Математички факултет
Универзитет у Београду

мастер рад

Екстремне статистике поретка и примјене у неживотном осигурању

Студент:
Оља Латиновић 1039/2015

Ментор:
др Павле Младеновић

10.11.2016.

Садржај

1 Увод

Теорија екстремних вриједности је алатка која се користи за разматрање вјероватноћа везаних за екстремне, односно ријетке догађаје. Корисна је при моделовању утицаја структурног лома. Користи се у моделовању понашања максимума или минимума у низу (реп расподјеле). Међутим, имплементација ове теорије има много препрека, укључујући оскудицу података са екстремним вриједностима, одређивање да ли се ради о расподјели са тешким репом, одабир прага гдје реп почиње, одабир метода за процјену параметара и друге.

До недавно, приступ вриједност-при-ризику (VaR) је био стандардан за индустрију управљања ризиком. VaR мјери колико би портфолио могао изгубити са датом вјероватноћом у одређеном периоду времена и при нормалним условима на тржишту. Овај приступ је критикован, највише из разлога што већина параметарских метода користе процјењивање нормалном расподјелом. Користећи овакву апроксимацију, ризик високих вриједности је потцијењен, посебно за расподјеле са тешким реповима, које су честе у финансијским подацима. Непараметарске методе се не ослањају на природу емпиријске расподјеле, али се ту такође појављују проблеми. Нпр. оне не могу да се користе за процјене за вриједности изван граница датог узорка.

Кључни резултат до којег су дошли Фишер[1] и Типет[2] 1928. о могућим граничним вриједностима максимума у узорку био је заслужан да се појави идеја да је теорија екстремних вриједности специфична, много другачија од класичне теорије граничних вриједности. Од тог тренутка десетине књига је написано о статистичким аспектима екстрема.

Обезбијеђене су тачније расподјеле које одговарају екстремним догађајима, а неке од њих су пригодне и за рјешавање проблема везаних за веома висока одступања, што је врло корисно за предвиђање ломова и великих штета. За разлику од VaR метода, не користе се претпоставке о природи расподјеле свих опсервација.

[1]Ronald Alymer Fisher (1890-1962), енглески статистичар и биолог
[2]Leonard Henry Caleb Tippett (1902-1985), енглески статистичар

Основе теорије екстремних вриједности су развијене 1928. године. Испод је дата кратка историја развоја ове теорије.

1928: Основе и асимптотска анализа развијена од стране Фишера и Типета

1940-е: Асимптотска теорија уједињена и проширена од стране Гнеденка[3] и Фон Мизеса[4]

1950-е: Гумбел[5] и Јенкинсон[6] користе асимптотске расподјеле за статистичко моделовање

1970-е: Класични закони граничних вриједности генерализовани од стране Пикандса[7] (и других)

1980-е: Надоградња теорије стационарних процеса, повезивање са теоријом тачкастих процеса, коришћење регресионих техника у анализи екстремних вриједности

1990-е: Развој мултиваријационих модела, повећање примјене у финансијама

2000-е: Примјена у испитивању климатских промјена

[3] Борис Владимирович Гнеденко (1912-1995), совјетски математичар, Колмогоровљев студент
[4] Ludwig von Mises (1881-1973), аустријски економски теоретичар
[5] Emil Julius Gumbel (1891-1966), њемачки математичар
[6] A. F. Jenkinson
[7] James Pickands III

2 Значај теорије екстремних вриједности у осигурању

Једна од најистакнутијих примјена теорије екстремних вриједности може да се види у неживотном осигурању. Неки портфолији имају тенденцију да повремено садрже велике штете које нарушавају солвентност портфолија, или чак угрожавају компанију. Осим великих непогода и несрећа као што су земљотреси, урагани, авионске несреће и остало, постоји и даље значајан број великих штета које се дешавају. С времена на вријеме, осигурање од аутоодговорности може бити у обавези да исплати велику штету. Још чешће, дешавају се велике штете у осигурању од пожара, а посебно пожара у индустрији, које могу да доведу и до губитка земљишта, раскида уговора и привремене незапослености.

Осигуравајуће друштво мора да се заштити од ризика које проузрокује портофлио који може да садржи штету која се, заправо, посматра више као екстремна него као просјечна. У ту сврху креирају се уговори о реосигурању гдје одређен проценат ризика на себе преузима реосигуравајуће друштво. Од велике је важности зато процјена репа расподјеле коју штета има, да би се склопио одговарајући уговор о реосигурању, као и да би се одредила одговарајућа премија.

Осим поменуте улоге теорије екстремних вриједности у осигурању, велике штете су утицале и на промјену финансијског тржишта везаног са осигурање. Област реосигурања стално доживљава пораст и по интезитету и по величини износа штета због што природних, што катастрофа изазваних од стране човјека. Финансијска индустрија је из тих разлога нудила осигуране производе у пољу осигурања од катастрофа. Неки од примјера су *CAT* (од енгл. *catastrophe* – катастрофа) фјучерс уговори, опције и обвезнице. Илустрације ради, један од примјера оваквих вриједносних папира су европске обвезнице које имају додатну купонску исплату која се исплаћује у случају да се деси неки прецизно дефинисани катастрофални догађај.

Ипак, вратићемо се сад на уговоре о реосигурању и рећи нешто више о њиховим врстама.

2.1 Уговори о реосигурању

Уговори о реосигурању се склапају између два осигуравајућа друштва ради прерасподјеле ризика. Осигуравач дио премије осигурања плаћа реосигуравачу као премију реосигурања, а овај се заузврат обавезује да ће му у случају настанка осигураног догађаја надокнадити дио обавезе према осигуранику. Премда се уговор о реосигурању и уговор о осигурању односе на исти предмет осигурања, они су међусобно независни. Стога се осигураник са одштетним захтјевом не може обратити реосигуравачу, који му често није ни познат, већ искључиво осигуравачу.

