (2.43) и (2.44), следи да

$$a_{n,[u_n]}(1) - a_{n,[u_n]}(2) + a_{n,[u_n]}(3) \geq P\{M_n > [u_n]\} \geq a_{n,[u_n]}(1) - a_{n,[u_n]}(2). \tag{2.58}$$

Заменом (2.39), (2.40) и (2.57) у (2.58), добија се

$$\begin{aligned}\left(\frac{e^{-x}}{e^{c_n}} - \frac{e^{-2x}}{2e^{2c_n}} + \frac{e^{-3x}}{6e^{3c_n}}\right)(1+o(1)) &\geq P\{M_n > [u_n]\} \\ &\geq \left(\frac{e^{-x}}{e^{c_n}} - \frac{e^{-2x}}{2e^{2c_n}}\right)(1+o(1)),\end{aligned}$$

одакле директно следи

$$P\{M_n > [u_n]\} = \left(\frac{e^{-x}}{e^{c_n}} - \frac{e^{-2x}}{2e^{2c_n}} + o\left(\frac{1}{e^{2c_n}}\right)\right)(1+o(1)), \quad \text{кад } n \to \infty. \tag{2.59}$$

Заменом (2.59) у леву страну једнакости (2.33), добија се низ релација

$$\begin{aligned}
&P\{M_n \leq n(x+c_n+a_n) \mid M_n > n(c_n+a_n)\} - (1-e^{-x}) \\
&= 1 - \frac{P\{M_n > n(x+c_n+a_n)\}}{P\{M_n > n(c_n+a_n)\}} - (1-e^{-x}) \\
&= -\frac{P\{M_n > n(x+c_n+a_n)\}}{P\{M_n > n(c_n+a_n)\}} + e^{-x} \\
&= e^{-x} - \left(\frac{e^{-x}}{e^{c_n}} - \frac{e^{-2x}}{2e^{2c_n}} + o\left(\frac{1}{e^{2c_n}}\right)\right)(1+o(1)) \times \\
&\qquad\qquad \times \left(\left(\frac{1}{e^{c_n}} - \frac{1}{2e^{2c_n}} + o\left(\frac{1}{e^{2c_n}}\right)\right)(1+o(1))\right)^{-1} \\
&= e^{-x} - e^{-x}\left(\frac{1}{e^{c_n}} - \frac{e^{-x}}{2e^{2c_n}} + o\left(\frac{1}{e^{2c_n}}\right)\right)e^{c_n} \times \\
&\qquad\qquad \times \left(1 - \frac{1}{2e^{c_n}} + o\left(\frac{1}{e^{c_n}}\right)\right)^{-1}(1+o(1)) \\
&= e^{-x} - e^{-x}\frac{1}{e^{c_n}}\left(1 - \frac{e^{-x}}{2e^{c_n}} + o\left(\frac{1}{e^{c_n}}\right)\right)e^{c_n}\left(1 + \frac{1}{2e^{c_n}} + o\left(\frac{1}{e^{c_n}}\right)\right)(1+o(1)) \\
&= e^{-x}\left(1 - \left(1 - \frac{e^{-x}}{2e^{c_n}} + \frac{1}{2e^{c_n}} + o\left(\frac{1}{e^{c_n}}\right)\right)\right)(1+o(1)) \\
&\sim e^{-x}(e^{-x}-1)\frac{1}{2e^{c_n}}, \quad \text{кад } n \to \infty,
\end{aligned} \tag{2.60}$$

чиме је доказ теореме комплетиран. ■

(б) Ако је $c_n = O(n)$ или $c_n/n \to \infty$ кад $n \to \infty$, онда из (2.6) и (2.36), на исти начин као (2.41) и (2.42), за $m = 3$ следи једнакост

$$a_{n,[u_n]}(3) = \frac{e^{-3x} e^{\frac{-9c_n}{2n}}}{6e^{3c_n}}(1 + o(1)), \quad \text{кад } n \to \infty. \tag{2.61}$$

Заменом (2.41), (2.42) и (2.61) у (2.58), добија се

$$\begin{aligned}\left(\frac{e^{-x}}{e^{c_n} e^{\frac{c_n}{2n}}} - \frac{e^{-2x}}{2e^{2c_n} e^{\frac{2c_n}{n}}} + \frac{e^{-3x}}{6e^{3c_n} e^{\frac{9c_n}{2n}}}\right)(1 + o(1)) &\geq P\{M_n > [u_n]\} \\ &\geq \left(\frac{e^{-x}}{e^{c_n} e^{\frac{c_n}{2n}}} - \frac{e^{-2x}}{2e^{2c_n} e^{\frac{2c_n}{n}}}\right)(1 + o(1)),\end{aligned} \tag{2.62}$$

одакле директно следи

$$P\{M_n > [u_n]\} = \left(\frac{e^{-x}}{e^{c_n} e^{\frac{c_n}{2n}}} - \frac{e^{-2x}}{2e^{2c_n} e^{\frac{2c_n}{n}}} + o\left(\frac{1}{e^{2c_n} e^{\frac{2c_n}{n}}}\right)\right)(1+o(1)), \quad n \to \infty. \tag{2.63}$$

Заменом (2.62) у леву страну једнакости (2.34), добија се низ релација

$$\begin{aligned}
&P\{M_n \leq n(x + c_n + a_n) \mid M_n > n(c_n + a_n)\} - (1 - e^{-x}) \\
&= 1 - \frac{P\{M_n > n(x + c_n + a_n)\}}{P\{M_n > n(c_n + a_n)\}} - (1 - e^{-x}) \\
&= -\frac{P\{M_n > n(x + c_n + a_n)\}}{P\{M_n > n(c_n + a_n)\}} + e^{-x} \\
&= e^{-x} - \left(\frac{e^{-x}}{e^{c_n} e^{\frac{c_n}{2n}}} - \frac{e^{-2x}}{2e^{2c_n} e^{\frac{2c_n}{n}}} + o\left(\frac{1}{e^{2c_n} e^{\frac{2c_n}{n}}}\right)\right)(1 + o(1)) \times \\
&\qquad \times \left(\left(\frac{1}{e^{c_n} e^{\frac{c_n}{2n}}} - \frac{1}{2e^{2c_n} e^{\frac{2c_n}{n}}} + o\left(\frac{1}{e^{2c_n} e^{\frac{2c_n}{n}}}\right)\right)(1 + o(1))\right)^{-1} \\
&= e^{-x} - e^{-x}\left(\frac{1}{e^{c_n} e^{\frac{c_n}{2n}}} - \frac{e^{-x}}{2e^{2c_n} e^{\frac{2c_n}{n}}} + o\left(\frac{1}{e^{2c_n} e^{\frac{2c_n}{n}}}\right)\right) e^{c_n} e^{\frac{c_n}{2n}} \times \\
&\qquad \times \left(1 - \frac{1}{2e^{c_n} e^{\frac{3c_n}{2n}}} + o\left(\frac{1}{e^{c_n} e^{\frac{3c_n}{2n}}}\right)\right)^{-1} (1 + o(1))
\end{aligned}$$

$$
\begin{aligned}
&= e^{-x} - e^{-x}\frac{1}{e^{c_n}e^{\frac{c_n}{2n}}}\left(1 - \frac{e^{-x}}{2e^{c_n}e^{\frac{3c_n}{2n}}} + o\left(\frac{1}{e^{c_n}e^{\frac{3c_n}{2n}}}\right)\right)e^{c_n}e^{\frac{c_n}{2n}} \times \\
&\qquad\qquad \times\left(1 + \frac{1}{2e^{c_n}e^{\frac{3c_n}{2n}}} + o\left(\frac{1}{e^{c_n}e^{\frac{3c_n}{2n}}}\right)\right)(1+o(1)) \\
&= e^{-x}\left(1 - \left(1 - \frac{e^{-x}}{2e^{c_n}e^{\frac{3c_n}{2n}}} + \frac{1}{2e^{c_n}e^{\frac{3c_n}{2n}}} + o\left(\frac{1}{e^{c_n}e^{\frac{3c_n}{2n}}}\right)\right)\right)(1+o(1)) \\
&\sim e^{-x}(e^{-x}-1)\frac{1}{2e^{c_n}e^{\frac{3c_n}{2n}}}, \quad \text{кад } n \to \infty, \qquad (2.64)
\end{aligned}
$$

одакле директно следи тврђење теореме. ■

Доказ Теореме 2.3.4 За $m = 3$ из (2.53) се добија

$$
\binom{n}{3}P\{Y_{n3} > [u_n]\} = \frac{e^{-3x}e^{\frac{3c_n}{n}}}{6e^{3c_n}}(1+o(1)), \quad \text{кад } n \to \infty. \qquad (2.65)
$$

Из Бонферонијеве неједнакости, на исти начин као (2.48), следи

$$
\begin{aligned}
nP\{Y_{n1} > [u_n]\} - \binom{n}{2}P\{Y_{n2} > [u_n]\} + \binom{n}{3}P\{Y_{n3} > [u_n]\} &\geq P\{M_n > [u_n]\} \\
\geq nP\{Y_{n1} > [u_n]\} - \binom{n}{2}P\{Y_{n2} > [u_n]\}. \qquad (2.66)
\end{aligned}
$$

Заменом (2.54), (2.55) и (2.65) у (2.66) добија се

$$
\begin{aligned}
\left(\frac{e^{-x}e^{\frac{c_n}{n}}}{e^{c_n}} - \frac{e^{-2x}e^{\frac{2c_n}{n}}}{2e^{2c_n}} + \frac{e^{-3x}e^{\frac{3c_n}{n}}}{6e^{3c_n}}\right)(1+o(1)) &\geq P\{M_n > [u_n]\} \\
\geq \left(\frac{e^{-x}e^{\frac{c_n}{n}}}{e^{c_n}} - \frac{e^{-2x}e^{\frac{2c_n}{n}}}{2e^{2c_n}}\right)(1+o(1)), \qquad (2.67)
\end{aligned}
$$

одакле директно следи

$$
P\{M_n > [u_n]\} = \left(\frac{e^{-x}e^{\frac{c_n}{n}}}{e^{c_n}} - \frac{e^{-2x}e^{\frac{2c_n}{n}}}{2e^{2c_n}} + o\left(\frac{e^{\frac{2c_n}{n}}}{e^{2c_n}}\right)\right)(1+o(1)), \text{ кад } n \to \infty. \qquad (2.68)
$$

Заменом (2.68) у леву страну једнакости (2.35) добија се низ релација

$$
\begin{aligned}
&P\{M_n \le n^2(x+c_n+\log n) \mid M_n > n^2(c_n+\log n)\} - (1-e^{-x}) \\
&= 1 - \frac{P\{M_n > n^2(x+c_n+\log n)\}}{P\{M_n > n^2(c_n+\log n)\}} - (1-e^{-x}) \\
&= -\frac{P\{M_n > n^2(x+c_n+\log n)\}}{P\{M_n > n^2(c_n+\log n)\}} + e^{-x} \\
&= e^{-x} - \left(\frac{e^{-x}e^{\frac{c_n}{n}}}{e^{c_n}} - \frac{e^{-2x}e^{\frac{2c_n}{n}}}{2e^{2c_n}} + o\left(\frac{e^{\frac{2c_n}{n}}}{e^{2c_n}}\right)\right)(1+o(1)) \times \\
&\qquad\qquad \times \left(\left(\frac{e^{\frac{c_n}{n}}}{e^{c_n}} - \frac{e^{\frac{2c_n}{n}}}{2e^{2c_n}} + o\left(\frac{e^{\frac{2c_n}{n}}}{e^{2c_n}}\right)\right)(1+o(1))\right)^{-1} \\
&= e^{-x} - e^{-x}\left(\frac{e^{\frac{c_n}{n}}}{e^{c_n}} - \frac{e^{-x}e^{\frac{2c_n}{n}}}{2e^{2c_n}} + o\left(\frac{e^{\frac{2c_n}{n}}}{e^{2c_n}}\right)\right)\frac{e^{c_n}}{e^{\frac{c_n}{n}}} \times \\
&\qquad\qquad \times \left(1 - \frac{e^{\frac{c_n}{n}}}{2e^{c_n}} + o\left(\frac{e^{\frac{c_n}{n}}}{e^{c_n}}\right)\right)^{-1}(1+o(1)) \\
&= e^{-x} - e^{-x}\frac{e^{\frac{c_n}{n}}}{e^{c_n}}\left(1 - \frac{e^{-x}e^{\frac{c_n}{n}}}{2e^{c_n}} + o\left(\frac{e^{\frac{c_n}{n}}}{e^{c_n}}\right)\right)\frac{e^{c_n}}{e^{\frac{c_n}{n}}} \times \\
&\qquad\qquad \times \left(1 + \frac{e^{\frac{c_n}{n}}}{2e^{c_n}} + o\left(\frac{e^{\frac{c_n}{n}}}{e^{c_n}}\right)\right)(1+o(1)) \\
&= e^{-x}\left(1 - \left(1 - \frac{e^{-x}e^{\frac{c_n}{n}}}{2e^{c_n}} + o\left(\frac{e^{\frac{c_n}{n}}}{e^{c_n}}\right)\right)\left(1 + \frac{e^{\frac{c_n}{n}}}{2e^{c_n}} + o\left(\frac{e^{\frac{c_n}{n}}}{e^{c_n}}\right)\right)\right) \times \\
&\qquad\qquad \times (1+o(1)) \\
&= e^{-x}\left(1 - \left(1 - \frac{e^{-x}e^{\frac{c_n}{n}}}{2e^{c_n}} + \frac{e^{\frac{c_n}{n}}}{2e^{c_n}} + o\left(\frac{e^{\frac{c_n}{n}}}{e^{c_n}}\right)\right)\right)(1+o(1)) \\
&\sim e^{-x}(e^{-x}-1)\frac{e^{\frac{c_n}{n}}}{2e^{c_n}}, \qquad \text{кад } n \to \infty,
\end{aligned}
\tag{2.69}
$$

одакле директно следи тврђење теореме. ■

Глава 3

Оцењивање параметара генералисаних Паретових расподела

Нека је $(x_1, x_2, \ldots, x_N)$ узорак чији чланови представљају реализације неке случајне величине X са непознатом функцијом расподеле F. Да би се функција расподеле прекорачења прага u, $F^{(u)}$, апроксимирала неком од генералисаних Паретових расподела, потребно је одредити високи ниво u, а онда, на основу узорка који чине прекорачења $(x_1 - u, x_2 - u, \ldots, x_n - u)$, где су $x_1, x_2, \ldots, x_n$ чланови почетног узорка већи од u, оценити параметре облика и размере одговарајуће двопараметарске генералисане Паретове расподеле.

Двопараметарска фамилија генералисаних Паретових расподела (у даљем тексту $\mathrm{GP}(\gamma, \sigma)$) је задата на следећи начин:

$$W_{\gamma,\sigma}(x) = \begin{cases} 1 - e^{-\frac{x}{\sigma}}, & x \geq 0, \gamma = 0; \\ 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-\frac{1}{\gamma}}, & x \geq 0, \gamma > 0; \\ 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-\frac{1}{\gamma}}, & x \in \left[0, -\frac{\sigma}{\gamma}\right], \gamma < 0. \end{cases} \tag{3.1}$$

Оцењивања параметара и квантила генералисаних Паретових расподела и даље је актуелна тема истраживања. Неки од новијих радова из те области су Castillo and Daoudi (2009) [8] и Mackay et al. (2011) [29]. Детаљан преглед постојећих метода оцењивања дат је у раду Bermudez and Kotz (2010) [7].

У поглављима 3.1 и 3.2 описане су неке од постојећих метода оцењивања параметара GP(γ, σ) расподеле. У поглављу 3.3 приказана је нова метода за поправку несагласних оцена. Поглавље 3.4 бави се упоређивањем особина различитих оцена квантила $W_{\gamma,\sigma}$ расподеле и могућношћу њихове примене у финансијама.

Сва израчунавања и симулације расподела урађени су уз помоћ програмског пакета MATLAB®, release 2007a.

Одабир високог нивоа

Посебан проблем представља одабир високог нивоа, изнад кога је оправдана апроксимација генералисаним Паретовим расподелама. Постоји више начина да се тај ниво одреди. Једна од познатих метода користи **узорачко средње прекорачење**.

Узорачко средње прекорачење је:

$$e_N(u) = \frac{\sum\limits_{i \le N} (x_i - u) I(u < x_i)}{\sum\limits_{i \le N} I(u < x_i)}$$

и то је једна емпиријска оцена средњег прекорачења за функцију расподеле F.