Уговори о реосигурању су најчешће подијељени на сљедећи начин:

- *Појединачни уговор о реосигурању* - покриће одређеног ризика, склапа се након закључивања уговора о осигурању а најчешће су то поморски и ваздушни каско и индустријски ризици

- *Општи уговор о реосигурању* - унапријед се за одређени удио или портфолио осигурања склапа прије закључивања уговора о осигурању

Постоје два начина утврђивања премије. Први начин је пропорционално реосигурање. Тада реосигуравач преузима на себе унапријед одређен постотак сваке штете која се деси, а добија тај исти постотак укупне премије коју је осигураник исплатио. Други начин је непропорционално реосигурање гдје је износ реосигуравачеве обавезе одређен висином одштете коју осигуравач треба да плати осигуранику. По томе се разликује од пропорционалног реосигурања код којег као полазна основа служи осигурани ризик, односно сума осигурања. Реосигуравач плаћа свој дио тек ако штета премаши самопридржај друштва.

Даље, пропорционално осигурање може бити:

- *Квотно (енгл. quota share reinsurance)* - осигуравач предаје реосигуравачу (или реосигуравачима) уговорени постотак (нпр. 40%) сваког ризика у одређеној врсти осигурања. Реосигуравач са тим удјелом учествује у свакој штети, као што са преузете обавезе добија одговарајући дио премије осигурања. Код оваквог начина уступања ризика не врши се никакав одабир, те се у реосигурање предају како мали, тако и велики ризици

- *Ексцедентно (реосигурање вишка ризика, енгл. surplus reinsurance)* – покрива само оне ризике код којих сума осигурања прелази одређену границу, тј. осигуравачев самопридржај. Значи да реосигурање не дејствује тамо гдје граница самопридржаја није прекорачена, а тамо гдје јесте дјелује, али само за износ прекорачења. Реосигуравач прима износ премије сразмјеран количини ризика који носи, а све штете које се јављају плаћа у истој размјери.

Могуће су и комбинације ове двије врсте реосигурања.

Основне форме непропорционалног реосигурања су:

- *Реосигурање вишка штета (енгл. excess of loss reinsurance)* – до одређене висине, односно самопридржаја, терет штете је на осигуравачу. Реосигурање дјелује тек изнад тога, када штету, по одбитку осигуравачевог дијела, преузима реосигуравач. Приликом уговарања он утврђује горњу границу своје обавезе. За разлику од реосигурања вишка ризика, гдје је мјерен ризик, овдје се обавеза реосигуравајућег друштва мјери искључиво полазећи од висине штете.

- *Реосигурање вишка губитка (енгл. stop loss reinsurance)* – реосигуравач преузима обавезу за накнаду штете изнад уговореног износа у одређеној врсти осигурања за извјесно временско раздобље, обично за годину дана. Осигуравач прије тога одреди свој самопридржај везан за једногодишње обавезе. Реосигурање му покрива само вишак штета као збирни цјелогодишњи износ који прелази самопридржај (с тим што се висина те обавезе по правилу ограничава). Наравно, осигуравач не може утврдити самопридржај тако да сваку годину завршава повољно превељујући губитке на реосигуравача. Због тога се у пракси могући губици из посла обично дијеле.

У уговорима о реосигурању за реосигуравајућу компанију значајан је износ одређеног броја максималних штета са којима се компанија суочила. Нека је k фиксиран број и $N(t)$ укупан број одштета. Означимо са $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(N(t))}$ статистике поретка формиране на основу скупа свих одштета.

Једна могућност код уговора је да реосигуравајућа компанија покрије k највећих одштета, односно збир

$$X_{(N(t))} + X_{(N(t)-1)} + \ldots + X_{(N(t)-k+1)}.$$

Друга могућност била би да реосигуравајуће друштво исплати сав вишак појединачних одштета у односу на k-ту највећу одштету, тј. суму

$$\sum_{j=1}^{k-1} (X_{N(t)-j+1} - X_{N(t)-k+1}) = \sum_{j=1}^{N(t)} (X_j - X_{N(t)-k+1}).$$

Проучавање својстава ових збирова захтијева примјену теорије екстремних вриједности, која се управо бави расподјелама и граничним расподјелама максимума у фамилији случајних величина.

3 Теорија екстремних вриједности

Због наведених догађаја, статистичка анализа екстремума је кључ за многе од проблема у управљању ризиком везаних за осигурање, реосигурање и финансије. У наставку ћемо дискутовати и приказати важне резултате теорије екстремних вриједности, уз поједностављене услове: претпоставићемо да су штете независне, једнако расподијељене и, наравно, увијек позитивне случајне величине.

Нека је $X_1, X_2, ...$ низ независних и једнако расподијељених случајних величина, а M_n максимум првих n, односно

$$M_n = \max\{X_2, X_2, ..., X_n\}. \tag{3.0.1}$$

Класична теорија екстремних вриједности бави се расподјелом максималне статистике M_n, посебно у случају када $n \to \infty$. Бавићемо се само максимумом, с обзиром да минималне вриједности нису значајне у осигурању, а уосталом, сви резултати се могу примијенити на минималну статистику ако посматрамо тривијалану релацију

$$m_n = \min\{X_2, X_2, ..., X_n\} = -\max\{-X_2, -X_2, ..., -X_n\}. \tag{3.0.2}$$

Нека X_i имају функцију расподјеле F. Тада је расподјела максималне статисике M_n једнака:

$$\begin{aligned} P\{M_n \leq x\} = P\{X_1 \leq x, X_2 \leq x, ..., X_n \leq x\} = \\ = P\{X_1 \leq x\} \cdot P\{X_2 \leq x\} \cdot \ldots \cdot P\{X_n \leq x\} = F^n(x). \end{aligned} \tag{3.0.3}$$

Централни резултат класичне теорије екстремних вриједности је теорема о екстремалним типовима. Она прецизира могуће облике граничне расподјеле максимума у низу једнако расподијељених независних случајних величина.

Могуће граничне расподјеле припадају класи расподјела које имају својство тзв. *максималне стаблности*. Показује се да се та класа састоји из три фамилије које називамо расподјелама екстремних вриједности.

У наставку ће, у првом одјељку, бити ријечи о конвергенцији функција расподјела уопште, о чему говори Хинчинова теорема, а у другом о појмовима максималне стабилности, као и области привлачења који имају велику улогу у теорији екстремних вриједности. Затим, у трећем одјељку долазимо до најзначајнијег резултата ове области статистике, односно до теореме о екстремалним типовима. С обзиром да главну улогу у поменутој теореми имају три расподјеле екстремних вриједности, у четвртом одјељку су дате неке од њихових карактеристика. У наставку су прецизирани услови за припадност свакој од три области привлачења. На крају ће бити ријечи о генералисаној Паретовој расподјели за коју се показује да је асимптотска расподјела репа расподјеле неке случајне промјенљиве, а која је од великог значаја за моделовање екстремних штета у осигурању.

3.1 Хинчинова теорема

Наредна, Хинчинова[8] теорема говори о конвергенцији функција расподјела.

Теорема 3.1.1 (Хинчин)**.** *Нека је (F_n) низ функција расподјеле, и G недегенерисана функција расподјеле. Нека су $a_n > 0$ и b_n, $n \in \mathbb{N}$ константе такве да важи*

$$\lim_{n\to\infty} F_n(a_n x + b_n) = G(x). \tag{3.1.1}$$

Тада, за неку недегенерисану расподјелу G_* и константе $\alpha_n > 0, \beta_n$, важи

$$\lim_{n\to\infty} F_n(\alpha_n x + \beta_n) = G_*(x) \tag{3.1.2}$$

ако и само ако

$$\lim_{n\to\infty} a_n^{-1}\alpha_n = a \ \text{и} \ \lim_{n\to\infty} a_n^{-1}(\beta_n - b_n) = b \tag{3.1.3}$$

за неко $a > 0$ и b, и тада

$$G_*(x) = G(ax + b). \tag{3.1.4}$$

Доказ. Ако означимо $\alpha_n' = a_n^{-1}\alpha_n, \beta_n' = a_n^{-1}(\beta_n - b_n)$ и $F_n'(x) = F_n(a_n x + b_n)$, можемо изразе 3.1.1, 3.1.2 и 3.1.3 написати као

$$\lim_{n\to\infty} F_n'(x) = G(x) \tag{3.1.5}$$

[8] Александр Яковлевич Хинчин (1894-1959), совјетски математичар

$$\lim_{n\to\infty} F'_n(\alpha'_n x + \beta'_n) = G_*(x) \tag{3.1.6}$$

$$\lim_{n\to\infty} \alpha'_n = a \text{ и } \lim_{n\to\infty} \beta'_n = b \text{ за неке } a > 0, b. \tag{3.1.7}$$

Ако 3.1.5 и 3.1.7 важе, слиједи да важи и 3.1.6, те $G_*(x) = G(ax+b)$. Тако из 3.1.1 и 3.1.3 слиједе 3.1.2 и 3.1.4.

Како претпостављамо да је функција расподјеле G_* недегенерисана, постоје двије различите тачке x' и x'' такве да важи $0 < G_*(x') < 1$ и $0 < G_*(x'') < 1$.

Низ $(\alpha'_n x' + \beta'_n)$ мора бити ограничен. У супротном, можемо изабрати низ (n_k) такав да је $\alpha'_{n_k} x' + \beta'_{n_k} \to \pm\infty$, што би, с обзиром на 3.1.5 и чињеницу да је G функција расподјеле, значило да је гранична вриједност за $F'_{n_k}(\alpha'_{n_k} x' + \beta'_{n_k})$ или 0 или 1, а то је у контрадикцији са 3.1.6 за $x = x'$. Како је $(\alpha'_n x' + \beta'_n)$ ограничен низ, на сличан начин можемо добити и да је $(\alpha'_n x'' + \beta'_n)$ ограничен, из чега слиједи да су низови (α'_n) и (β'_n) оба ограничена.

Дакле, постоје константе a и b и низ цијелих бројева (n_k) такви да важи

$$\lim_{k\to\infty} \alpha'_{n_k} = a \text{ и } \lim k \to \infty \beta'_{n_k} = b,$$

те, као у претходном случају, слиједи да је

$$\lim_{k\to\infty} F'_{n_k}(\alpha'_{n_k} x' + \beta'_{n_k}) = G(ax+b), \tag{3.1.8}$$

а одатле, како је по 3.1.6, $G(ax+b) = G_*(x)$ функција расподјеле, мора бити $a > 0$.

Када би постојао други низ цијелих бројева (m_k) за који важи $\alpha'_{m_k} \to a' > 0$ и $\beta'_{m_k} \to b'$, имали бисмо $G(a'x + b') = G_*(x) = G(ax+b)$, одакле би по посљедици 3.3.1 важило $a' = a$ и $b' = b$. Зато је $\alpha'_n \to a$ и $\beta'_n \to b$, чиме је доказ завршен. ∎

3.2 M-стабилност и област привлачења

У овом дијелу увешћемо значајне појмове када је ријеч о теорији екстремних вриједности, а то су појмови максималне стабилности и области привлачења расподјела.

Дефиниција 3.2.1. Нека су $a_n > 0$ и b_n $(n \in N)$ реални бројеви и нека за случајну величину M_n важи:

$$\lim_{n\to\infty} P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = G(x), \tag{3.2.1}$$

гдје је G недегенерисана функција расподјеле. Кажемо да G одређује *граничну расподјелу линеарно нормираног максимума* M_n, а a_n и b_n су *нормирајуће константе*.
Претходна једнакост се своди на

$$P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = P\{X_1 \leq a_n x + b_n, \ldots, X_n \leq a_n x + b_n\} =$$

$$= P\{X_1 \leq a_n x + b_n\}^n = F^n(a_n x + b_n) \tag{3.2.2}$$

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

Желимо да испитамо која је расподјела могућа као гранична расподјела, односно, ког облика је функција расподјеле $G(x)$.
Прије тога, уводимо појам типа расподјеле.

Дефиниција 3.2.2. Двије функције расподјеле G_1 и G_2 су истог типа ако важи

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

за неке константе $a > 0, b$.