Познато је да је за све генералисане Паретове расподеле средње прекорачење линеарна функција, па се зато бира се она вредност прага u, за коју важи да је узорачко средње прекорачење приближно линеарна функција, почевши од те вредности u.

Ако се изабере претерано низак праг, онда нису испуњени сви услови Теореме 1.2.3, тј. апроксимација овим типом расподела није одговајућа и добиће се пристрасне оцене непознатих параметара. С друге стране, ако је одабрани праг превише висок, изнад прага остаје превише мало чланова узорка и добијене оцене ће имати велику дисперзију. Уопште, повећањем прага повећава се дисперзија, а смањује пристрасност оцене, и обрнуто. Ова метода, као и неке друге, разматрана је у књизи Reiss (1997) [36].

3.1 Класичне методе оцењивања

3.1.1 Метод момената

Под одређеним условима, непрекидне расподеле су једнозначно одређене својим моментима. Из тога следи да је познавање момената до одређеног реда довољно за одређивање непознатих параметара расподеле.

Централни моменти случајне величине X са $\mathrm{GP}(\gamma, \sigma)$ расподелом дати су са

$$\mathrm{E}(X^r) = r! \frac{\sigma^r}{\gamma^{r+1}} \frac{\Gamma\left(\frac{1}{\gamma} - r\right)}{\Gamma\left(1 + \frac{1}{\gamma}\right)}, \quad \gamma < \frac{1}{r}, \tag{3.2}$$

одакле следи да су средња вредност и дисперзија ове случајне величине:

$$\mathrm{E}(X) = m_{W_{\gamma,\sigma}} = \frac{\sigma}{1-\gamma}, \quad \gamma < 1, \tag{3.3}$$

$$\mathrm{var}_{W_{\gamma,\sigma}} = \frac{\sigma^2}{(1-\gamma)^2(1-2\gamma)}, \quad \gamma < \frac{1}{2}. \tag{3.4}$$

Нека је $(x_1, x_2, \ldots, x_n)$ узорак из $\mathrm{GP}(\gamma, \sigma)$ расподеле, $\overline{x}$ узорачка средина и s^2 узорачка дисперзија. Оцене параметара (γ, σ) **методом момената** (ММ) добијају се заменом средње вредности и дисперзије у (3.3) и (3.4) њиховим узорачким еквивалентима и решавањем одговарајућег система једначина. Оне су дате са:

$$\hat{\gamma} = \frac{1}{2}\left(1 - \frac{\overline{x}^2}{s^2}\right) \quad \text{и} \quad \hat{\sigma} = \frac{1}{2}\overline{x}\left(\frac{\overline{x}^2}{s^2} + 1\right). \tag{3.5}$$

Ове оцене су дефинисане у раду Hosking and Wallis (1987) [20].

3.1.2 Метод пондерисаних момената

Пондерисани моменти $M_{p,r,s}$ случајне величине X са функцијом расподеле F дати су изразима

$$M_{p,r,s} = E[X^p (F(X))^r (1 - F(X))^s], \quad p, r, s \in \mathbb{R}. \tag{3.6}$$

У случају GP(γ, σ) расподеле, могу се користити једноставније релације

$$\alpha_s = M_{1,0,s} = E[X(1-F(X))^s] = \frac{\sigma}{(s+1)(s+1-\gamma)}, \quad \gamma < 1, s = 0, 1, 2, \ldots \tag{3.7}$$

из којих се добија

$$\gamma = 2 - \frac{\alpha_0}{\alpha_0 - 2\alpha_1} \quad \text{и} \quad \sigma = \frac{2\alpha_0\alpha_1}{\alpha_0 - 2\alpha_1}. \tag{3.8}$$

Нека је $x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}$ варијациони низ статистика поретка за узорак $(x_1, x_2, \ldots, x_n)$ из GP(γ, σ) расподеле и нека је

$$a_k = \frac{1}{n}\sum_{j=1}^{n}(1 - p_{j:n})^k x_{j:n} \quad \text{где је} \quad p_{j:n} = \frac{j - 0.35}{n}. \tag{3.9}$$

Оцене параметара γ и σ **методом пондерисаних момената** (МПМ) добијају се заменом α_0 и α_1 из (3.8) њиховим узорачким еквивалентима (3.9) и дате су са

$$\hat{\gamma} = 2 - \frac{a_0}{a_0 - 2a_1} \quad \text{и} \quad \hat{\sigma} = \frac{2a_0a_1}{a_0 - 2a_1}. \tag{3.10}$$

Овај метод оцењивања уведен је у раду Greenwood et al. (1979) [17].

3.1.3 Метод максималне веродостојности

Нека је H_θ расподела чије параметре треба оценити, h_θ одговарајућа густина расподеле и $(x_1, x_2, \ldots, x_n)$ узорак на основу кога се врши оцењивање. **Метода максималне веродостојности** (ММВ) састоји се у томе да се за оцену параметра θ узима она вредност за коју је вероватноћа реализације узорка $(x_1, x_2, \ldots, x_n)$ највећа. Функција веродостојности базирана на датом узорку је

$$L(\theta; x) = \prod_{i=1}^{n} h_\theta(x_i) I_{\theta \in D(\theta)},$$

где је $D(\theta)$ скуп дозвољених вредности параметра θ. Треба одредити θ које максимизира функцију веродостојности, односно, због једноставности, њен логаритам.

У случају $\mathrm{GP}(\gamma, \sigma)$ расподеле, логаритам функције веродостојности је

$$\log L(\gamma,\sigma;x) = \begin{cases} -n\log n - \left(\frac{1}{\gamma}+1\right)\sum\limits_{i=1}^{n}\log\left(1+\frac{\gamma}{\sigma}x_i\right), & \gamma \neq 0, \\ -n\log n - \frac{1}{\sigma}\sum\limits_{i=1}^{n} x_i, & \gamma = 0. \end{cases}$$

Оцене параметра γ и σ се онда добијају као решења система једначина:

$$\frac{\partial(\log L(\gamma,\sigma;x))}{\partial\gamma} = 0,$$

$$\frac{\partial(\log L(\gamma,\sigma;x))}{\partial\sigma} = 0.$$

У општем случају, решења овог система се не добијају у експлицитном облику, већ применом нумеричких метода.

3.2 Остале методе

3.2.1 Неке модификације класичних метода

Последњих година појавило се више модификација класичних метода за оцењивање параметара $\mathrm{GP}(\gamma, \sigma)$ расподеле. Неки од њих су:

1. *уопштени метод момената:* дефинисан у раду Ashkar and Ouarda (1996), користи трансформисани облик $\mathrm{GP}(\gamma, \sigma)$ расподеле, чиме се добија специјалан случај Пирсонове расподеле типа 3.

2. *метод L-момената:* дефинисан у раду Hosking (1990), уместо обичних момената случајне величине са $\mathrm{GP}(\gamma, \sigma)$ расподелом користи линеарну комбинацију очекивања статистика поретка, тј. линеарну комбинацију пондерисаних момената. L-моменат реда r случајне величине X дат је са

$$\lambda_{\mathrm{r}} = \frac{1}{r}\sum_{i=0}^{r-1}(-1)^i\binom{r-1}{i}E(X_{r-i:r}), \quad r \in \mathbb{N}.$$

Оцене непознатих параметара добијају се заменом одређеног броја L-момената њиховим узорачким еквивалентима и решавањем добијеног система једначина.

3. *метод L-момената вишег реда:* дефинисан у раду Wang (1997), користи уопштену верзију L-момената за горњи део расподеле. L-моменат вишег реда r случајне величине X дат је са
$$\lambda_{\mathrm{r}}^{\eta} = \frac{1}{r}\sum_{i=0}^{r-1}(-1)^i\binom{r-1}{i}E(X_{\eta+r-i:\eta+r}), \quad r, \eta \in \mathbb{N}.$$
Оцене непознатих параметара добијају се исто као у претходном случају.

4. *уопштени метод пондерисаних момената:* дефинисан у раду Rasmussen (2001), заснива се на замени било које две реалне вредности s_1, s_2 у једнакости (3.7) и решавању добијеног система једначина. Добијене оцене су:
$$\hat{\sigma} = a_{s_2}(s_2+1)(s_2+1+\gamma), \quad \hat{\gamma} = \frac{a_{s_2}(s_2+1)^2 - a_{s_1}(s_1+1)^2}{a_{s_2}(s_2+1) - a_{s_1}(s_1+1)}.$$

3.2.2 Робусне методе

Пожељна особина оцена параметара неке расподеле је робусност, тј. примењљивост оцене при малом одступању од претпоставки модела. То важи и при моделовању екстремних вредности неког узорка.

Робусне оцене параметара облика и размере за $W_{\gamma,\sigma}$ расподелу дате су у радовима Peng and Welsh (2001) [34] и Juárez and Schucany (2004) [25]. Ове методе су рачунски захтевније од осталих. Резултати Монте Карло експеримената показују да се робусне оцене добро понашају на контаминираним узорцима (оним који не потичу из чистих расподела), а да у случају чистих узорака постоје ефикасније оцене од њих.

3.2.3 *Elemental percentile method*

Elemental percentile method (*EPM*) је метод оцењивања параметара GP(γ, σ) расподеле дефинисан у раду Castillo and Hadi (1997) [5]. Састоји се из две фазе:

1. налажење одређеног броја почетних оцена параметара γ и σ;
2. извођење коначних оцена параметара γ и σ на основу почетних.

1. Нека је $x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}$ варијациони низ статистика поретка за случајни узорак $(x_1, x_2, \ldots, x_n)$ из $\mathrm{GP}(\gamma, \sigma)$ расподеле и нека је

$$W_{\gamma,\sigma}(x_{i:n}) = p_{i:n}, \quad W_{\gamma,\sigma}(x_{j:n}) = p_{j:n}, \quad p_{i:n} = \frac{i-\alpha}{n+\beta}. \tag{3.11}$$

   (У раду [5] се препоручује коришћење параметара $\alpha = 0$ и $\beta = 1$.)

   Кад се уведе смена $\delta = \sigma/\gamma$ у једначини (3.11) и елиминише γ из добијеног система, добија се

$$\log(1-p_{i:n})\log\left(1+\frac{x_{j:n}}{\delta}\right) - \log(1-p_{j:n})\log\left(1+\frac{x_{i:n}}{\delta}\right) = 0. \tag{3.12}$$

   Из Теореме 1 из рада [5] следи да једначина (3.12), која има само једну непознату, δ, има реално решење $\delta(i,j)$, које може да се одреди методом половљења интервала. Заменом $\delta(i,j)$ у (3.11) добијају се оцене $\gamma(i,j)$ и $\sigma(i,j)$ за параметре γ и σ.

2. Кад су оцене $\gamma(i,j)$ и $\sigma(i,j)$ одређене за све могуће комбинације (i,j), онда се коначне оцене параметара γ и σ добијају као

$$\hat{\gamma} = m\left\{\gamma(1,2), \gamma(1,3), \ldots, \gamma(n-1,n)\right\}; \tag{3.13}$$

$$\hat{\sigma} = m\left\{\sigma(1,2), \sigma(1,3), \ldots, \sigma(n-1,n)\right\}, \tag{3.14}$$

   где је $m\{a_1, a_2, \ldots, a_n\}$ медијана узорка $\{a_1, a_2, \ldots, a_n\}$.

За велике узорке, број свих могућих парова (i,j) је велики, као и број потребних израчунавања. Један од начина да се тај проблем превазиђе је да се, уместо свих могућих парова, посматрају само случајеви $i \in \{1, 2, \ldots, n-1\}$ и $j = n$. То поједностављење је примењено и у овом раду.

Elemental percentile method има важну предност над неким другим методама оцењивања, а то је да оцене параметара постоје за све могуће вредности γ и σ, без ограничења.

3.3 Несагласност оцена параметара са узорком

Ако је $\gamma < 0$, онда је десни крај носача $\mathrm{GP}(\gamma, \sigma)$ расподеле једнак $-\sigma/\gamma$, тј. то је коначан број који зависи од параметара облика и размере. При оцењивању непознатих параметара оцењује се и ова величина. Дешава се да су добијене оцене несагласне са узорком на основу кога се врши оцењивање, тј. постоје чланови узорка који су већи од овако оцењеног десног краја носача расподеле. Ако је $x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}$ варијациони низ статистика поретка за тај узорак, несагласност се појављује кад је $-\sigma/\gamma < x_{n:n}$.

Овај проблем је први пут примећен у раду Dupuis (1996) [9], за случај методе момената и методе пондерисаних момената. У истом раду одређене су и вероватноће оваквих догађаја. У раду Castillo and Hadi (1997) [5] предложена је метода оцењивања (*elemental percentile method*) код које се не појављује овај проблем. Детаљна студија проблема рађена на симулираним узорцима из $\mathrm{GP}(\gamma, \sigma)$ расподеле приказана је у раду Ashkar and Tatsambon (2007), [1].

Упркос томе што данас постоје ефикасније методе оцењивања параметара $\mathrm{GP}(\gamma, \sigma)$ расподеле, поправка традиционалних оцена, као што су оцене методом момената и методом пондерисаних момената, значајна је због њихове једноставности и због тога што нису рачунски захтевне.

Посебна врста несагласности се појављује у случају кад је добијена оцена параметра облика позитивна, а постоји разлог да се верује (на пример, закључак добијен визуелним методама, нека нова информација која је у међувремену постала доступна, природа самог проблема...) да је параметар облика, заправо, негативан.

Један могући начин да се превазиђе проблем несагласности, у случају оцена методом момената и методом пондерисаних момената, дат је у раду Dupuis and Tsao (1998), [10]. Проблем је решен увођењем условног ограничења за параметар облика, чиме се исправља несагласност у случају да се она јави. Оне су у раду

[10] назване **хибридне оцене** и дате су са

$$\sigma_h = \hat{\sigma}, \quad \gamma_h = \begin{cases} -\frac{\hat{\sigma}}{x_{n:n}}, & \text{ако је } x_{n:n}\hat{\gamma} + \hat{\sigma} < 0 \text{ и } \hat{\gamma} < 0; \\ \hat{\gamma}, & \text{у другом случају,} \end{cases} \tag{3.15}$$

где су $\hat{\sigma}$ и $\hat{\gamma}$ обичне ММ или МПМ оцене.

Помоћу серије Монте Карло експеримената показано је да су хибридне оцене параметра облика, у много случајева, ефикасније (мање пристрасне и са мањом средњеквадратном грешком) од оригиналних оцена. Објашњење те чињенице је да хибридна оцена користи више података о узорку него обична ММ или МПМ оцена.

Техника за поправку несагласних оцена параметара $\mathrm{GP}(\gamma, \sigma)$ расподела предложена о овом раду полази од сличне идеје, али је општија (зависи од два додатна параметра), па се може применити и на друге методе оцењивања и друге расподеле. Такође, новом методом се могу побољшати особине оцена и параметра облика и параметра размере истовремено, што није случај са хибридном оценом предложеном у раду [10].

3.3.1 Предложена корекција

Нека је $x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}$ варијациони низ статистика поретка за узорак из $\mathrm{GP}(\gamma, \sigma)$ расподеле чије параметре треба оценити. Нека су $\hat{\gamma}$ и $\hat{\sigma}$ оцене параметара облика и размере добијене неком од метода оцењивања, такве да постоји несагласност са узорком, тј. $\hat{\gamma} < 0$ и $x_{n:n} > -\hat{\sigma}/\hat{\gamma}$. Циљ је одредити поправљене оцене за параметре размере и облика, $\tilde{\sigma}$ и $\tilde{\gamma}$, такве да важи $-\tilde{\sigma}/\tilde{\gamma} \geq x_{n:n}$.
Нове оцене параметара, $\tilde{\gamma}$ и $\tilde{\sigma}$ се дефинишу као

$$\tilde{\sigma} = \hat{\sigma} + \alpha p, \quad \tilde{\gamma} = \hat{\gamma} + (1-\alpha)p, \quad \alpha \in \mathbb{R}. \tag{3.16}$$

Ако важи услов

$$\hat{\gamma} + (1-\alpha)p < 0, \tag{3.17}$$

онда из неједнакости

$$\frac{\hat{\sigma} + \alpha p}{\hat{\gamma} + (1-\alpha)p} \leq -x_{n:n}, \tag{3.18}$$

следи

$$p \geq \frac{-x_{n:n}\hat{\gamma} - \hat{\sigma}}{\alpha + (1-\alpha)x_{n:n}}, \quad \text{ако је} \quad \alpha + (1-\alpha)x_{n:n} > 0, \tag{3.19а}$$

$$p \leq \frac{-x_{n:n}\hat{\gamma} - \hat{\sigma}}{\alpha + (1-\alpha)x_{n:n}}, \quad \text{ако је} \quad \alpha + (1-\alpha)x_{n:n} < 0. \tag{3.19б}$$

Корекција дата са

$$p = \frac{-x_{n:n}\hat{\gamma} - \hat{\sigma}}{\alpha + (1-\alpha)x_{n:n}} \tag{3.20}$$

решава проблем несагласности у оба случаја.