Показаће се касније да G расподјела припада једном од сљедећа три типа:

1. тип (Гумбелова расподјела): $G_0(x) = \exp(-e^{-x}), \ -\infty < x < \infty$
2. тип (Фрешеова расподјела): $G_{1,\alpha}(x) = \begin{cases} 0, & \text{за } \ x \leq 0, \\ \exp(-x^{-\alpha}), & \text{за } \ x > 0. \end{cases}$
3. тип (Вејбулова расподјела): $G_{2,\alpha}(x) = \begin{cases} \exp(-(-x)^{\alpha}), & \text{за } \ x \leq 0, \\ 1, & \text{за } \ x > 0. \end{cases}$

за неко $\alpha > 0$.

Дефиниција 3.2.3. Недегенерисана функција расподјеле G је *максимум стабилна (М-стабилна)* ако за сваки природан број $n \geq 2$, постоје реалне константе $a_n > 0$ и b_n такве да за сваки реалан број x важи једнакост

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

Ову дефиницију сада можемо да формулишемо и на сљедећи начин, користећи појам типа расподјеле.

Дефиниција 3.2.4. Недегенерисана функција расподјеле G је М-стабилна ако је за свако $n \geq 2$ функција расподјеле G^n истог типа као G.

Хинчинова теорема дакле, користећи ову дефиницију, показује да ако је (F_n) низ функција расподјеле такав да је

$$\lim_{n\to\infty} F_n(a_n x + b_n) = G_1, \ \lim_{n\to\infty} F_n(\alpha_n x + \beta_n) = G_2 \ (a_n > 0, \alpha_n > 0),$$

онда су G_1 и G_2 истог типа, под условом да се ради о недегенерисаним расподјелама. Функције расподјеле очито могу да се подијеле у еквивалнтне класе, а те класе називамо *типовима* и кажемо да су G_1 и G_2 еквивалентне ако важи $G_2(x) = G_1(ax + b)$ за неке $a > 0, b$.

Ако су G_1 и G_2 истог типа (односно $G_2(x) = G_1(ax + b)$) и $F \in D(G_1)$, тј. $F^n(a_n x + b_n) \to G_1$ за неке $a_n > 0, b_n$ онда је 3.1.3 задовољено за $\alpha_n = a_n a$, $\beta_n = b_n + a_n b$ па је $F^n(\alpha_n x + \beta_n) \to G_2$ по Хинчиновој теореми, те је $F \in D(G_2)$. Из овога слиједи да ако су G_1 и G_2 истог типа, онда оне имају и исту област привлачења, односно тада је и $D(G_1) = D(G_2)$. Исто тако, видимо из Хинчинове теореме да ако нека функција расподјеле F припада и $D(G_1)$ и $D(G_2)$, онда су G_1 и G_2 истог типа.

Резултат ових разматрања је сљедећи:

Области привлачења $D(G_1)$ и $D(G_2)$ су једнаке ако и само ако су G_1 и G_2 истог типа. Односно, област привлачења функције расподјеле G зависи искључиво од типа G.

Везу између класе М-стабилних функција расподјела и недегенерисаних расподјела које се појављују као граничне у 3.2.1 даје нам сљедећа теорема.

Теорема 3.2.1. *(а) Недегенерисана функција расподјеле G је М-стабилна ако и само ако постоји низ функција расподјела F_n и константе $a_n > 0$ и b_n такве да за свако $k = 1, 2, ...$ важи*

$$\lim_{n\to\infty} F_n(a_{nk}x + b_{nk}) = G^{1/k}(x). \tag{3.2.3}$$

(б) Посебно, ако је G недегенерисана, област привлачења $D(G)$ није празан скуп ако и само ако је G М-стабилна расподјела. Тада, такође, важи $G \in D(G)$.
Тако добијамо да се класа недегенерисаних функција рсподјела G које се појављују као граничне расподјеле у 3.2.1 (за независне, једнако расподијељене случајне величине $X_1, X_2, \dots$) подудара са класом М-стабилних функција расподјела.

Доказ. (а) Нека важи да је G М-стабилна и означимо $F_n = G^n$. Тада имамо $G^n(a_nx + b_n) = G(x)$ за неке $a_n > 0$ и b_n, и важи

$$F_n(a_{nk}x + b_{nk}) = (G^{nk}(a_{nk}x + b_{nk}))^{1/k} = G(x)^{1/k},$$

из чега слиједи израз 3.2.3.
Обрнуто, нека је G недегенерисана функција расподјеле, онда је и $G^{1/k}$ недегенерисана за свако k. Нека 3.2.3 важи за свако k, Хинчинова теорема каже да је $G^{1/k}(x) = G(\alpha_k x + \beta_k)$ за неке $\alpha_k > 0$ и β_k, тако да је G М-стабилна.

(б) Ако је расподјела G М-стабилна, тада имамо $G^n(a_nx + b_n) = G(x)$ за неке $a_n > 0$ и b_n, па када $n \to \infty$, видимо да је $G \in D(G)$.
Обрнуто, ако $D(G)$ није празан скуп, постоји функција расподјеле F таква да важи $F \in D(G)$, односно таква да је $F^n(a_nx + b_n) \to G(x)$. Тада важи и $F^{nk}(a_{nk}x + b_{nk}) \to G(x)$, односно $F^n(a_{nk}x + b_{nk}) \to G^{1/k}(x)$. Сада 3.2.3 важи за $F_n = F^n$, па по (а) важи да је G М-стабилна функција расподјеле, чиме је доказ завршен.

■

Посљедица 3.2.1. Ако је G М-стабилна функција расподјеле, онда постоје реалне функције $a(s) > 0$ и $b(s)$ дефинисане за $s > 0$ тако да

$$G^s(a(s)x + b(s)) = G(x), \text{ за све } x, s > 0 \tag{3.2.4}$$

Доказ. С обзиром да је G М-стабилна, постоје $a_n > 0, b_n$ такви да за сваки природан број $n \geq 2$ и за сваки реалан број x важи једнакост

$$G^n(a_n x + b_n) = G(x), \tag{3.2.5}$$

а отуда и

$$\lim_{n\to\infty} G^{[ns]}(a_{[ns]}x + b_{[ns]}) = G(x)$$

гдје [ ] означава цијели дио броја. Из претходне релације видимо да важи

$$G^n(a_{[ns]}x + b_{[ns]}) \to G^{1/s}(x). \tag{3.2.6}$$

С обзиром на чињеницу да је $G^{1/s}$ недегенерисана, из 3.2.6 и 3.2.5, и из Хинчинове теореме слиједи да, ако је $\alpha_n = a_{[ns]}$ и $\beta_n = b_{[ns]}$, важи $G(a(s)x + b(s)) = G^{1/s}(x)$ за неке $a(s) > 0$ и $b(s)$.

■

3.3 Теорема о екстремалним типовима

Такозване расподјеле екстремних вриједности су три параметарске фамилије, које су познате као Гумбелова, Фрешеова и Вејбулова расподјела, чије смо стандардне представнике видјели у одјељку који говори о М-стабилним расподјелама и области привлачења.

Показаћемо да је функција расподјеле М-стабилна ако и само ако је истог типа као једна од функција расподјела екстремних вриједности.

Напомена: У наставку ћемо расподјеле које су истог типа као једна од расподјела екстремних вриједности такође називати расподјелама екстремних вриједности.

За потребе рада, прво ћемо приказати неке резултате везане за инверзе монотоних функција. Доказ наредне леме и њене посљедице може да се пронађе у [1].

Лема 3.3.1. *Нека је $\psi(x)$ неопадајућа непрекидна с десна функција. Инверз функције ψ^{-1} дефинишемо на отвореном интервалу $(\inf\{\psi(x)\}, \sup\{\psi(x)\})$ као*

$$\psi^{-1}(y) = \inf\{x; \psi(x) \geq y\}.$$

Тада:

(а) Ако су $a > 0, b$ и c константе, и $H(x) = \psi(ax + b) - c$, онда важи

$$H^{-1}(y) = a^{-1}(\psi^{-1}(y + c) - b).$$

(б) Ако је ψ^{-1} непрекидна, онда је $\psi^{-1}(\psi(x)) = x$

(в) Ако је G недегенерисана функција расподјеле, онда постоје $y_1 < y_2$ такви да су $G^{-1}(y_1) < G^{-1}(y_2)$ добро дефинисани и коначни.

Посљедица 3.3.1. Ако је G недегенерисана функција расподјеле, $a > 0, \alpha > 0, b$ и β константе такве да важи $G(ax + b) = G(\alpha x + \beta)$ за свако x, тада је $a = \alpha$ и $b = \beta$.

Теорема 3.3.1. *Свака М-стабилна расподјела је расподјела екстремних вриједности, тј. једнака је $G(ax + b)$ за $a > 0$ и b гдје је $G(x)$*

1. тип (Гумбелова расподјела): $G_0(x) = \exp(-e^{-x}), \; -\infty < x < \infty;$

2. тип (Фрешеова расподјела): $G_{1,\alpha}(x) = \begin{cases} 0, & \text{за } x \leq 0, \\ \exp(-x^{-\alpha}), & \text{за } x > 0; \end{cases}$

3. тип (Вејбулова расподјела): $G_{2,\alpha}(x) = \begin{cases} \exp(-(-x)^{\alpha}), & \text{за } x \leq 0, \\ 1, & \text{за } x > 0, \end{cases}$

за неко $\alpha > 0$. Параметар $\alpha > 0$ се назива параметар облика.

Обрнуто, свака расподјела екстремних вриједности је М-стабилна.

Доказ. Јасно је да је свака расподјела екстремних вриједности М-стабилна јер имамо, за нпр. расподјелу 1. типа сљедећи израз:

$$\{\exp(-e^{-(ax+b)})\}^n = \exp(-e^{-(ax+b-\ln n)}),$$

а слично важи и за расподјеле 2. и 3. типа.

За доказ директног смјера користимо Ханов[9] доказ (1976).

Ако је G максимално стабилна функција расподјеле, тада једнакост 3.2.4 важи за све реалне бројеве $s > 0$ и све x. Када је $0 < G(x) < 1$, по посљедици 3.2.1 логаритмовањем добијамо

$$-s \ln G(a(s)x + b(s)) = -\ln G(x),$$

односно, када ову једнакост логаритмујемо и помножимо са -1 имамо

$$-\ln(-\ln(G(a(s)x + b(s))) - \ln s = -\ln(-\ln(G(x)).$$

Из својства М-стабилних расподјела за $n = 2$, види се да G не може да има скок на крајњим тачкама. Зато је неопадајућа функција $\psi(x) = -\ln(-\ln G(x))$ таква да важи

$$\inf\{\psi(x)\} = -\infty, \; \sup\{\psi(x)\} = +\infty,$$

те ψ има инверзну функцију $U(y)$ дефинисану за свако $y \in R$. Даље, имамо

$$\psi(a(s)x + b(s)) - \ln s = \psi(x).$$

По леми 3.3.1 (а) важи

$$\frac{U(y + \ln s) - b(s)}{a(s)} = U(y).$$

Ако од овог израза одузмемо тај израз за $y = 0$ имамо

$$\frac{U(y + \ln s) - U(\ln s)}{a(s)} = U(y) - U(0).$$

[9] Laurens de Haan (1937 -)

Означимо $z = \ln s, \tilde{a}(z) = a(e^z)$ и $\tilde{U}(y) = U(y) - U(0)$, тада, за све реалне бројеве y, z :

$$\tilde{U}(y+z) - \tilde{U}(y) = \tilde{U}(y)\tilde{a}(z) \tag{3.3.1}$$

Замјењујући y и z видимо да важи

$$\tilde{U}(y)(1-\tilde{a}(z)) = \tilde{U}(z)(1-\tilde{a}(y)) \tag{3.3.2}$$

Могућа су два случаја:

(а) $\tilde{a}(z) = 1$ за свако z, гдје из 3.3.1 имамо

$$\tilde{U}(y+z) = \tilde{U}(y) + \tilde{U}(z).$$

Једина монотоно растућа функција која је рјешење ове једначине је функција $\tilde{U}(y) = \rho y$, за неко $\rho > 0$. Тада је $U(y) - U(0) = \rho y$ односно имамо

$$\psi^{-1}(y) = U(y) = \tilde{U}(y) + U(0) = \rho y + U(0).$$

С обзиром да је претходна функција непрекидна, лема 3.3.1 (б) нам даје

$$x = \psi^{-1}(\psi(x)) = \rho\psi(x) + U(0),$$

Односно $\psi(x) = (x - U(0))/\rho$, тако да важи

$$G(x) = exp\{-e^{-(x-U(0))/\rho}\}, \text{ када } 0 < G(x) < 1.$$

Као што је речено, G не може да има скок на крајњим тачкама, те она има форму изнад за свако x, односно, она јесте 1. типа (припада фамилији Гумбелове расподјеле).