Неједнакост (3.17), где је p задато са (3.20), важи за свако $\hat{\gamma} < 0$, ако су изрази

$$\alpha\hat{\gamma} - (1-\alpha)\hat{\sigma} \quad \text{и} \quad \alpha + (1-\alpha)x_{n:n} \tag{3.21}$$

различитог знака. Због једноставности, за параметар α се у даљем раду претпоставља да припада интервалу

$$\alpha \in \left(-\frac{\hat{\sigma}}{\mid \hat{\sigma} + \hat{\gamma} \mid}, \frac{\hat{\sigma}}{\mid \hat{\sigma} + \hat{\gamma} \mid}\right), \quad \hat{\gamma} \neq -\hat{\sigma}, \tag{3.22}$$

који садржи део решења одговарајућег система неједначина.

Улога параметра α је да контролише која од оцена параметара, ($\hat{\gamma}$ или $\hat{\sigma}$, ће бити "више промењена". "Природан" избор за α би биле вредности из интервала $[0, 1]$. Међутим, може се показати (у следећем одељку), да неке друге вредности α из интервала (3.22) дају поправљене оцене са бољим особинама (мање пристрасне и са мањом средњеквадратном грешком).

Такође, могуће је увести додатни параметар, β, који означава ниво корекције, тј. контролише колико ће највећи члан узорка бити удаљен од новооцењеног десног краја носача расподеле. У том случају, поправљене оцене задовољавају једнакост

$$\frac{\tilde{\sigma}}{\tilde{\gamma}} = \frac{\hat{\sigma} + \alpha p}{\hat{\gamma} + (1-\alpha)p} = -\beta x_{n:n}, \quad \text{где је} \quad \beta \geq 1. \tag{3.23}$$

Коначно, оцене параметра размере и облика су дате са

$$\tilde{\sigma} = \begin{cases} \hat{\sigma} + \alpha p, & \text{ако је } x_{n:n}\hat{\gamma} + \hat{\sigma} < 0 \text{ и } \hat{\gamma} < 0, \text{ или } \hat{\gamma} > 0; \\ \hat{\sigma}, & \text{у другом случају,} \end{cases} \tag{3.24а}$$

$$\tilde{\gamma} = \begin{cases} \hat{\gamma} + (1-\alpha)p, & \text{ако је } x_{n:n}\hat{\gamma} + \hat{\sigma} < 0 \text{ и } \hat{\gamma} < 0, \text{ или } \hat{\gamma} > 0; \\ \hat{\gamma}, & \text{у другом случају}, \end{cases} \tag{3.24б}$$

где је

$$p = \frac{-\beta\hat{\gamma}x_{n:n} - \hat{\sigma}}{\alpha + (1-\alpha)\beta x_{n:n}} \quad \text{и} \quad \beta \geq 1. \tag{3.25}$$

У случају $\hat{\gamma} > 0$ (кад је оцена параметра облика позитивна, а природа проблема налаже другачије) поправљене оцене имају исти облик. Услов $\hat{\gamma} > 0$ није обавезан. Ако се изостави, позитивне оцене параметра облика неће бити исправљене.

Неке особине поправљених оцена су:

1. Ако је $\alpha = 0$ параметар размере је оцењен оригиналном (непоправљеном) методом, а несагласност је поправљена. За $\beta = 1$ ово је, заправо, хибридна оцена предложена у раду [10].

2. За $\alpha = 1$ параметар облика је оцењен оригиналном (непоправљеном) методом, а несагласност је поправљена.

3. За $x_{n:n}\hat{\gamma}+\hat{\sigma} > 0$ (случај када нема несагласности), корекција се не примењује. Оба параметра су оцењена почетном методом.

3.3.2 Симулације

Да би се одредиле одговарајуће вредности α из интервала (3.22) за корекцију ММ и МПМ оцена, параметар α је дефинисан као

$$\alpha = \begin{cases} i\hat{\sigma}, & \hat{\gamma} = -\hat{\sigma}; \\ \max\left\{ i\hat{\sigma}, j\frac{\hat{\sigma}}{|\hat{\sigma}+\hat{\gamma}|} \right\}, & \hat{\gamma} \neq -\hat{\sigma}, \end{cases} \tag{3.26}$$

где је $i, j \in \{-1, -0.9, -0.8, \ldots, 0.8, 0.9, 1\}$, и где су $\hat{\gamma}, \hat{\sigma}$ почетне ММ или МПМ оцене, без корекције. Члан $i\hat{\sigma}$ се додаје да би се спречило да α узме јако велике вредности у случају кад је $\hat{\sigma}+\hat{\gamma}$ блиско нули. Параметер β дефинисан је на интервалу $[1, 2]$.

После извршеног великог броја Монте Карло експеримената, са симулираним подацима из $\mathrm{GP}(\gamma, \sigma)$ расподела са разним вредностима параметара, примећено је да су за параметар α прихватљиве вредности добијене за $i, j \leq 0$. За ове вредности, осим што се исправља несагласност, добијају се оцене непознатих параметара које су мање пристрасне и са мањом средњеквадратном грешком од почетних. Оне се добро понашају чак и за јако велике и јако мале вредности $\hat{\sigma} > 0$. Прецизније:

1. ако $\hat{\sigma} \downarrow 0$, из (3.26) следи $\alpha \to 0$;
2. ако $\hat{\sigma} \uparrow \infty$, из (3.26) следи $\alpha \to j < 0$,

па следи да α увек остаје у прихватљивом интервалу.

Што се тиче параметра β, показује се да вредности из интервала $[1, 1.5]$ задовољавајуће.

У другој фази експеримента, извршена је детаљна анализа оцена добијених за конкретан избор параметара α и β. За параметар α је узето

$$\alpha = \begin{cases} -0.5\hat{\sigma}, & \hat{\gamma} = -\hat{\sigma}; \\ \max\left\{-0.5\hat{\sigma}, -\frac{0.9\hat{\sigma}}{|\hat{\sigma}+\hat{\gamma}|}\right\}, & \hat{\gamma} \neq -\hat{\sigma}, \end{cases} \tag{3.27}$$

што је добијено случајним одабиром пара (i, j) из прихватљивог интервала.

Симулације за конкретан избор параметара α и β

Да би се проверило понашање поправљених ММ и МПМ оцена, генерисано је по 1000 симулираних узорака из GP(γ, σ) расподеле за сваку од комбинација $(n, \gamma, \sigma, \beta)$ величине узорка, $n \in \{15, 50, 100\}$, параметара расподеле, $\gamma \in \{-0.2, -0.4, -0.6, -0.8, -1, -2\}$ и $\sigma \in \{0.1, 0.2, 0.4, 0.6, 0.8, 1, 2, 5, 10, 20\}$, као и параметра корекције, $\beta \in \{1, 1.01, 1.5\}$. Параметар α је дефинисан као у (3.27).

Почетне ММ и МПМ оцене упоређене су са поправљеним оценама у односу на следеће карактеристике: пристрасност и средњеквадратна грешка оцена параметара облика и размере, пристрасност и средњеквадратна грешка оцена 99% и 99.9%-квантила подељених са стварним вредностима квантила који се оцењују (Табела 3.1) и број симулираних узорака (од 1000 укупно) на којима се појавила несагласност.

Добијени резултати се могу интерпретирати на следећи начин:

1. Број симулираних узорака на којима се појавила несагласност (Табела 3.2), слаже се са бројем таквих узорака добијених у претходним студијама [1, 5, 10].

2. Поправљене ММ и МПМ оцене увек су сагласне са узорцима на основу којих се вршило оцењивање, тј. несагласност је комплетно нестала.

3. За обе методе, ММ и МПМ, код свих посматраних узорака, средњеквадратна грешка поправљених оцена параметара облика и размере је мања него код почетних оцена. Такође, у већини случајева, пристрасност поправљене оцене је мања него код почетне оцене. Једини изузетак су узорци из расподеле GP$(-2, \sigma)$.

4. За обе методе, ММ и МПМ, средњеквадратна грешка оцена 99% и 99.9%-квантила је мања код поправљених него код почетних оцена. Што се тиче пристрасности оцена, у неким случајевима је благо повећана, у другима смањена. Дакле, поправљене оцене квантила су прихватљиве исто колико и оцене пре корекције.

5. Најзначајнија разлика у понашању поправљених оцена (најмања пристрасност и средњеквадратна грешка) добијени су за величину узорка $n = 15$.

Резултати добијени за $\sigma = 1$ (овај случај је одабран зато што је често приказан у литератури) и $\beta = 1.01$ су приказани у Табелама 3.2 - 3.5. Сви остали експерименти доводе до закључака сличних већ наведенима.

Додатно, да би се упоредило понашање поправљених и хибридних оцена, одређени су количници апсолутних пристрасности и количници средњеквадратних грешака оцена параметара облика, размере, 99% и 99.9%-квантила. Кад је неки од ових количника мањи од 1 поправљена оцена је ефикаснија од хибридне, а ако је количник већи од 1 важи обрнуто. Ови резултати приказани су у Табели 3.6.

Табела 3.1: Стварне вредности 99% и 99.9%-квантила

γ	99%-квантил	99.9%-квантил
-0.2	3.01	3.74
-0.4	2.10	2.34
-0.6	1.56	1.64
-0.8	1.22	1.25
-1	0.99	1.00
-2	0.50	0.50

Табела 3.2: Број несагласних симулираних узорака за ММ и МПМ методу (број узорака где су добијене позитивне оцене параметра облика дат је у загради)

		ММ	МПМ	ММ	МПМ	ММ	МПМ
γ	n	15		50		100	
-0.2		87 (134)	106 (209)	34 (49)	94 (109)	45 (7)	86 (37)
-0.4		129 (44)	136 (102)	163 (1)	201 (16)	166 (0)	206 (0)
-0.6		228 (11)	240 (32)	306 (0)	325 (0)	302 (0)	322 (0)
-0.8		283 (2)	249 (16)	352 (0)	352 (0)	382 (0)	381 (0)
-1		339 (0)	291 (3)	395 (0)	377 (0)	444 (0)	437 (0)
-2		438 (0)	301 (0)	473 (0)	400 (0)	507 (0)	465 (0)

Табела 3.3: Пристрасност (средњеквадратна грешка) при оцењивању параметара облика и размере, у случајевима кад су добијене позитивне оцене параметра облика. Позитивне оцене параметра облика су поправљене. Методе: 1-ММ, 2-поправљена ММ, 3-МПМ, 4-поправљена МПМ

	n	γ	1	2	3	4
пар. γ						
	15	-0.2	0.1(0.34)	0.14(0.27)	0.08(0.37)	0.16(0.30)
		-0.4	0.09(0.40)	0.08(0.32)	0.07(0.41)	0.08(0.32)
		-0.6	0.11(0.52)	0.05(0.38)	0.08(0.48)	0.04(0.37)
		-0.8	0.13(0.64)	0.04(0.45)	0.08(0.56)	0.02(0.44)
		-1	0.12(0.69)	-0.01(0.46)	0.04(0.57)	-0.03(0.46)
	50	-0.2	0.02(0.15)	0.03(0.14)	0.02(0.18)	0.04(0.14)
		-0.4	0.03(0.18)	0.01(0.15)	0.02(0.20)	0.01(0.17)
	100	-0.2	0.02(0.11)	0.02(0.10)	0.01(0.13)	0.02(0.11)
пар. σ						
	15	-0.2	0.11(0.44)	0.11(0.42)	0.09(0.46)	0.11(0.43)
		-0.4	0.09(0.45)	0.08(0.41)	0.08(0.45)	0.07(0.41)
		-0.6	0.1(0.49)	0.07(0.42)	0.07(0.45)	0.05(0.41)
		-0.8	0.12(0.54)	0.07(0.44)	0.08(0.48)	0.05(0.42)
		-1	0.09(0.50)	0.03(0.38)	0.05(0.42)	0.02(0.37)
	50	-0.2	0.02(0.20)	0.02(0.20)	0.02(0.22)	0.02(0.21)
		-0.4	0.02(0.20)	0.02(0.20)	0.02(0.22)	0.01(0.20)
	100	-0.2	0.02(0.15)	0.01(0.15)	0.01(0.16)	0.01(0.15)

Табела 3.4: Пристрасност (средњеквадратна грешка) при оцењивању параметара облика и размере. Позитивне оцене параметра облика нису поправљене. Методе: 1-ММ, 2-поправљена ММ, 3-МПМ, 4-поправљена МПМ

	n	γ	1	2	3	4
пар. γ						
	15	-0.2	0.1(0.35)	0.09(0.32)	0.08(0.38)	0.06(0.34)
		-0.4	0.1(0.43)	0.07(0.34)	0.08(0.43)	0.05(0.36)
		-0.6	0.1(0.53)	0.05(0.40)	0.07(0.49)	0.02(0.41)
		-0.8	0.16(0.65)	0.06(0.45)	0.1(0.56)	0.03(0.45)
		-1	0.12(0.72)	-0.01(0.47)	0.04(0.58)	-0.03(0.47)
		-2	0.46(1.97)	-0.31(0.77)	0.09(1.09)	-0.19(0.79)
	50	-0.2	0.03(0.16)	0.02(0.15)	0.02(0.19)	0.01(0.17)
		-0.4	0.02(0.17)	0.01(0.16)	0.02(0.20)	0(0.17)
		-0.6	0.02(0.23)	-0.01(0.18)	0.01(0.24)	-0.02(0.20)
		-0.8	0.03(0.28)	-0.02(0.21)	0.02(0.28)	-0.03(0.22)
		-1	0.05(0.34)	-0.02(0.25)	0.03(0.32)	-0.02(0.25)
		-2	0.1(0.72)	-0.23(0.43)	0.02(0.57)	-0.15(0.41)
	100	-0.2	0.01(0.10)	0.01(0.10)	0.01(0.12)	0.01(0.11)
		-0.4	0.01(0.12)	0(0.11)	0.01(0.14)	0(0.12)
		-0.6	0(0.14)	-0.02(0.12)	0(0.16)	-0.03(0.13)
		-0.8	0.01(0.19)	-0.02(0.15)	0.01(0.19)	-0.03(0.15)
		-1	0.02(0.23)	-0.03(0.17)	0.01(0.23)	-0.04(0.18)
		-2	0.06(0.47)	-0.16(0.31)	0.02(0.39)	-0.11(0.29)
пар. σ						
	15	-0.2	0.1(0.44)	0.1(0.43)	0.09(0.46)	0.08(0.44)
		-0.4	0.09(0.46)	0.08(0.42)	0.08(0.45)	0.06(0.42)
		-0.6	0.09(0.49)	0.06(0.42)	0.07(0.45)	0.05(0.41)
		-0.8	0.13(0.55)	0.08(0.44)	0.1(0.48)	0.06(0.42)
		-1	0.09(0.52)	0.03(0.39)	0.05(0.44)	0.01(0.38)
		-2	0.21(0.86)	-0.12(0.36)	0.06(0.50)	-0.05(0.37)
	50	-0.2	0.03(0.21)	0.03(0.20)	0.02(0.22)	0.02(0.22)
		-0.04	0.02(0.20)	0.02(0.19)	0.02(0.21)	0.01(0.20)
		-0.6	0.02(0.22)	0.01(0.21)	0.02(0.23)	0(0.21)
		-0.8	0.03(0.23)	0.01(0.21)	0.02(0.23)	0(0.21)
		-1	0.04(0.25)	0.01(0.21)	0.03(0.24)	0.01(0.21)
		-2	0.04(0.31)	-0.09(0.19)	0.02(0.25)	-0.05(0.19)
	100	-0.2	0.01(0.14)	0.01(0.14)	0.01(0.15)	0.01(0.15)
		-0.4	0.01(0.15)	0.01(0.14)	0.01(0.15)	0(0.15)
		-0.6	0(0.14)	-0.01(0.14)	0(0.15)	-0.01(0.14)
		-0.8	0.01(0.16)	0(0.14)	0.01(0.16)	0(0.14)
		-1	0.01(0.17)	0(0.15)	0.01(0.16)	-0.01(0.15)
		-2	0.02(0.20)	-0.06(0.14)	0.01(0.17)	-0.04(0.13)

Табела 3.5: Пристрасност (средњеквадратна грешка) при оцењивању 99% и 99.9%-квантила. Позитивне оцене параметра облика нису поправљене. Методе: 1-ММ, 2-поправљена ММ, 3-МПМ, 4-поправљена МПМ

	n	γ	1	2	3	4
99% кв.						
	15	-0.2	-0.03(0.29)	-0.03(0.29)	0.02(0.36)	0.03(0.36)
		-0.4	-0.02(0.24)	-0.01(0.23)	0.02(0.30)	0.03(0.29)
		-0.6	0.01(0.21)	0.03(0.20)	0.04(0.26)	0.06(0.25)
		-0.8	0.02(0.19)	0.04(0.17)	0.04(0.22)	0.06(0.21)
		-1	0.04(0.18)	0.06(0.16)	0.06(0.20)	0.08(0.19)
		-2	0.03(0.13)	0.07(0.12)	0.05(0.13)	0.07(0.12)
	50	-0.2	-0.01(0.16)	-0.01(0.16)	0.01(0.20)	0.02(0.19)
		-0.4	-0.01(0.13)	0(0.12)	0(0.16)	0.02(0.14)
		-0.6	0.01(0.11)	0.03(0.10)	0.02(0.14)	0.04(0.12)
		-0.8	0.01(0.10)	0.03(0.08)	0.01(0.12)	0.04(0.10)
		-1	0.01(0.09)	0.03(0.07)	0.01(0.09)	0.04(0.08)
		-2	0.01(0.06)	0.03(0.05)	0.02(0.05)	0.03(0.05)
	100	-0.2	-0.01(0.11)	0(0.11)	0(0.14)	0.01(0.13)
		-0.4	0(0.09)	0.01(0.08)	0.01(0.11)	0.02(0.10)
		-0.6	0.01(0.08)	0.02(0.07)	0.01(0.09)	0.03(0.08)
		-0.8	0(0.07)	0.02(0.06)	0.01(0.08)	0.03(0.06)
		-1	0.01(0.07)	0.03(0.05)	0.01(0.07)	0.03(0.06)
		-2	0(0.04)	0.02(0.03)	0.01(0.03)	0.02(0.03)
99.9% кв.						
	15	-0.2	0.02(0.45)	0.03(0.45)	0.17(0.76)	0.18(0.76)
		-0.4	0.04(0.36)	0.05(0.35)	0.13(0.56)	0.14(0.55)
		-0.6	0.07(0.34)	0.09(0.33)	0.14(0.50)	0.16(0.49)
		-0.8	0.06(0.27)	0.08(0.26)	0.14(0.37)	0.12(0.36)
		-1	0.07(0.25)	0.10(0.24)	0.11(0.32)	0.13(0.32)
		-2	0.04(0.15)	0.07(0.14)	0.06(0.15)	0.07(0.15)
	50	-0.2	0.01(0.25)	0.01(0.25)	0.06(0.35)	0.07(0.34)
		-0.4	0.01(0.19)	0.02(0.18)	0.04(0.26)	0.05(0.25)
		-0.6	0.02(0.16)	0.04(0.15)	0.04(0.20)	0.06(0.19)
		-0.8	0.02(0.13)	0.05(0.12)	0.03(0.15)	0.05(0.14)
		-1	0.02(0.11)	0.05(0.09)	0.03(0.12)	0.05(0.11)
		-2	0.01(0.07)	0.03(0.06)	0.02(0.05)	0.03(0.05)
	100	-0.2	0(0.16)	0(0.16)	0.02(0.22)	0.03(0.22)
		-0.4	0.01(0.13)	0.02(0.12)	0.02(0.17)	0.03(0.16)
		-0.6	0.02(0.11)	0.03(0.10)	0.03(0.13)	0.05(0.12)
		-0.8	0.01(0.09)	0.03(0.07)	0.01(0.10)	0.04(0.09)
		-1	0.01(0.08)	0.03(0.06)	0.02(0.08)	0.04(0.07)
		-2	0(0.04)	0.02(0.03)	0.01(0.03)	0.02(0.03)

Табела 3.6: Однос апсолутних пристрасности (средњеквадратних грешака) поправљених и хибридних оцена. Методе: 1-ММ, 2-МПМ

n	γ	1	2	1	2
		пар. γ		пар. σ	
15	-0.2	0.96 (0.97)	0.92 (0.97)	0.94 (0.97)	0.88 (0.97)
	-0.4	0.86 (0.94)	0.71 (0.95)	0.83 (0.94)	0.75 (0.94)
	-0.6	0.57 (0.90)	0.11 (0.93)	0.66 (0.89)	0.58 (0.92)
	-0.8	0.09 (0.83)	6.61 (0.91)	0.46 (0.82)	0.38 (0.89)
	-1	0.41 (0.75)	13.22 (0.88)	0.21 (0.72)	0.16 (0.87)
	-2	1.03 (0.42)	29.67 (0.75)	0.63 (0.39)	1.62 (0.73)
50	-0.2	0.96 (0.99)	0.77 (0.98)	0.93 (0.99)	0.72 (0.97)
	-0.4	0.26 (0.97)	1.88 (0.96)	0.60 (0.96)	0.16 (0.95)
	-0.6	2.51 (0.93)	1.64 (0.94)	0.09 (0.92)	0.61 (0.92)
	-0.8	2.47 (0.90)	1.74 (0.92)	0.31 (0.89)	1.42 (0.90)
	-1	3.21 (0.87)	1.95 (0.90)	0.57 (0.85)	2.16 (0.89)
	-2	8.41 (0.66)	7.12 (0.75)	2.69 (0.61)	7.70 (0.72)
100	-0.2	0.97 (0.99)	0.83 (0.98)	0.95 (0.99)	0.82 (0.98)
	-0.4	0.25 (0.97)	2.76 (0.96)	0.68 (0.96)	0.47 (0.95)
	-0.6	3.14 (0.93)	1.95 (0.94)	0.29 (0.93)	0.04 (0.92)
	-0.8	2.96 (0.91)	2.14 (0.92)	0.05 (0.89)	0.34 (0.90)
	-1	4.00 (0.88)	2.55 (0.90)	0.28 (0.86)	0.65 (0.88)
	-2	6.80 (0.73)	50.71 (0.78)	2.18 (0.69)	2.93 (0.76)
		99% кв.		99.9% кв.	
15	-0.2	0.99 (1.00)	0.10 (0.58)	0.11 (1.33)	1.00 (1.00)
	-0.4	1.11 (1.00)	0.28 (0.66)	1.32 (1.40)	1.01 (1.00)
	-0.6	1.03 (1.00)	0.48 (0.71)	4.83 (1.37)	1.01 (1.00)
	-0.8	1.03 (1.00)	0.65 (0.76)	1.80 (1.31)	1.01 (1.00)
	-1	1.04 (1.00)	0.78 (0.79)	1.39 (1.26)	1.02 (1.00)
	-2	1.05 (1.00)	0.96 (0.84)	1.14 (1.19)	1.04 (1.00)
50	-0.2	0.99 (1.00)	0.06 (0.47)	0.06 (0.99)	1.01 (1.00)
	-0.4	1.05 (1.00)	0.20 (0.58)	0.44 (1.21)	1.02 (1.00)
	-0.6	1.06 (0.99)	0.41 (0.68)	4.12 (1.33)	1.03 (1.00)
	-0.8	1.07 (0.99)	0.63 (0.76)	2.56 (1.28)	1.05 (1.00)
	-1	1.09 (0.99)	0.80 (0.82)	1.56 (1.21)	1.07 (1.00)
	-2	1.15 (1.01)	1.00 (0.83)	1.32 (1.22)	1.14 (1.01)
100	-0.2	0.97 (1.00)	0.04 (0.40)	0.03 (0.78)	1.01 (1.00)
	-0.4	1.06 (0.99)	0.15 (0.52)	0.24 (1.01)	1.03 (1.00)
	-0.6	1.09 (0.99)	0.36 (0.64)	1.37 (1.25)	1.07 (0.99)
	-0.8	1.12 (0.99)	0.59 (0.74)	4.87 (1.29)	1.10 (0.99)
	-1	1.14 (1.00)	0.79 (0.81)	1.84 (1.22)	1.12 (1.00)
	-2	1.25 (1.02)	1.06 (0.82)	1.49 (1.27)	1.26 (1.02)

Робусност

Да би се упоредила робусност почетних и поправљених ММ и МПМ оцена, серија Монте Карло експеримената је изведена на следећи начин (по угледу на радове [25] и [34]).

Дефинишу се два типа контаминације:

1. *контаминација параметра облика:* случајни узорци се генеришу из мешавине расподела $0.9W_{\gamma,\sigma} + 0.1W_{2\gamma,\sigma}$;

2. *контаминација параметра размере:* случајни узорци се генеришу из мешавине расподела $0.9W_{\gamma,\sigma} + 0.1W_{\gamma,2\sigma}$.

За сваку комбинацију (σ, γ, n) параметара расподеле $\sigma \in \{0.1, 1, 10\}$, $\gamma \in \{-0.2, -0.4, -0.6, -0.8, -1, -2\}$ и величине узорка $n \in \{15, 50, 100\}$, генерисано је по 1000 узорака са контаминацијом сваког од типова. Параметар α је задат са (3.27), а $\beta = 1.01$.

Примећено је да се променом параметра σ резултати не мењају много, као и да разлика у ефикасности између почетних и поправљених оцена постаје све мање значајна кад величина узорка расте. Због тога су наведени само резултати за случај $\sigma = 1$ и $n = 15$ и приказани су у Табелама 3.7 и 3.8.

Закључци су следећи:

1. Обе поправљене оцене (ММ и МПМ) боље реагују на контаминацију првог него на контаминацију другог типа. То такође важи за почетне (непоправљене) ММ и МПМ оцене.

2. Поправљене оцене параметра облика и параметра размере имају мање средњеквадратне грешке у поређењу са оценама истих параметара пре корекције (осим у случају $\gamma = -2$). У већини случајева поправљене оцене су и мање пристрасне.

3. Поправљене оцене квантила имају мало већу пристрасност него оцене пре корекције. С друге стране, средњеквадратна грешка поправљених оцена квантила је мања.

Табела 3.7: Пристрасност (средњеквадратна грешка) добијени за узорке са контаминацијом првог типа. Методе: 1-ММ, 2-поправљена ММ, 3-МПМ, 4-поправљена МПМ

	n	γ	1	2	3	4
пар. γ						
	15	-0.2	0.11(0.34)	0.1(0.31)	0.09(0.37)	0.07(0.33)
		-0.4	0.1(0.4)	0.07(0.33)	0.08(0.41)	0.05(0.35)
		-0.6	0.09(0.48)	0.03(0.36)	0.07(0.46)	0.01(0.37)
		-0.8	0.09(0.57)	0(0.4)	0.05(0.51)	-0.03(0.4)
		-1	0.1(0.69)	-0.05(0.43)	0.03(0.57)	-0.07(0.44)
		-2	0.18(1.51)	-0.49(0.74)	-0.13(0.92)	-0.39(0.73)
пар. σ						
	15	-0.2	0.08(0.41)	0.08(0.39)	0.07(0.43)	0.06(0.41)
		-0.4	0.06(0.42)	0.05(0.39)	0.05(0.42)	0.04(0.39)
		-0.6	0.05(0.43)	0.03(0.38)	0.04(0.41)	0.01(0.37)
		-0.8	0.05(0.46)	0(0.37)	0.02(0.41)	-0.01(0.36)
		-1	0.05(0.48)	-0.02(0.35)	0.01(0.40)	-0.03(0.35)
		-2	0.06(0.63)	-0.21(0.34)	-0.05(0.41)	-0.16(0.33)
99% кв.						
	15	-0.2	-0.06(0.29)	-0.06(0.29)	-0.02(0.35)	-0.01(0.34)
		-0.4	-0.03(0.24)	-0.02(0.23)	-0.01(0.29)	0.01(0.28)
		-0.6	-0.02(0.21)	0.01(0.19)	0.01(0.25)	0.03(0.23)
		-0.8	0(0.19)	0.03(0.17)	0.02(0.22)	0.05(0.20)
		-1	0.01(0.18)	0.04(0.15)	0.03(0.20)	0.06(0.18)
		-2	0.02(0.13)	0.06(0.11)	0.04(0.12)	0.06(0.11)
99.9% кв.						
	15	-0.2	-0.02(0.45)	-0.01(0.44)	0.11(0.70)	0.12(0.69)
		-0.4	0.02(0.37)	0.03(0.36)	0.1(0.54)	0.11(0.53)
		-0.6	0.04(0.31)	0.06(0.30)	0.09(0.43)	0.11(0.42)
		-0.8	0.04(0.27)	0.07(0.25)	0.08(0.35)	0.11(0.34)
		-1	0.04(0.23)	0.08(0.22)	0.07(0.29)	0.1(0.28)
		-2	0.03(0.14)	0.07(0.13)	0.05(0.14)	0.07(0.14)

Табела 3.8: Пристрасност (средњеквадратна грешка) добијени за узорке са контаминацијом другог типа. Методе: 1-ММ, 2-поправљена ММ, 3-МПМ, 4-поправљена МПМ

	n	γ	1	2	3	4
пар. γ						
	15	-0.2	0.04(0.33)	0.03(0.3)	0.02(0.37)	0.01(0.33)
		-0.4	0(0.39)	-0.03(0.34)	-0.01(0.41)	-0.04(0.36)
		-0.6	-0.04(0.48)	-0.09(0.39)	-0.05(0.47)	-0.09(0.4)
		-0.8	-0.08(0.58)	-0.15(0.45)	-0.08(0.53)	-0.16(0.45)
		-1	-0.11(0.7)	-0.24(0.52)	-0.12(0.6)	-0.24(0.52)
		-2	-0.27(1.52)	-0.92(1.1)	-0.36(1.03)	-0.87(1.08)
пар. σ						
	15	-0.2	0.13(0.44)	0.13(0.42)	0.12(0.46)	0.11(0.44)
		-0.4	0.1(0.44)	0.09(0.41)	0.1(0.45)	0.08(0.42)
		-0.6	0.08(0.45)	0.06(0.40)	0.08(0.44)	0.06(0.40)
		-0.8	0.06(0.46)	0.03(0.38)	0.06(0.44)	0.03(0.38)
		-1	0.05(0.48)	-0.01(0.36)	0.05(0.43)	0(0.36)
		-2	0.01(0.62)	-0.25(0.34)	-0.02(0.43)	-0.22(0.33)
99% кв.						
	15	-0.2	0.14(0.45)	0.14(0.44)	0.2(0.54)	0.2(0.54)
		-0.4	0.21(0.47)	0.22(0.46)	0.24(0.52)	0.26(0.52)
		-0.6	0.26(0.49)	0.28(0.49)	0.27(0.52)	0.3(0.53)
		-0.8	0.3(0.51)	0.33(0.52)	0.29(0.51)	0.34(0.54)
		-1	0.33(0.53)	0.38(0.55)	0.31(0.50)	0.38(0.55)
		-2	0.35(0.50)	0.52(0.66)	0.31(0.44)	0.52(0.65)
99.9% кв.						
	15	-0.2	0.29(0.80)	0.29(0.80)	0.48(1.24)	0.49(1.24)
		-0.4	0.39(0.84)	0.4(0.83)	0.49(1.12)	0.51(1.12)
		-0.6	0.44(0.84)	0.46(0.84)	0.48(1.00)	0.52(1.00)
		-0.8	0.45(0.81)	0.49(0.81)	0.46(0.88)	0.51(0.90)
		-1	0.45(0.77)	0.51(0.78)	0.43(0.78)	0.51(0.82)
		-2	0.39(0.58)	0.58(0.74)	0.34(0.51)	0.57(0.72)

3.3.3 Примери

Понашање поправљених ММ и МПМ оцена илустровано је на два конкретна узорка: *Bilbao waves data* (анализиран је у раду [5]) и *Fish river data* (анализиран је у раду [1]), код којих се, приликом апроксимације генералисаним Паретовим расподелама, појавила несагласност.

За параметре γ и σ, као и 99% и 99.9%-квантиле $\mathrm{GP}(\gamma, \sigma)$ расподеле одређене су следеће оцене: ММ и МПМ (пре корекције), ММ и МПМ (хибридна оцена предложена у раду [10]), ММ и МПМ (са поправкама предложеним у овом раду), ММВ и *EPM* (оцене код којих се обично не појављује несагласност). Резултати су дати у Табелама 3.10 и 3.12.