(б) $\tilde{a}(z) \neq 1$ за неко z, па 3.3.2 даје

$$\tilde{U}(y) = \frac{\tilde{U(z)}}{1-\tilde{a}(z)}(1-\tilde{a}(y)) = c(1-\tilde{a}(y)) \tag{3.3.3}$$

гдје је

$$c = \tilde{U}(z) = \frac{\tilde{U}(z)}{1-\tilde{a}(z)} \neq 0.$$

Када би $\tilde{U}(z)$ било једнако нули, то би значило да је $\tilde{U}(y) = 0$ за свако y па би $U(y) = U(0)$ било константно.
Из израза 3.3.1 сада добијамо

$$c(1-\tilde{a}(y+z)) - c(1-\tilde{a}(z)) = c(1-\tilde{a}(y))\tilde{a}(z),$$

што нам даје $\tilde{a}(y+z) = \tilde{a}(y)\tilde{a}(z)$. Међутим, како је $\tilde{a}$ монотона функција (3.3.3), а једине монотоне неконстантне функције која су рјешења ове једначине имају облик $\tilde{a}(y) = e^{\rho y}$ за $\rho \neq 0$. Отуда и из 3.3.3 произилази

$$\psi^{-1}(y) = U(y) = U(0) + c(1 - e^{\rho y}).$$

Како је $-\ln(-\ln G(x))$ растућа функција, имамо да је и U растућа, па мора бити $c < 0$ ако је $\rho > 0$ и $c > 0$ ако је $\rho < 0$. По леми 3.3.1 (б) важи

$$x = \psi^{-1}(\psi(x)) = U(0) + c(1 - e^{\rho\psi(x)}) = U(0) + c(1 - (-\ln G(x))^{-\rho}),$$

гдје је $0 < G(x) < 1$,

$$G(x) = exp\left\{-\left(1 - \frac{x - U(0)}{c}\right)^{-1/\rho}\right\}.$$

Из непрекидности расподјеле G на било којој од коначних крајњих тачака, видимо да је G или 2. или 3. типа, гдје је

$$\alpha = \frac{1}{\rho} \text{ или } \alpha = -\frac{1}{\rho}$$

у зависности од тога да ли је $\rho > 0$ $(c < 0)$ или $\rho < 0$ $(c > 0)$

■

Слиједи главни резултат ових разматрања, теорема о екстремалним типовима.

Теорема 3.3.2 (Теорема о екстремалним типовима)**.** *Нека је $M_n = max\{X_1, X_2, \ldots, X_n\}$ гдје су X_i независне, једнако расподијељене случајне величине са функцијом расподјеле F. Ако за неке константе $a_n > 0, b_n$ важи*

$$\lim_{n\to\infty} P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = \lim_{n\to\infty} F^n(a_n x + b_n) = G(x), \tag{3.3.4}$$

за свако $x \in C(G)$, гдје је $C(G)$ скуп тачака непрекидности неке недегенерисане функције расподјеле G, онда је функција расподјеле G истог типа као нека од функција расподјела екстремних вриједности.

Обрнуто, свака функција расподјеле G која је расподјела екстремних вриједности може да се појави као гранична вриједност у 3.3.4, а када је G функција расподјеле сваког од X_i, тада је $G(x)$ управо гранична вриједност из 3.3.4.

Доказ. Ако 3.3.4 важи, из теореме 3.2.1 слиједи да је G М-стабилна, па, по теореми 3.3.1 она је функција расподјеле екстремних вриједности.

Обрнуто, ако је G функција расподјеле екстремних вриједности, она је М-стабилна на основу теореме 3.3.1, а теорема 3.2.1 (б) показује да $G \in D(G)$, из чега произилази дата теорема. ■

3.4 Расподјеле екстремних вриједности

С обзиром на резултат теореме о екстремалним типовима, тј. с обзиром на значај расподјела екстремних вриједности, у овом одјељку ће бити приказане основне особине те три расподјеле: Гумбелове, Фрешеове и Вејбулове.

3.4.1 Гумбелова расподјела

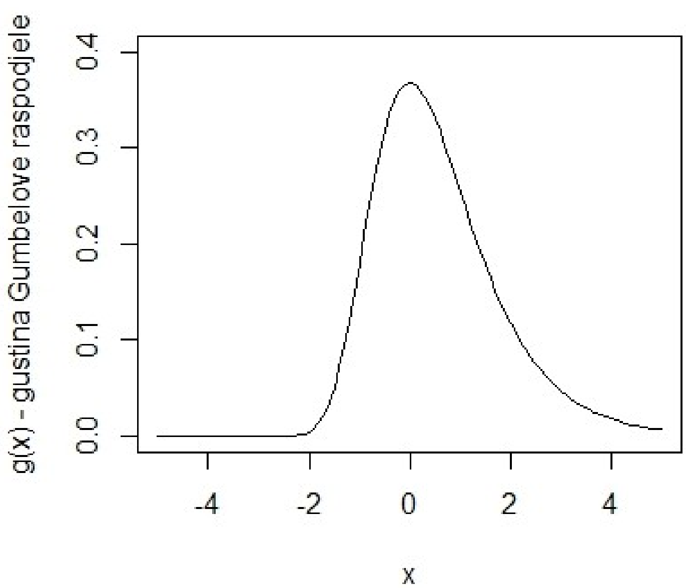
Функција расподјеле:

$$G_0(x) = e^{-e^{-x}}, \ -\infty < x < \infty;$$

Густина:

$$g_0(x) = e^{-x-e^{-x}}, -\infty < x < \infty.$$

График густине:



Слика 1: Густина Гумбелове расподјеле

Математичко очекивање и дисперзија:

$$E(X) = \gamma,$$

гдје је $\gamma = 0.57792...$ Ојлерова константа.

$$D(X) = \frac{\pi^2}{6}.$$

3.4.2 Фрешеова расподјела са параметром $\alpha > 0$

Функција расподјеле:

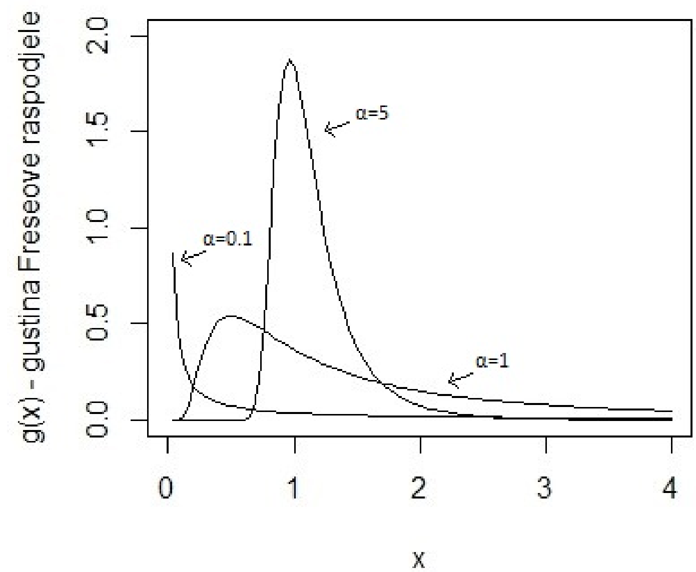
$$G_{1,\alpha}(x) = \begin{cases} 0, & \text{за} \;\; x \leq 0, \\ e^{-x^{-\alpha}}, & \text{за} \;\; x > 0; \end{cases}$$

где је $\alpha > 0$ параметар облика.

Густина:

$$g_{1,\alpha}(x) = \alpha x^{-(1+\alpha)} e^{-x^{-\alpha}}, x \geq 0.$$

График густине:



Слика 2: Густина Фрешеове расподјеле

Математичко очекивање и дисперзија:

m_r - момент r-тог реда

$$m_r = \Gamma\left(1 - \frac{r}{\alpha}\right), \alpha > r,$$

$$E(X) = \Gamma\left(1 - \frac{1}{\alpha}\right), \alpha > 1,$$

$$\mu_2 = D(X) = \Gamma\left(1 - \frac{2}{\alpha}\right) - \Gamma^2\left(1 - \frac{1}{\alpha}\right), \alpha > 2.$$

Важи

$$E(X) = +\infty \text{ ако је } 0 < \alpha \leq 1,$$

$$D(X) = +\infty \text{ ако је } 0 < \alpha \leq 2.$$

3.4.3 Вејбулова расподјела са параметром $\alpha > 0$

Функција расподјеле:

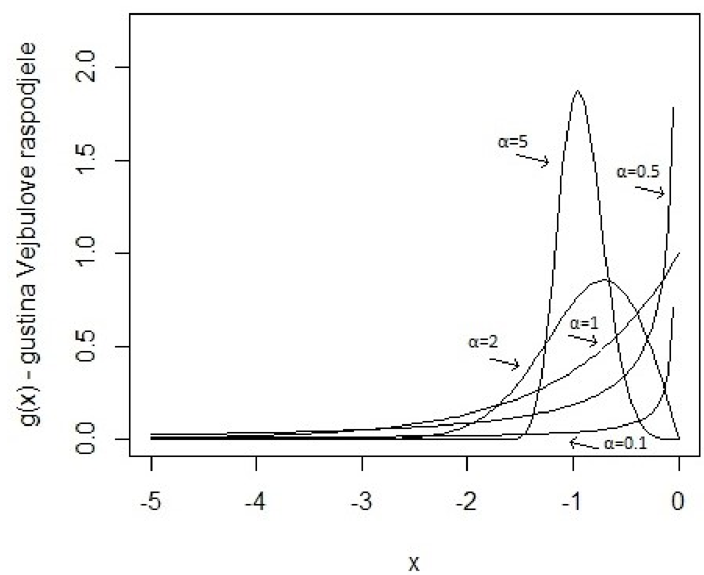
$$G_{2,\alpha}(x) = \begin{cases} e^{-(-x)^{\alpha}}, & \text{за } \ x \le 0, \\ 1, & \text{за } \ x > 0, \end{cases}$$

гдје је $\alpha > 0$ параметар облика.

Густина:

$$g_{2,\alpha}(x) = \alpha(-x)^{\alpha-1}e^{-(-x)^{\alpha}}, x \le 0.$$

График густине:



Слика 3: Густина Вејбулове расподјеле

Математичко очекивање и дисперзија:

$$m_r = (-1)^r\Gamma\left(1 + \frac{r}{\alpha}\right), r > 0$$

$$E(X) = -\Gamma\left(1 + \frac{1}{\alpha}\right)$$

$$\mu_2 = D(X) = \Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right), \alpha > 0$$

3.5 Асимптотско понашање вјероватноће $P\{M_n \leq u_n\}$

Израз

$$P\left\{\frac{M_n - b_n}{a_n} \leq x\right\}$$

можемо да напишемо као $P\{M_n \leq u_n\}$ гдје је $u_n = u_n(x) = a_n x + b_n$. Конвергенција израза се захтијева за свако x. Значајно је посматрати низ (u_n), који не мора да зависи од параметра x, или који може бити и у компликованијем облику од линеарне функције која је разматрана изнад.

Гранична расподјела максимума M_n одређена је асимптотским понашањем репа $1 - F(x)$ при $x \to x_F$ гдје је

$$x_F = sup\{t : F(t) < 1\},$$

тј. $F(x) < 1$ за све $x < x_F$ и $F(x) = 1$ за све $x \geq x_F$.

Наводимо два важна тврђења при проучавању асимптотског понашања максимума случајних величина.