Да би се оценио квалитет апроксимације, коришћене су следеће дефиниције грешке:

$$\mathrm{ASAE} = \frac{1}{n}\sum_{i=1}^{n} \frac{|x_{i:n} - \hat{x}_{i:n}|}{x_{n:n} - x_{1:n}}, \quad \hat{x}_{i:n} = \frac{\hat{\sigma}}{\hat{\gamma}}\left(1 - (1 - p_{i:n})^{\hat{\gamma}}\right), \quad p_{i:n} = \frac{i}{n+1},$$

и

$$\mathrm{SSQ} = \sum_{i=1}^{n} (F_{\hat{\gamma},\hat{\sigma}}(x_{i:n}) - \hat{F}(x_{i:n}))^2,$$

где је $\hat{F}$ емпиријска функција расподеле.

Bilbao waves data

Узорак под именом *Bilbao waves data* анализиран је у раду [5], да би се оценила ефикасност *elemental percentile* оцене. Подаци, који су апроксимирани $\mathrm{GP}(\gamma, \sigma)$ расподелом, састоје се од 179 опсервација које се односе на периоде морских таласа (у секундама), измерених у заливу Билбао јануара 1997. Подаци потичу од Maritime Climate Program of CEDEX, Spain. Сва мерења су већа од 7 секунди, па су то, заправо, прекорачења прага $u = 7$. Ове вредности, поређане по величини и умањене за вредност прага, дате су у Табели 3.9.

Када су за параметре одговарајуће расподеле одређене ММ и МПМ оцене примећено је да овај узорак показује несагласност. Код ММ оцене, 13 чланова узорка је било веће од оцењене горње границе, а код МПМ то је било 17 чланова.

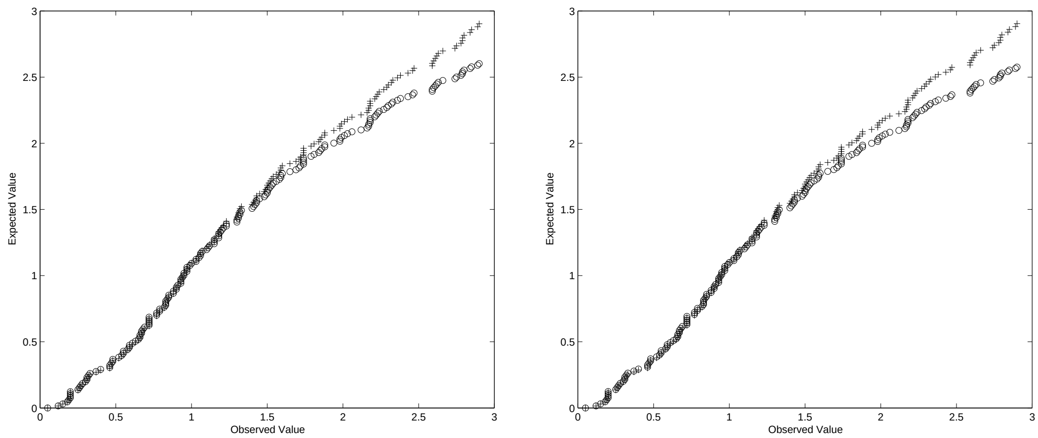
Оцењене грешке (ASAE и SSQ, Табела 3.10) показују да су оцене предложене у овом раду дале бољу апроксимацију него хибридне. На основу qq-дијаграма (Слика 3.1) може се закључити да обе поправљене оцене дају бољу апроксимацију репа расподеле него оцене пре корекције. Такође, Слика 3.2 и Слика 3.3 показују да су обе поправљене оцене једнако прихватљиве као *EPM* и MMB. Нарочито добра сагласност ових метода је примећена у репу расподеле.

Табела 3.9: *The Bilbao waves data*: прекорачења прага од 7 секунди, поређана од најмањег до највећег

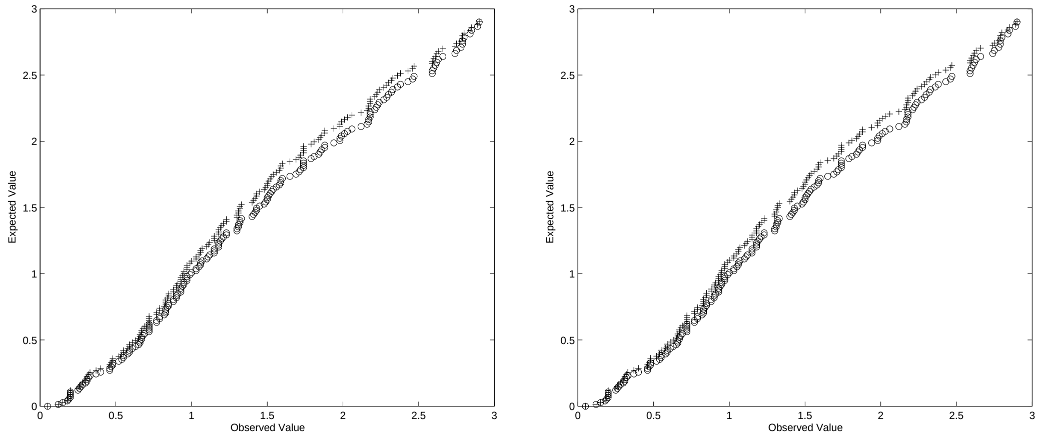
0.05	0.12	0.15	0.18	0.19	0.20	0.20	0.20	0.20	0.25	0.26
0.27	0.28	0.30	0.31	0.31	0.32	0.33	0.37	0.40	0.46	0.46
0.47	0.48	0.48	0.52	0.54	0.55	0.55	0.58	0.59	0.59	0.61
0.63	0.65	0.66	0.66	0.67	0.67	0.68	0.69	0.72	0.72	0.72
0.72	0.72	0.77	0.77	0.79	0.79	0.82	0.83	0.83	0.83	0.84
0.85	0.85	0.88	0.88	0.90	0.90	0.91	0.93	0.93	0.93	0.94
0.95	0.95	0.97	0.97	0.97	0.99	1.00	1.03	1.03	1.05	1.06
1.06	1.07	1.10	1.11	1.12	1.15	1.15	1.15	1.18	1.18	1.18
1.19	1.20	1.21	1.23	1.23	1.30	1.30	1.31	1.31	1.32	1.32
1.33	1.4	1.41	1.42	1.43	1.43	1.45	1.48	1.49	1.50	1.50
1.51	1.52	1.53	1.54	1.56	1.58	1.59	1.59	1.60	1.65	1.69
1.71	1.72	1.74	1.74	1.74	1.74	1.79	1.81	1.84	1.85	1.86
1.88	1.88	1.94	1.98	1.98	1.99	2.01	2.03	2.06	2.12	2.16
2.17	2.17	2.18	2.18	2.18	2.21	2.22	2.23	2.24	2.27	2.29
2.30	2.32	2.33	2.36	2.38	2.43	2.46	2.47	2.59	2.59	2.60
2.61	2.62	2.63	2.66	2.74	2.75	2.78	2.79	2.79	2.80	2.84
2.85	2.89	2.90								

Табела 3.10: *The Bilbao waves data*: параметар облика ($\hat{\gamma}$), размере ($\hat{\sigma}$) оцене квантила, оцењени десни крај носача расподеле ($-\hat{\sigma}/\hat{\gamma}$), број чланова узорка већих од оцењене горње границе (d), грешке (ASAE, SSQ)

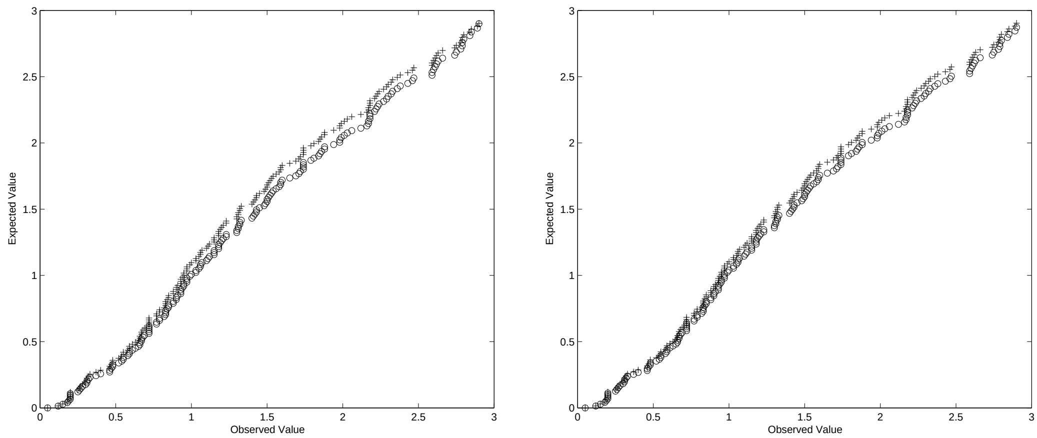
	ММ (ориг.)	ММ (хиб.)	ММ (попр.)	МПМ (ориг.)	МПМ (хиб.)	МПМ (попр.)	*EPM*	ММВ
$\hat{\gamma}$	-1.052	-0.949	-0.91	-1.075	-0.959	-0.918	-0.815	-0.861
$\hat{\sigma}$	2.748	2.748	2.667	2.78	2.78	2.688	2.399	2.501
99%	2.593	2.863	2.885	2.567	2.865	2.886	2.874	2.851
$99,9\%$	2.611	2.896	2.924	2.584	2.896	2.924	2.932	2.899
$-\hat{\sigma}/\hat{\gamma}$	2.613	2.900	2.929	2.585	2.900	2.929	2.943	2.906
d	13	0	0	17	0	0	0	0
ASAE	0.036	0.043	0.039	0.037	0.045	0.041	0.026	0.03
SSQ	0.543	0.692	0.627	0.558	0.725	0.649	0.446	0.476



Слика 3.1: qq-дијаграм за узорак *Bilbao waves data*: (а) поређење оригиналне (о) и кориговане ММ оцене (+); (б) поређење оригиналне (о) и кориговане МПМ оцене (+)



Слика 3.2: qq-дијаграм за узорак *Bilbao waves data*: (а) поређење *EPM* (o) и кориговане ММ оцене (+); (б) поређење *EPM* (o) и кориговане МПМ оцене (+)



Слика 3.3: qq-дијаграм за узорак *Bilbao waves data*: (а) поређење ММВ (o) и кориговане ММ оцене (+); (б) поређење ММВ (o) и кориговане МПМ оцене (+)

Fish river data

У раду [1] аутори су дали пример, скуп података који потиче из хидрологије, на коме су и ММ и МПМ оцена показале несагласност. Узорак се састоји од 42 опсервације (дневни проток реке Fish, Canada, регистрован у периоду 1981-1999.), које су мање од унапред дефинисаног прага.
Подаци поређани по величини, у данима, дати су у Табели 3.11.

Резултати приказани у Табели 3.12 показују одличну сагласност између обе поправљене оцене и ММВ оцене, која се сматра најбољом у случају овог узорка (наведено у [1]).

Табела 3.11: *The Fish river data*: прекорачења поређана од најмањег до највећег

7	7	9	9	11	12	15
17	18	20	20	22	22	24
28	29	30	31	31	32	34
34	35	41	41	47	49	53
57	59	59	60	62	68	72
74	76	78	79	92	101	111

Табела 3.12: *The Fish river data*: параметар облика ($\hat{\gamma}$), размере ($\hat{\sigma}$) оцене квантила, оцењени десни крај носача расподеле ($-\hat{\sigma}/\hat{\gamma}$), број чланова узорка већих од оцењене горње границе (d), грешке (ASAE, SSQ)

	ММ (ориг.)	ММ (хиб.)	ММ (попр.)	МПМ (ориг.)	МПМ (хиб.)	МПМ (попр.)	*EPM*	ММВ
$\hat{\gamma}$	-0.705	-0.649	-0.643	-0.754	-0.668	-0.661	-0.486	-0.560
$\hat{\sigma}$	72.092	72.092	72.063	74.170	74.170	74.125	64.309	65.451
99%	98.294	105.423	106.301	95.313	105.884	106.773	118.146	108.041
$99,9\%$	101.490	109.750	110.788	97.828	109.902	110.946	127.631	114.472
$-\hat{\sigma}/\hat{\gamma}$	102.275	111.000	112.110	98.366	111.000	112.110	132.226	116.918
d	1	0	0	3	0	0	0	0
ASAE	0.030	0.033	0.034	0.031	0.038	0.038	0.030	0.027
SSQ	0.275	0.322	0.327	0.287	0.361	0.367	0.284	0.246

3.4 Оцењивање квантила генералисаних Паретових расподела и примена у финансијама

3.4.1 Робусност оцена високих квантила

Познато је да ММ, МПМ и *EPM* оцена, генерално, нису робусне. Међутим, кад је у питању конкретан проблем оцене високих квантила расподеле (на пример, 95% квантила и 99% квантила који се често користе у финансијама), можда је могуће међу овим оценама пронаћи једну или комбинацију више оцена које дају задовољавајуће резултате.

Да би се проверила робусност наведених оцена 95% квантила и 99% квантила GP(γ, σ) расподеле, по 1000 симулираних узорака (величине узорка $n \in \{15, 30, 45\}$) је генерисано из сваке од следећих расподела:

1. из расподеле $W_{\gamma,\sigma}$;
2. из мешавине расподела $0.9W_{\gamma,\sigma} + 0.1W_{2\gamma,\sigma}$;
3. из мешавине расподела $0.9W_{\gamma,\sigma} + 0.1W_{\gamma,2\sigma}$,

где $\gamma \in \{-1, -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8, 1\}$, а параметар σ увек има вредност 1. Тачне вредности квантила који се оцењују дате су у Табели 3.13, а резултати (пристрасност и средњеквадратна грешка оцене, подељени са тачним вредностима квантила) у Табелама 3.14 - 3.22.

Резултати симулираних експеримената

Резултати експеримената са контаминираним и неконтаминираним расподелама могу се сумирати на следећи начин:

1. Резултати добијени за неконтаминиране расподеле (Табеле 3.14, 3.17, 3.20) сагласни су са резултатима добијеним у радовима [1], [5], [20] и [29]. Пристрасност и средњеквадратна грешка оцена оба квантила опадају са величином узорка. Има неколико изузетака за случајеве $\gamma < 0$, али то може бити због

начина на који су симулирани случајни узорци из $\mathrm{GP}(\gamma, \sigma)$ расподеле.

2. За све величине узорка и $\gamma \in \{0.4, 0.6, 0.8, 1\}$, све три оцене много боље реагују на контаминацију другог типа (пристрасност и средњеквадратна грешка се не мењају значајно), него на контаминацију првог типа. У овим случајевима може се препоручити МПМ оцена.

3. За све величине узорка и $\gamma \in \{-0.4, -0.6, -0.8, -1\}$, све три оцене боље реагују на контаминацију првог типа, него на контаминацију другог типа. У овим случајевима *EPM* оцена се може препоручити ако постоји извесна сигурност око оцене параметра σ, тј. не постоји контаминација другог типа. У осталим случајевима, ММ и МПМ оцене су бољи избор.

4. За све величине узорка и $\gamma \in \{-0.2, 0, 0.2\}$, све три оцене добро реагују на контаминацију оба типа. У овим случајевима могу се препоручити ММ и МПМ оцена, зато што су им и пристрасност и средњеквадратна оцена мање него код *EPM* оцене.

Табела 3.13: Стварне вредности 95% и 99%-квантила

γ	99%-квантил	99.9%-квантил
1	19.00	99.00
0.8	12.48	48.51
0.6	8.39	24.75
0.4	5.79	13.27
0.2	4.10	7.56
0	3.00	4.61
-0.2	2.25	3.01
-0.4	1.75	2.10
-0.6	1.39	1.56
-0.8	1.14	1.22
-1	0.95	0.99

Табела 3.14: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 15$, без контаминације

	γ	ММ	МПМ	*EPM*
95% кв.				
	-1	-0.01 (0.12)	0.01 (0.12)	0 (0.07)
	-0.8	-0.01 (0.13)	0 (0.14)	0.01 (0.1)
	-0.6	-0.02 (0.15)	-0.01 (0.16)	0.03 (0.14)
	-0.4	-0.03 (0.18)	-0.02 (0.19)	0.07 (0.21)
	-0.2	-0.04 (0.22)	-0.03 (0.23)	0.13 (0.34)
	0	-0.06 (0.28)	-0.04 (0.29)	0.25 (0.59)
	0.2	-0.08 (0.36)	-0.06 (0.36)	0.47 (1.17)
	0.4	-0.08 (0.54)	-0.09 (0.44)	0.93 (2.91)
	0.6	-0.01 (1.21)	-0.11 (0.57)	2.04 (10.03)
	0.8	0.25 (4.04)	-0.11 (1.12)	5.28 (43.89)
	1	1.12 (16)	0.03 (3.9)	16.78 (216.29)
99% кв.				
	-1	0.03 (0.18)	0.05 (0.2)	0.03 (0.13)
	-0.8	0.02 (0.19)	0.05 (0.22)	0.05 (0.19)
	-0.6	0.01 (0.21)	0.04 (0.26)	0.1 (0.29)
	-0.4	-0.01 (0.24)	0.03 (0.3)	0.19 (0.49)
	-0.2	-0.04 (0.29)	0.01 (0.36)	0.37 (0.94)
	0	-0.09 (0.36)	-0.01 (0.45)	0.82 (2.18)
	0.2	-0.16 (0.46)	-0.05 (0.56)	2.1 (7.68)
	0.4	-0.25 (0.62)	-0.11 (0.68)	7.56 (52.45)
	0.6	-0.31 (1.07)	-0.2 (0.83)	45.76 (592.01)
	0.8	-0.28 (2.6)	-0.27 (1.4)	448.66 (8323.17)
	1	-0.03 (7.62)	-0.25 (3.64)	5608.99 (127343.78)

Табела 3.15: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 15$, контаминација са $W_{2\gamma,1}$, *: у овим случајевима су пристрасност и средњеквадратна грешка већи од 10^6

	γ	ММ	МПМ	*EPM*
95% кв.				
	-1	-0.03 (0.13)	-0.01 (0.13)	0 (0.08)
	-0.8	-0.04 (0.14)	-0.03 (0.14)	0.01 (0.1)
	-0.6	-0.05 (0.16)	-0.04 (0.16)	0.02 (0.15)
	-0.4	-0.05 (0.18)	-0.05 (0.19)	0.05 (0.22)
	-0.2	-0.06 (0.22)	-0.05 (0.23)	0.11 (0.34)
	0	-0.06 (0.28)	-0.04 (0.29)	0.25 (0.59)
	0.2	-0.04 (0.41)	-0.03 (0.39)	0.6 (1.71)
	0.4	0.1 (1.99)	0.01 (0.71)	2.09 (17.42)
	0.6	1.35 (27.59)	0.28 (6.7)	17.16 (344.15)
	0.8	17.62 (456.29)	4.15 (109.76)	293.14 (7931.7)
	1	275.26 (7961.39)	66.05 (1914.74)	6533.39 (194002.23)
99% кв.				
	-1	0.01 (0.18)	0.03 (0.2)	0.04 (0.14)
	-0.8	0 (0.19)	0.02 (0.22)	0.06 (0.2)
	-0.6	-0.02 (0.21)	0.01 (0.25)	0.09 (0.3)
	-0.4	-0.03 (0.24)	-0.01 (0.29)	0.17 (0.49)
	-0.2	-0.06 (0.29)	-0.02 (0.35)	0.34 (0.91)
	0	-0.09 (0.36)	-0.01 (0.45)	0.82 (2.18)
	0.2	-0.12 (0.53)	0.01 (0.66)	3.18 (16.54)
	0.4	-0.07 (2.17)	0.05 (1.38)	72.68 (1532.36)
	0.6	0.83 (23.25)	0.43 (11.01)	8711.38 (254249.93)
	0.8	10.83 (291.91)	5.02 (136.8)	*
	1	130.79 (3799.35)	61.11 (1779.95)	*

Табела 3.16: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 15$, контаминација са $W_{\gamma,2}$

	γ	ММ	МПМ	*EPM*
95% кв.				
	-1	0.2 (0.33)	0.19 (0.31)	0.39 (0.59)
	-0.8	0.18 (0.32)	0.17 (0.3)	0.36 (0.57)
	-0.6	0.15 (0.3)	0.15 (0.29)	0.34 (0.56)
	-0.4	0.12 (0.29)	0.12 (0.29)	0.33 (0.58)
	-0.2	0.08 (0.3)	0.1 (0.31)	0.36 (0.66)
	0	0.05 (0.33)	0.07 (0.34)	0.46 (0.89)
	0.2	0.02 (0.4)	0.04 (0.4)	0.69 (1.51)
	0.4	0.02 (0.6)	0 (0.47)	1.18 (3.44)
	0.6	0.09 (1.35)	-0.03 (0.61)	2.4 (11.09)
	0.8	0.37 (4.33)	-0.03 (1.2)	5.95 (46.3)
	1	1.31 (16.61)	0.12 (4.05)	18.37 (222.04)
99% кв.				
	-1	0.33 (0.53)	0.31 (0.5)	0.7 (1.14)
	-0.8	0.3 (0.51)	0.29 (0.51)	0.7 (1.21)
	-0.6	0.26 (0.49)	0.27 (0.52)	0.71 (1.33)
	-0.4	0.21 (0.47)	0.24 (0.52)	0.77 (1.58)
	-0.2	0.14 (0.45)	0.2 (0.54)	0.95 (2.13)
	0	0.05 (0.45)	0.15 (0.58)	1.44 (3.69)
	0.2	-0.06 (0.5)	0.08 (0.66)	2.93 (10.26)
	0.4	-0.17 (0.67)	-0.01 (0.75)	9.18 (58.58)
	0.6	-0.24 (1.16)	-0.12 (0.89)	51.42 (611.72)
	0.8	-0.21 (2.77)	-0.21 (1.49)	478.5 (8395.7)
	1	0.06 (7.9)	-0.19 (3.77)	5809.77 (127620.56)

Табела 3.17: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 30$, без контаминације

	γ	ММ	МПМ	*ЕРМ*
95% кв.				
	-1	0 (0.08)	0.01 (0.08)	0 (0.03)
	-0.8	-0.01 (0.09)	0 (0.09)	0 (0.05)
	-0.6	-0.01 (0.1)	0 (0.11)	0.01 (0.07)
	-0.4	-0.01 (0.12)	-0.01 (0.13)	0.03 (0.12)
	-0.2	-0.02 (0.15)	-0.02 (0.16)	0.06 (0.19)
	0	-0.03 (0.19)	-0.02 (0.2)	0.12 (0.31)
	0.2	-0.05 (0.25)	-0.03 (0.25)	0.22 (0.52)
	0.4	-0.06 (0.35)	-0.06 (0.31)	0.39 (0.96)
	0.6	-0.03 (0.68)	-0.1 (0.37)	0.68 (1.96)
	0.8	0.13 (2.05)	-0.15 (0.5)	1.22 (4.72)
	1	0.63 (7.86)	-0.16 (1.12)	2.28 (13.22)
99% кв.				
	-1	0.01 (0.12)	0.03 (0.13)	0.01 (0.04)
	-0.8	0.01 (0.13)	0.02 (0.15)	0.01 (0.07)
	-0.6	0.01 (0.14)	0.02 (0.17)	0.03 (0.11)
	-0.4	0 (0.16)	0.01 (0.2)	0.06 (0.19)
	-0.2	-0.02 (0.2)	0 (0.24)	0.14 (0.35)
	0	-0.06 (0.26)	-0.01 (0.31)	0.3 (0.7)
	0.2	-0.12 (0.34)	-0.03 (0.4)	0.64 (1.63)
	0.4	-0.22 (0.44)	-0.08 (0.49)	1.47 (5.19)
	0.6	-0.31 (0.67)	-0.18 (0.55)	3.87 (25.06)
	0.8	-0.34 (1.38)	-0.3 (0.66)	12.94 (156.64)
	1	-0.25 (3.82)	-0.41 (1.15)	57.45 (1080.44)

Табела 3.18: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 30$, контаминација са $W_{2\gamma,1}$

	γ	ММ	МПМ	*ЕРМ*
95% кв.				
	-1	-0.02 (0.09)	-0.02 (0.09)	-0.01 (0.04)
	-0.8	-0.03 (0.1)	-0.03 (0.1)	-0.01 (0.05)
	-0.6	-0.04 (0.11)	-0.03 (0.11)	0 (0.08)
	-0.4	-0.04 (0.12)	-0.04 (0.13)	0.01 (0.12)
	-0.2	-0.04 (0.15)	-0.04 (0.16)	0.04 (0.18)
	0	-0.03 (0.19)	-0.02 (0.2)	0.12 (0.31)
	0.2	-0.01 (0.28)	0 (0.28)	0.31 (0.71)
	0.4	0.09 (0.89)	0.03 (0.41)	0.8 (2.62)
	0.6	0.75 (8.16)	0.1 (1.16)	2.46 (15.11)
	0.8	6.43 (101.83)	0.75 (12.67)	10.21 (111.48)
	1	67.15 (1399.48)	8.22 (173.51)	56.98 (913.64)
99% кв.				
	-1	0 (0.12)	0.01 (0.13)	0.01 (0.05)
	-0.8	-0.01 (0.13)	0 (0.15)	0.01 (0.07)
	-0.6	-0.02 (0.15)	-0.01 (0.17)	0.03 (0.11)
	-0.4	-0.03 (0.17)	-0.02 (0.2)	0.06 (0.19)
	-0.2	-0.04 (0.2)	-0.02 (0.24)	0.12 (0.34)
	0	-0.06 (0.26)	-0.01 (0.31)	0.3 (0.7)
	0.2	-0.07 (0.39)	0.03 (0.48)	0.98 (3.16)
	0.4	-0.05 (1)	0.09 (0.77)	5.95 (55.92)
	0.6	0.37 (7.01)	0.15 (1.99)	92.63 (1731.35)
	0.8	3.77 (66.48)	0.86 (16.4)	2529.05 (62661.91)
	1	32.16 (681.79)	7.68 (167.42)	86679.07 (2435607.8)

Табела 3.19: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 30$, контаминација са $W_{\gamma,2}$

	γ	ММ	МПМ	*ЕРМ*
95% кв.				
	-1	0.23 (0.3)	0.2 (0.27)	0.45 (0.55)
	-0.8	0.21 (0.29)	0.18 (0.26)	0.4 (0.51)
	-0.6	0.18 (0.27)	0.16 (0.25)	0.35 (0.47)
	-0.4	0.15 (0.25)	0.14 (0.24)	0.3 (0.44)
	-0.2	0.12 (0.24)	0.12 (0.25)	0.28 (0.45)
	0	0.09 (0.25)	0.1 (0.27)	0.31 (0.52)
	0.2	0.05 (0.29)	0.07 (0.3)	0.39 (0.72)
	0.4	0.03 (0.4)	0.04 (0.35)	0.56 (1.16)
	0.6	0.07 (0.76)	-0.02 (0.4)	0.87 (2.22)
	0.8	0.24 (2.16)	-0.07 (0.52)	1.45 (5.07)
	1	0.78 (8.03)	-0.09 (1.16)	2.59 (13.74)
99% кв.				
	-1	0.33 (0.45)	0.27 (0.39)	0.71 (0.88)
	-0.8	0.31 (0.44)	0.26 (0.4)	0.66 (0.86)
	-0.6	0.28 (0.43)	0.25 (0.4)	0.61 (0.85)
	-0.4	0.24 (0.4)	0.23 (0.41)	0.56 (0.87)
	-0.2	0.18 (0.38)	0.2 (0.42)	0.56 (0.97)
	0	0.1 (0.36)	0.16 (0.44)	0.68 (1.32)
	0.2	0 (0.38)	0.11 (0.5)	1.03 (2.39)
	0.4	-0.13 (0.46)	0.03 (0.56)	1.94 (6.32)
	0.6	-0.23 (0.7)	-0.09 (0.59)	4.59 (26.59)
	0.8	-0.27 (1.43)	-0.24 (0.69)	14.3 (156.51)
	1	-0.18 (3.9)	-0.35 (1.17)	60.3 (1061.51)

Табела 3.20: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 45$, без контаминације

	γ	ММ	МПМ	*EPM*
95% кв.				
	-1	0 (0.07)	0 (0.07)	0 (0.03)
	-0.8	0 (0.07)	0 (0.08)	0 (0.04)
	-0.6	-0.01 (0.08)	0 (0.09)	0.01 (0.06)
	-0.4	-0.01 (0.1)	-0.01 (0.11)	0.02 (0.09)
	-0.2	-0.01 (0.12)	-0.01 (0.13)	0.05 (0.15)
	0	-0.02 (0.16)	-0.01 (0.16)	0.1 (0.25)
	0.2	-0.04 (0.21)	-0.02 (0.21)	0.18 (0.43)
	0.4	-0.04 (0.3)	-0.04 (0.26)	0.32 (0.77)
	0.6	0 (0.6)	-0.08 (0.31)	0.57 (1.46)
	0.8	0.21 (1.75)	-0.13 (0.38)	0.99 (2.93)
	1	0.87 (6.32)	-0.16 (0.67)	1.73 (6.36)
99% кв.				
	-1	0.01 (0.1)	0.02 (0.11)	0 (0.02)
	-0.8	0.01 (0.11)	0.02 (0.12)	0.01 (0.04)
	-0.6	0.01 (0.12)	0.02 (0.14)	0.02 (0.07)
	-0.4	0 (0.14)	0.01 (0.17)	0.05 (0.13)
	-0.2	-0.01 (0.17)	0.01 (0.21)	0.1 (0.25)
	0	-0.03 (0.22)	0 (0.26)	0.23 (0.52)
	0.2	-0.09 (0.3)	-0.01 (0.35)	0.49 (1.17)
	0.4	-0.18 (0.39)	-0.04 (0.44)	1.09 (3.1)
	0.6	-0.26 (0.6)	-0.13 (0.5)	2.59 (10.26)
	0.8	-0.27 (1.18)	-0.27 (0.55)	6.84 (40.86)
	1	-0.12 (3.08)	-0.39 (0.76)	20.77 (185.22)

Табела 3.21: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 45$, контаминација са $W_{2\gamma,1}$

	γ	ММ	МПМ	*ЕРМ*
95% кв.				
	-1	-0.02 (0.08)	-0.02 (0.08)	-0.01 (0.03)
	-0.8	-0.03 (0.08)	-0.03 (0.09)	-0.01 (0.04)
	-0.6	-0.03 (0.09)	-0.03 (0.1)	-0.01 (0.06)
	-0.4	-0.04 (0.1)	-0.04 (0.11)	0.01 (0.09)
	-0.2	-0.03 (0.12)	-0.03 (0.13)	0.03 (0.15)
	0	-0.02 (0.16)	-0.01 (0.16)	0.1 (0.25)
	0.2	0.01 (0.25)	0.02 (0.24)	0.28 (0.63)
	0.4	0.17 (1.28)	0.07 (0.38)	0.76 (2.61)
	0.6	1.64 (16.86)	0.19 (1.62)	2.45 (15.51)
	0.8	20.55 (270.87)	1.78 (23.17)	10.7 (106.74)
	1	310.34 (4705.64)	26.45 (395.68)	61.21 (784.19)
99% кв.				
	-1	-0.01 (0.1)	0 (0.11)	0 (0.03)
	-0.8	-0.01 (0.11)	-0.01 (0.12)	0.01 (0.05)
	-0.6	-0.02 (0.13)	-0.02 (0.15)	0.01 (0.08)
	-0.4	-0.03 (0.14)	-0.02 (0.17)	0.03 (0.14)
	-0.2	-0.04 (0.17)	-0.02 (0.21)	0.08 (0.25)
	0	-0.03 (0.22)	0 (0.26)	0.23 (0.52)
	0.2	-0.02 (0.36)	0.07 (0.45)	0.81 (2.39)
	0.4	0.07 (1.44)	0.16 (0.81)	4.92 (36.74)
	0.6	1.16 (14.59)	0.33 (2.83)	70.33 (862.93)
	0.8	13.1 (178.08)	2.25 (30.32)	1591.43 (23534.86)
	1	151.66 (2308.3)	25.59 (386.32)	41657.52 (685463.25)

Табела 3.22: Пристрасност (средњеквадратна грешка) оцена 95% и 99%-квантила, $n = 45$, контаминација са $W_{\gamma,2}$

	γ	ММ	МПМ	*ЕРМ*
95% кв.				
	-1	0.22 (0.27)	0.18 (0.23)	0.48 (0.55)
	-0.8	0.21 (0.26)	0.17 (0.22)	0.42 (0.49)
	-0.6	0.18 (0.24)	0.15 (0.21)	0.36 (0.44)
	-0.4	0.16 (0.22)	0.14 (0.2)	0.3 (0.4)
	-0.2	0.13 (0.21)	0.12 (0.2)	0.28 (0.4)
	0	0.11 (0.21)	0.11 (0.22)	0.3 (0.46)
	0.2	0.08 (0.25)	0.1 (0.26)	0.4 (0.7)
	0.4	0.09 (0.42)	0.08 (0.31)	0.62 (1.28)
	0.6	0.23 (1.15)	0.03 (0.36)	1.06 (2.66)
	0.8	0.8 (4.13)	-0.01 (0.53)	1.94 (6.04)
	1	2.91 (16.63)	0.07 (1.49)	3.81 (14.9)
99% кв.				
	-1	0.31 (0.4)	0.24 (0.31)	0.73 (0.85)
	-0.8	0.3 (0.39)	0.24 (0.32)	0.68 (0.81)
	-0.6	0.29 (0.39)	0.23 (0.32)	0.61 (0.78)
	-0.4	0.26 (0.37)	0.22 (0.33)	0.55 (0.76)
	-0.2	0.21 (0.35)	0.2 (0.34)	0.54 (0.81)
	0	0.16 (0.33)	0.18 (0.37)	0.63 (1.08)
	0.2	0.07 (0.36)	0.17 (0.47)	1.01 (2.12)
	0.4	-0.03 (0.49)	0.13 (0.59)	2.17 (6.12)
	0.6	-0.06 (1)	0.03 (0.63)	5.88 (23.36)
	0.8	0.12 (2.68)	-0.11 (0.75)	19.81 (105.56)
	1	0.88 (8.08)	-0.17 (1.49)	77.98 (511.07)

Оцењивање квантила непознате расподеле

Нека је $(x_1, x_2, \ldots, x_N)$ узорак чији чланови представљају реализације случајне величине X са непознатом функцијом расподеле F. Нека је познато да функција F припада области привлачења за прекорачења високог нивоа једне од генералисаних Паретових расподела, чији параметри облика и размере су већ оцењени са $\hat{\gamma}$ и $\hat{\sigma}$. Из једнакости (1.15) следи

$$F^{(u)}(x) = 1 - \frac{1 - F(x+u)}{1 - F(u)}, \quad x \geq 0, \tag{3.28}$$

па се десни реп расподеле F може приказати на следећи начин:

$$1 - F(x+u) = \left(1 - F^{(u)}(x)\right)(1 - F(u)), \quad x \geq 0. \tag{3.29}$$

где је u задати високи ниво.

Нека је n број чланова узорка већих од u. Тада се први члан на десној страни једнакости (3.29) може оценити са $1 - W_{\hat{\gamma},\hat{\sigma}}(x)$, други члан на десној страни са $\frac{n}{N}$, па се заменом у једнакост (3.29) добија оцена репа непознате расподеле F:

$$1 - F(x+u)\hat{} = \begin{cases} \frac{n}{N}\left(1 + \hat{\gamma}\frac{x}{\hat{\sigma}}\right)^{-\frac{1}{\hat{\gamma}}}, & \hat{\gamma} \neq 0; \\ \frac{n}{N} e^{-\frac{x}{\hat{\sigma}}}, & \hat{\gamma} = 0. \end{cases} \tag{3.30}$$

Ако је x_p p-квантил расподеле F, онда важи:

$$1 - p = \begin{cases} \frac{n}{N}\left(1 + \hat{\gamma}\frac{x_p - u}{\hat{\sigma}}\right)^{-\frac{1}{\hat{\gamma}}}, & \hat{\gamma} \neq 0; \\ \frac{n}{N} e^{-\frac{x_p - u}{\hat{\sigma}}}, & \hat{\gamma} = 0. \end{cases}$$

и, сређивањем једнакости, добија се оцена квантила x_p:

$$\hat{x_p} = \begin{cases} u + \frac{\hat{\sigma}}{\hat{\gamma}}\left(\frac{N}{n}(1-p)^{-\hat{\gamma}} - 1\right), & \hat{\gamma} \neq 0; \\ u - \hat{\sigma}\log\left(\frac{N}{n}(1-p)\right), & \hat{\gamma} = 0. \end{cases} \tag{3.31}$$

3.4.2 Оцењивање параметра вредности при ризику

Банкама, осигуравајућим друштвима и, уопште, свим учесницима на финансијским тржиштима је потребно да знају приближан степен изложености ризику при својим акцијама. Један од начина да се то утврди је оцењивање параметра **вредности при ризику** (VaR). Параметар VaR представља максимални губитак коме се излаже учесник на финансијском тржишту у неком одређеном временском периоду, за дату вероватноћу. На основу те вредности, учесник на финансијском тржишту може да одреди ниво капитала који му обезбеђује позицију у условима екстремних кретања.

Нека су I_0,I_1,$\ldots$,I_t вредности појединачног финансијског инструмента (на пример, цене акција неке компаније или девизни курс), дате у дискретним временским тренуцима $t \in \{0, 1, 2 \ldots\}$. Период је обично дан, недеља или месец. Тада је принос у периоду T (промена вредности тог финансијског инструмента у односу на почетну, у периоду T) једнак $(I_T - I_0)/I_0$.

Принос у периоду $T = 1$ (дневни, недељни, месечни) је онда

$$s_t = \frac{I_t - I_{t-1}}{I_{t-1}}$$

и назива се **аритметички принос**.

Иста величина се може изразити у облику

$$r_t = \log I_t - \log I_{t-1}$$

што се назива **логаритамски принос**.

Ове две величине су блиске у случају кад је I_t/I_{t-1} блиско вредности 1. Величина r_t је погоднија за статистичко моделовање због тога што је неограничена и због тога што неке економске законитости могу помоћу ње једноставније да се прикажу.

Ако је I_0 вредност финансијског инструмента у тренутку $t = 0$, и $r_1, r_2, \ldots, r_T$ приноси током периода T, онда је вредност I_T у тренутку T једнака

$$I_T = I_0 \exp \sum_{t \leq T} r_t. \tag{3.32}$$

Губитак коме је учесник на тржишту изложен током периода од T дана (недеља, месеци) је

$$G_T = \sum_{t \leq T} (-r_t).$$

Нека је β задата вероватноћа (обично 95% или 99%). Тада је $VaR(T;\beta)$, тј. максимални губитак за временски хоризонт T и дати ниво поверења β, одређен са

$$F_T = P\{G_T \leq VaR(T;\beta)\} = \beta$$

Значи, $VaR(T;\beta)$ је β-квантил функције расподеле губитка, F_T. Из једнакости (3.32) се добија

$$\begin{aligned} VaR(T;\beta)\hat{} &= -(I_T - I_0) = I_0\left(1 - \exp\sum_{t\leq T} r_t\right) \\ &= I_0\left(1 - \exp\{-F_T^{-1}(\beta)\}\right) \approx I_0 F_T^{-1}(\beta). \end{aligned}$$

Ова оцена параметра $VaR(T;\beta)$ се може користити у случају кад је $F_T^{-1}(\beta)$ довољно мало. Даље, треба наћи оцену величине $F_T^{-1}(\beta)$, на основу познатих вредности I_0,I_1,...,I_T. Неки од начина да се то уради су:

- *Коришћење емпиријске функције расподеле губитака*
  На основу претходних вредности финансијског инструмента, односно приноса и губитака, одреди се емпиријска функција расподеле губитака, а њен β - квантил је оцена за $F_T^{-1}(\beta)$. Мане овог метода су што није могуће оценити екстремне квантиле на основу узорка и што је ова оцена осетљива на појаву екстремних података у узорку.

- *Моделовање генералисаним Паретовим расподелама*
  Функција расподеле губитака се моделује једном од генералисаних Паретових расподела, $W_{\gamma,\mu,\sigma}$. Тада се за оцену β - квантила расподеле $W_{\gamma,\mu,\sigma}$ може користити формула (3.31) и добија се

$$VaR(T;\beta)\hat{} = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\gamma}}\left(\frac{N}{n}(1-\beta)^{-\hat{\gamma}} - 1\right). \tag{3.33}$$

Викенд ефекат у финансијама

Викенд ефекат је познати феномен везан за финансијска тржишта, који се односи се на разлику у приносима између понедељка и осталих дана у недељи. Прецизније, цене финансијских инструмената регистроване понедељком су значајно ниже него оне регистроване претходног петка, па могу бити непоуздане. Овој појави, која је запажана у дугим временским периодима и на великом броју међународних тржишта, посвећено је доста радова, а неки од најранијих су [15] и [27]. Постоји више техника за кориговање викенд ефекта, а неке од њих су (према [36]):

1. *рачинање приноса током радних дана:* у обзир се узимају само приноси остварени током радних дана;

2. *прескакање приноса од понедељка:* прескачу се дани кад цене нису регистроване, укључујући и наредни понедељак;

3. *расподела приноса од понедељка:* приноси од понедељка су једнако расподељени на суботу, недељу и понедељак.

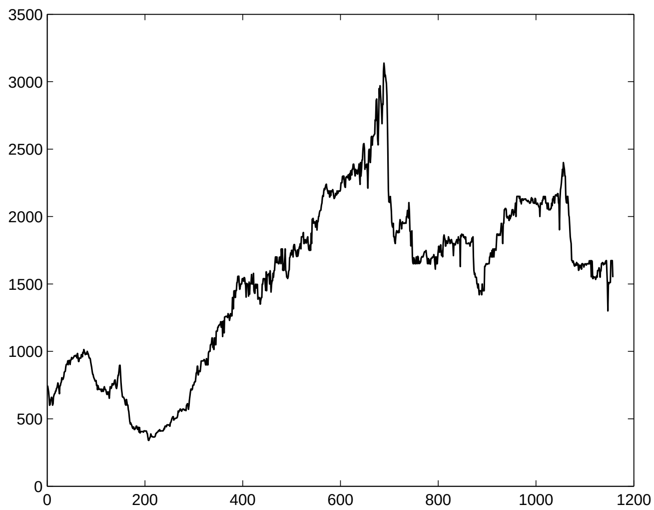
Један од начина да се умањи овај проблем је и робусан модел.

3.4.3 Пример - оцена VaR параметра

Као пример за претходно наведене технике, анализирани су дневни приноси цена акција компаније Тигар а. д. Пирот у периоду од 31. маја 2005. до 31. децембра 2009. Подаци се састоје од укупно 1157 дневних нивоа цена, регистрованих у току радних дана и преузети су са адресе www.belex.rs. Циљ је био оценити параметре $Var(1; 0.95)$ и $Var(1; 0.99)$.

Дневне стопе приноса су добијене као разлика логаритмованих нивоа цена у два узастопна дана. Губици су добијени множењем дневних стопа приноса са -1. Почетни узорак који се састоји од дневних нивоа цена показује линеарни тренд (Слика 3.4), док је серија дневних приноса стационарна (Слика 3.5). Вредност високог нивоа, потребног ради апроксимације генералисаним Паретовим расподелама, добијена је анализом графичког приказа функције узорачког средњег прекорачења. Ова функција је приближно линеарна почевши од вредности $u = 0.04$ (Слика 3.6), па је ова вредност одређена као високи праг. Ради поређења, у даљем раду анализиране су и вредности прага $u \in \{0.04, 0.05, 0.06, 0.07, 0.08, 0.09\}$.

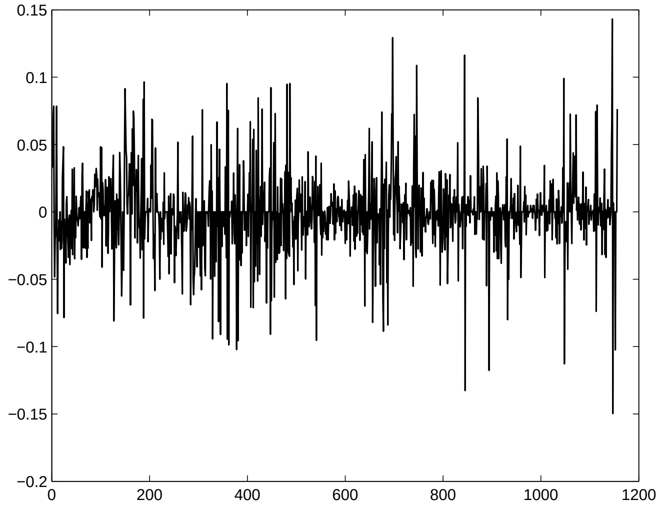
Оцене парамара $Var(1; 0.95)$ и $Var(1; 0.99)$ добијене методом рачунања приноса током радних дана, методом прескакања приноса од понедељка и методом расподеле приноса од понедељка дате су у Табелама 3.23 - 3.25.



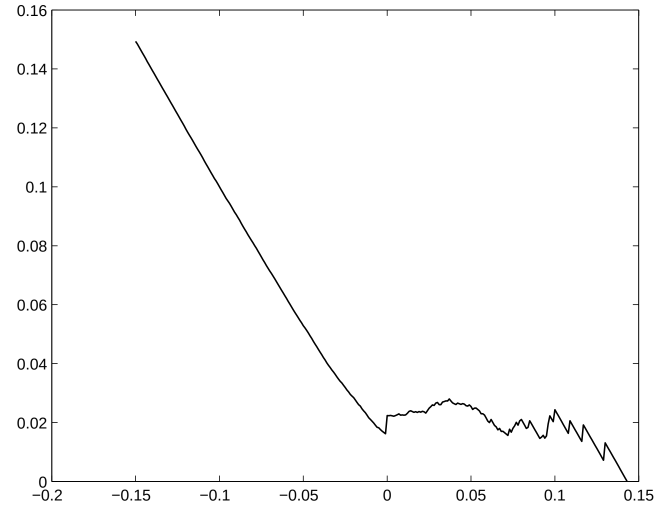
Слика 3.4: Цене акција компаније Тигар а. д. Пирот (оригинални подаци)

Табела 3.23: Оцене добијене методом рачунања приноса током радних дана

u	N_u	метод	γ	σ	$Var(1;0.95)$	$Var(1;0.99)$
0.04	75	*EPM*	-0.0984	0.0292	0.0475	0.0899
		MM	-0.2448	0.0328	0.0483	0.0893
		МПМ	-0.2549	0.0331	0.0484	0.0893
0.05	53	*EPM*	-0.1495	0.0310	0.0473	0.0922
		MM	-0.3615	0.0347	0.0470	0.0906
		МПМ	-0.4831	0.0377	0.0467	0.0907
0.06	43	*EPM*	-0.0020	0.0220	0.0535	0.0889
		MM	-0.1495	0.0236	0.0529	0.0882
		МПМ	-0.2380	0.0254	0.0522	0.0887
0.07	31	*EPM*	0.2331	0.0137	0.0621	0.0852
		MM	0.0642	0.0155	0.0605	0.0858
		МПМ	0.1356	0.0143	0.0615	0.0851
0.08	14	*EPM*	-0.0431	0.0247	0.0440	0.0848
		MM	-0.2158	0.0256	0.0378	0.0849
		МПМ	-0.2868	0.0271	0.0329	0.0852
0.09	11	*EPM*	0.3659	0.0131	0.0737	0.0895
		MM	0.0841	0.0143	0.0679	0.0894
		МПМ	0.2108	0.0123	0.0728	0.0895



Слика 3.5: Дневни приноси цена акција компаније Тигар а. д. Пирот



Слика 3.6: Тигар а. д. Пирот: узорачко средње прекорачење

Табела 3.24: Оцене добијене методом прескакања приноса од понедељка

u	N_u	метод	γ	σ	$Var(1; 0.95)$	$Var(1; 0.99)$
0.04	57	*EPM*	-0.0396	0.0275	0.0457	0.0882
		MM	-0.1922	0.0310	0.0463	0.0876
		МПМ	-0.1915	0.0310	0.0463	0.0876
0.05	40	*EPM*	-0.1164	0.0309	0.0481	0.0916
		MM	-0.2781	0.0322	0.0452	0.0887
		МПМ	-0.3686	0.0345	0.0448	0.0890
0.06	32	*EPM*	0.0824	0.0205	0.0526	0.0868
		MM	-0.0993	0.0227	0.0515	0.0866
		МПМ	-0.1556	0.0239	0.0509	0.0870
0.07	23	*EPM*	0.2346	0.0155	0.0600	0.0857
		MM	0.0536	0.0164	0.0588	0.0853
		МПМ	0.0676	0.0162	0.0615	0.0851
0.08	11	*EPM*	0.1335	0.0214	0.0521	0.0839
		MM	-0.0572	0.0216	0.0479	0.0839
		МПМ	-0.0689	0.0218	0.0473	0.0839
0.09	8	*EPM*	0.6055	0.0116	0.0776	0.0885
		MM	0.1417	0.0141	0.0682	0.0882
		МПМ	0.3222	0.0112	0.0751	0.0886

Табела 3.25: Оцене добијене методом расподеле приноса од понедељка

u	N_u	метод	γ	σ	$Var(1; 0.95)$	$Var(1; 0.99)$
0.04	57	*EPM*	-0.0396	0.0275	0.0304	0.0738
		MM	-0.1922	0.0310	0.0288	0.0747
		МПМ	-0.1915	0.0310	0.0288	0.0747
0.05	40	*EPM*	-0.1164	0.0309	0.0274	0.0766
		MM	-0.2781	0.0322	0.0250	0.0758
		МПМ	-0.3686	0.0345	0.0223	0.0766
0.06	32	*EPM*	0.0824	0.0205	0.0417	0.0744
		MM	-0.0993	0.0227	0.0380	0.0750
		МПМ	-0.1556	0.0239	0.0362	0.0755
0.07	23	*EPM*	0.2346	0.0155	0.0532	0.0757
		MM	0.0536	0.0164	0.0501	0.0759
		МПМ	0.0676	0.0162	0.0506	0.0758
0.08	11	*EPM*	0.1335	0.0214	0.0426	0.0721
		MM	-0.0572	0.0216	0.0347	0.0718
		МПМ	-0.0689	0.0218	0.0336	0.0717
0.09	8	*EPM*	0.6055	0.0116	0.0757	0.0836
		MM	0.1417	0.0141	0.0623	0.0807
		МПМ	0.3222	0.0112	0.0719	0.0831

Табела 3.26: Квалитет апроксимације (за праг $u = 0.04$)

метод	коефицијент детерминације	збир квадрата резидуала
1.		
EPM	0.9820	0.1136
MM	0.9868	0.0856
МПМ	0.9869	0.0852
2.,3.		
EPM	0.9851	0.0722
MM	0.9883	0.0588
МПМ	0.9883	0.0588

Поређење са индексима Београдске берзе

Индекси Београдске берзе уведени су ради бољег информисања потенцијалних инвеститора и потребе да се побољша транспарентност и упоредивост података на тржишту.

Да би се упоредиле оцене ризика добијене за компанију Тигар а. д. Пирот са максималним, односно просечним ризиком везаним за учешће на Београдској берзи, посматрани су следећи индекси:

1. *BELEX15*: водећи индекс Београдске берзе, који описује кретање цена 15 најликвиднијих српских акција и рачуна се у реалном времену;

2. *BELEXline*: општи индекс Београдске берзе, који се рачуна на крају радног дана.

Историјски подаци за оба индекса за период од 31. маја 2005. до 31. децембра 2009. преузети су са адресе www.belex.rs. Резултати оцене ризика дати су у Табели 3.27.

Пошто су оцене параметара $Var(1; 0.95)$ и $Var(1; 0.99)$ за компанију Тигар а. д. Пирот једнаке, редом, 0.048 и 0.09 (односи се на највећи узорак), може се закључити да улагање у акције ове компаније носи са собом два пута већи ризик од улагања у најликвидније акције Београдске берзе, односно, три пута већи ризик од генералног ризика за Београдску берзу.

Табела 3.27: Оцене ризика за индексе Београдске берзе

индекс	$Var(1; 0.95)$	$Var(1; 0.99)$
BELEX15	0.0246	0.0484
BELEXline	0.0176	0.0304

Закључак

Резултати добијени при оцењивању ризика за Тигар а. д. Пирот могу се сумирати на следећи начин:

1. Оцене $Var(1; 0.95)$ и $Var(1; 0.99)$ добијене методом прескакања приноса од понедељка не разликују се много од оних добијених рачунањем приноса током свих радних дана. То значи

да у разматраном примеру цене акција регистроване понедељком нису значајно ниже од цена регистрованих током осталих радних дана, тј. викенд ефекат није имао много утицаја.

2. Оцене параметара $Var(1; 0.95)$ и $Var(1; 0.99)$ добијене методом расподеле приноса од понедељка су значајно ниже од оних добијених методом прескакања приноса, иако су оцене параметара γ и σ исте као у том случају. То је због тога што се мења величина узорка у формули (3.31).

3. Примећено је да се оцене параметра γ значајно мењају са променом високог прага, односно, величине узорка на основу кога се врши оцењивање. То важи за све три методе оцењивања. Међутим, оцене оба параметра ризика, $Var(1; 0.95)$ и $Var(1; 0.99)$, су стабилне у свим случајевима, осим за $u = 0.09$ (што су оцене добијене на веома малом узорку). То је у складу са резултатима симулација датим у одељку 3.4.1, тј. са чињеницом да за γ блиско нули све три методе оцењивања дају задовољавајуће резултате. Такође, квалитет апроксимације добијен на основу највећег узорка (праг $u = 0.04$) је задовољавајући (Табела 3.26). На основу свега тога се може закључити да су оцене VaR параметра добијене на овај начин поуздане.

4. Још један показатељ поузданости оцена VaR параметра, тј. квантила расподеле, је *failure rate*, односно однос броја чланова узорка већих од оцењеног VaR параметра и величине целог узорка. Оцена параметра VaR је адекватна ако је тај однос приближно једнак $1-\beta$, где је β ниво поверења. У овом случају, добијене су вредности: приближно 0.052 за $Var(1; 0.95)$ оцену и приближно 0.011 за $Var(1; 0.99)$ оцену, што је задовољавајућа тачност и још једна потврда адекватности оцена.

Литература

[1] F. Ashkar and C. N. Tatsambon. Revisiting some estimation methods for the generalized Pareto distribution. *Journal of Hydrology*, 346(136-143), 2007.

[2] A. Balkema and L. de Haan. Residual life time at great age. *Annals of Probability*, 2(792-804), 1974.

[3] L. E. Baum and P. Billingsley. Asymptotic distributions for the coupon collector's problem. *Ann. Math. Statist.*, 36(1835-1839), 1965.

[4] S. Boneh and V. G. Papanicolau. General asymptotic estimates for the coupon collector problem. *Journal of Computational and Applied Mathematics*, 67(277-289), 1996.

[5] E. Castillo and A. S. Hadi. Fitting the generalized Pareto distribution to data. *Journal of the American Statistical Association*, 92(440)(1609-1620), 1997.

[6] B. Dawkins. Siobhan's problem: the coupon collector revisited. *The American Statistician*, 45(1)(76-82), 1991.

[7] P. de Zea Bermudez and S. Kotz. Parameter estimation of the generalized Pareto distribution–Parts 1 and 2. *Jornal of Statistical Planning and inference*, 140(1353-1373), 2010.

[8] J. del Castillo and J. Daoudi. Estimation of the generalized Pareto distribution. *Statistics and Probability Letters*, 79(684-688), 2009.

[9] D. J. Dupuis. Estimating the probability of obtaining nonfeasible parameter estimates of the generalized Pareto distribution. *Journal of Statistical Computation and Simulation*, 54(197-209), 1996.

[10] D. J. Dupuis and M. Tsao. A hybrid estimator for generalized Pareto and extreme-value distributions. *Communications in Statistics - Theory and Methods*, 27(4)(925-941), 1998.

[11] P. Embrechts, C. Kluppelberg, and T. Mikosch. *Modelling extremal events*. Springer Verlag, Berlin, 2003.

[12] P. Erdös and A. Rényi. On a classical problem of probability theory. *Magyar Tud. Akad. Mat. Kutato Int. Kozl.*, 9(133-141), 1961.

[13] W. Feller. *An Introduction to Probability Theory and its Applications, vol. 2*. Wiley, 1966.

[14] R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc. Cambridge Phil. Soc.*, 24(180-190), 1928.

[15] K. R. French. Stock returns and the weekend effect. *Journal of Financial Economics*, 8(55-69), 1980.

[16] B. V. Gnedenko. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, 44(423–453), 1943.

[17] J. A. Greenwood, J. M. Landwehr, N. C. Matalas, and R. Wallis. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15(5)(1049-1054), 1979.

[18] L. de Haan. *On regular variation and its application to the weak convergence of sample extremes*. Mathematical Centre Tracts, Amsterdam, 1970.

[19] L. Holst. On birthday, collectors, occupancy and other classical urn problems. *International Statistical Review*, 54(15-27), 1986.

[20] J. R. M. Hosking and J. R. Wallis. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29(3)(339-349), 1987.

[21] J. Jocković. Quantile estimation for the generalized Pareto distributions with application to finance. *Yugoslav Journal of Operations Research*, (DOI:10.2298/YJOR110308013J).

[22] J. Jocković. Correcting certain estimation methods for the generalized Pareto distribution. In A. Migdalas, A. Sifaleras, C. K. Georgiadis, J. Papathanasiou, and E. Stiakakis, editors, *Optimization Theory, Decision Making, and Operations Research Applications (Proceedings of the 1st International Symposium and 10th Balkan Conference on Operational Research)*, volume 31, 2013.

[23] J. Jocković and P. Mladenović. Coupon collector's problem and generalized Pareto distributions. *Journal of Statistical Planning and Inference*, 141(2348-2352), 2011.

[24] N. L. Johnson and S. Kotz. *Urn models and their application, an approach to modern discrete probability theory.* John Wiley and Sons, New York-London-Sydney-Toronto, 1977.

[25] S. Juárez and W. Schucany. Robust and efficient estimation for the generalized Pareto distribution. *Extremes*, 7(237-251), 2004.

[26] J. Karamata. Sur un mode de croissance régulière des fonctions. *Mathematica (Cluj)*, 3(33-48), 1930.

[27] D. Keim and R. Stambaugh. A further investigation of the weekend effect in stock returns. *Journal of Finance*, 39(819-835), 1984.

[28] M. R. Leadbetter, G. Lindgren, and H. Rootzen. *Extremes and related properties of random sequences and processes.* Springer, New York, 1983.

[29] E. Mackay, P. Challenor, and A. Bahaj. A comparison of estimators for the generalised Pareto distribution. *Ocean Engeneering*, 38(1338-1346), 2011.

[30] P. Mladenović. Limit theorems for the maximum terms of a sequence of random variables with marginal geometric distributions. *Extremes*, 2(405-419), 1999.

[31] P. Mladenović. *Ekstremne vrednosti slučajnih nizova.* Matematički fakultet, Beograd, 2002.

[32] P. Mladenović. A generalization of the Mejzler-de Haan theorem. *Theory Probab. Appl.*, 50(141-153), 2002.

[33] P. Mladenović. Limit distributions for the problem of collecting pairs. *Bernoulli*, 14(2)(419-439), 2008.

[34] L. Peng and A. H. Welsh. Robust estimation of the generalized Pareto distribution. *Extremes*, 4(1)(53-65), 2001.

[35] J. Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(119-131), 1975.

[36] R.-D. Reiss and M. Thomas. *Statistical analysis of extreme values*. Birkhauser Verlag, Basel, 1997.

[37] S. M. Ross. The hypergeometric coupon collection problem and its dual. *Journal of Industrial and Systems Engineering*, 1(1)(1-7), 2007.

[38] E. Samuel-Cahn. Asymptotic distributions for occupancy and waiting time problems with positive probability of falling through the cells. *The Annals of Probability*, 2(3)(515-521), 1974.

[39] R. von Mises. La distribution de la plus grande de n valeurs. *Rev. Math. Union Interbalcanique*, 1(141-160), 1936.

Биографија аутора

Јелена (Мирослав) Јоцковић рођена је 1979. године у Београду. Завршила је Математичку гимназију у Београду 1998. године.

На Математичком факултету у Београду, смер Теоријска математика и примене, дипломирала је 2003. године са просечном оценом 9.32. На истом факултету је 2009. године завршила магистарске студије, на смеру Вероватноћа и математичка статистика. Магистарски рад под насловом *Генералисане Паретове расподеле у теорији екстремних вредности и примене*, који је радила под руководством проф. др Павла Младеновића, одбранила је 24. 3. 2009. године.

Од 2003. године запослена је на Катедри за физику и математику Фармацеутског факултета у Београду, где је тренутно у звању асистента. У периоду 2004-2008. радила је и као професор-спољни сарадник у Математичкој гимназији у Београду.

Јелена Јоцковић се бави научно-истраживачким радом у области Вероватноће и математичке статистике, а занима је и примена математике у другим областима.

До сада је објавила следеће научне радове:

1. J. Petrović, J. Jocković, S. Ibrić, J. Parojčić, Z. Djurić (2008): *Mathematical modeling of diclofenac sodium's release from polyethylene oxide matrices*, Journal of controlled release, 132 (3), e44-e45.

2. J. Petrović, S. Ibrić, J. Jocković, J. Parojčić, Z. Djurić (2009): *Determination of the percolation thresholds for polyethylene oxide and polyacrylic acid matrix tablets*, Journal of Drug Delivery Science and Technology, 19 (5), 359-364.

3. J. Petrović, J. Jocković, S. Ibrić, J. Parojčić, Z. Djurić (2009): *Modeling of diclofenac sodium diffusion from swellable and water-soluble polyethylene oxide matrices*, Journal of Pharmacy and Pharmacology, 61 (11), 1449-1456.

4. Lj. Djekić, M. Primorac, J. Jocković (2011): *Phase behaviour, microstructure and ibuprofen solubilization capacity of pseudo-ternary nonionic microemulsions*, Journal of Molecular Liquids, 160, 81-87.

5. J. Jocković, P. Mladenović (2011): *Coupon collector's problem and generalized Pareto distributions*, Journal of Statistical Planning and Inference, 141, 2348-2352.

6. Lj. Solomun, S. Ibrić, V. Pejanović, J. Djuriš, J. Jocković, P. Stanković, Z. Vujić (2012): *In silico methods in stability testing of hydrocortisone, powder for injections: Multiple regression analysis versus dynamic neural network*, Chemical Industry, DOI:10.2298/HEMIND120207023S

7. J. Jocković (2012): *Quantile estimation for the generalized Pareto distributions with application to finance*, Yugoslav Journal of Operations Research, DOI:10.2298/YJOR110308013J

8. J. Jocković (2013): *Correcting certain estimation methods for the generalized Pareto distribution*, Optimization Theory, Decision Making, and Operations Research Applications (Proceedings of BALCOR 2011), Springer Proceedings in Mathematics and Statistics, 31 (*in press*).

Имала је следећа саопштења на међународним конференцијама:

1. J. Petrović, S. Ibrić, J. Jocković, I. Tomić, A. M. Antić: *Determination of percolation threshold and drug release mechanism from polyethylene oxide metrix tablets*, PBP 2008, April 7–10, 2008, Barcelona, Spain.

2. J. Petrović, J. Jocković, S. Ibrić, J. Parojčić, Z. Djurić: *Mathematical modeling of diclofenac sodium's release from polyethylene oxide matrices*, ESCDD 2008, April 2-4, 2008, Noordwijk aan Zee, Netherlands.

3. J. Jocković: *Generalized Pareto distributions in extreme value theory*, MICOM 2009, September 16 – 20, 2009, Ohrid, Macedonia.

4. J. Jocković: *Raspodele ekstremnih vrednosti, generalisane Paretove raspodele i neke njihove primene*, SYM-OP-IS 2009, September 22 - 25, 2009, Ivanjica, Serbia.

5. P. Mladenović, J. Jocković: *Limit theorems for coupon collector's problem and generalized Pareto distributions*, EVA 2011, June 27-July 1 2011, Lyon, France.

6. J. Jocković: *Improving certain estimation methods for the generalized Pareto distribution*, BALCOR 2011, September 22 - 24, 2011, Thessaloniki, Greece.

Прилог 1.

Изјава о ауторству

Потписани-а Јелена Јоцковић

број уписа ______________________

Изјављујем

да је докторска дисертација под насловом

"Стохастички модели прекорачења високог нивоа и проблеми чекања"

- резултат сопственог истраживачког рада,
- да предложена дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, 10. септембра 2012.

Јелена Јоцковић

Прилог 2.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора __Јелена Јоцковић

Број уписа ____

Студијски програм __Вероватноћа и математичка статистика

Наслов рада __"Стохастички модели прекорачења високог нивоа и проблеми чекања"

Ментор __проф. др Павле Младеновић

Потписани __Јелена Јоцковић

изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла за објављивање на порталу **Дигиталног репозиторијума Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис докторанда

У Београду, 10. septembra 2012.

Јелена Јоцковић

Прилог 3.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

"Стохастички модели прекорачења високог нивоа и проблеми чекања"

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство
2. Ауторство - некомерцијално
3. Ауторство – некомерцијално – без прераде
4. Ауторство – некомерцијално – делити под истим условима
5. Ауторство – без прераде
6. Ауторство – делити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на полеђини листа).

Потпис докторанда

У Београду, 10. септембра 2012.

Јелена Јоцковић