Теорема 3.5.1. *Нека је (X_n) низ независних случајних величина са заједничком функцијом рсподјеле F, $M_n = max\{X_1, \ldots, X_n\}$, (u_n) низ реалних бројева и $0 \leq \tau \leq \infty$. Тада важи*

$$\lim_{n\to\infty} n(1 - F(u_n)) = \tau, \tag{3.5.1}$$

ако и само ако је

$$\lim_{n\to\infty} P\{M_n \leq u_n\} = e^{-\tau}. \tag{3.5.2}$$

Доказ. (a) Прво разматрамо случај када је $0 \leq \tau < \infty$. Ако важи

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

тада је

$$P\{M_n \leq u_n\} = F^n(u_n) = \{1 - (1 - F(u_n))\}^n$$
$$= \left\{1 - \frac{\tau}{n} + o\left(\frac{1}{n}\right)\right\}^n \to e^{-\tau}, n \to \infty.$$

Обрнуто, ако претпоставимо да 3.5.2 важи, тада мора бити $1 - F(u_n) \to 0$ када $n \to \infty$. У супротном би постојао позитиван број δ и подниз (n_k) такав да за све n_k важи неједнакост $1 - F(u_{n_k}) \geq \delta$, па би при $k \to \infty$ важило

$$P\{M_{n_k} \leq u_{n_k}\} = \{1 - (1 - F(u_{n_k})\}^{n_k} \to 0,$$

што је у контрадикцији са 3.5.2. Дакле, заиста важи $1 - F(u_n) \to 0$. На основу тога, из 3.5.2 добијамо да је

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

из чега, заједно са чињеницом да је $\ln(1-x) \sim -x$ када $x \to 0$, добијамо

$$n(1 - F(u_n))(1 + o(1)) \to \tau,$$

а одатле слиједи 3.5.1.

(б) Нека је сада $\tau = +\infty$ и нека важи 3.5.2 . Претпоставимо да 3.5.1 не важи. Тада постоји низ (n_k) природних бројева, такав да при $k \to \infty$ важи $n_k \to \infty$ и

$$n_k(1 - F(u_{n_k})) \to \tau_0 < +\infty.$$

На основу првог дијела доказа, важи $\{M_{n_k} \leq u_{n_k}\} \to e^{-\tau_0} \neq 0$, што је у овом случају у контрадикцији са 3.5.2, што потврђује да за $\tau = +\infty$ из 3.5.2 слиједи 3.5.1. Слично се доказује да 3.5.2 слиједи из 3.5.1 за $\tau = +\infty$.

■

Имамо сљедећу посљедицу ове теореме.

Посљедица 3.5.1. (а) $M_n \to x_F(\leq \infty)$ са вјероватноћом 1 када $n \to \infty$.

(б) Ако важи $x_F < \infty$ и $F(x_F - 0) < 1$ (односно ако F има скок у крајњој десној тачки), и ако за низ (u_n) важи $P\{M_n \leq u_n\} \to \rho$ када $n \to \infty$, тада је или $\rho = 0$ или $\rho = 1$.

Доказ. (а) Ако је $\lambda < x_F(\leq \infty), 1 - F(\lambda) > 0$ онда важи 3.5.1 за $u_n \equiv \lambda, \tau = \infty$ па 3.5.2 даје $\lim_{n\to\infty} P\{M_n \leq \lambda\} = 0$. Како је очито $P\{M_n > x_F\} = 0$ за свако n, слиједи да $M_n \to x_F$ у вјероватноћи. Како је низ (M_n) монотон, онда он и скоро сигурно конвергира.

(б) Претпоставимо сада да је $x_F < \infty$ и $F(x_F - 0) < 1$. Нека је низ (u_n) такав да је $P\{M_n \leq u_n\} \to \rho$. Како је $0 \leq \rho \leq 1$, можемо користити $\rho = e^{-\tau}$ гдје је $0 \leq \tau \leq \infty$, па по теореми 3.5.1 долазимо до $n(1 - F(u_n)) \to \tau$. Ако је $u_n < x_F$ за било које n, како је тада $1 - F(u_n) \geq 1 - F(X_F - 0) > 0$, мора бити $\tau = \infty$. Једина друга могућност је да је $u_n \geq x_F$ за све довољно велике вриједности n, што нам даје $n(1 - F(u_n)) = 0$ па је $\tau = 0$.
Дакле, $\tau = +\infty$ или $\tau = 0$, односно важи или $\rho = 0$ или $\rho = 1$, чиме је завршен доказ.

■

У сљедећем поглављу разматраћемо опште питање области привлачења екстремних расподјела користећи теорему 3.5.1. Међутим, најважнију позицију заузимају нормално расподијељени низови. Зато ће прво бити приказано како теорема 3.5.1 може директно да се примијени на граничну расподјелу 1. типа, односно на Гумбелову расподјелу, за низ независних случајних величина са нормалном расподјелом.

У наставку ћемо функцију расподјеле за нормално расподијељене случајне величине обиљежавати са $\Phi(x)$, а густину расподјеле са $\varphi(x)$. Користиће се познато својство репа расподјеле Φ :

$$1 - \Phi(u) \sim \frac{\varphi(u)}{u} \text{ када } u \to \infty. \tag{3.5.3}$$

Теорема 3.5.2. *Нека је (X_n) низ независних једнако расподијељених случајних величина са стандардном нормалном расподјелом. Тада асимптотска расподјела максимума $M_n = max(X_1, X_2, \ldots, X_n)$ има расподјелу истог типа као Гумбелова расподјела. Тачније, важи*

$$\lim_{n\to\infty} P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = exp(-e^{-x}), \tag{3.5.4}$$

гдје је

$$a_n = \frac{1}{\sqrt{2\ln n}},$$
$$b_n = \sqrt{2\ln n} - \frac{\ln\ln n + \ln 4\pi}{2\sqrt{2\ln n}}.$$

Доказ. Ако уведемо ознаку $\tau = e^{-x}$, израз 3.5.1 можемо написати као

$$1 - \Phi(u_n) = \frac{1}{n}e^{-x}.$$

Како је $1 - \Phi(u_n) \sim \varphi(u_n)/u_n$, имамо

$$\lim_{n\to\infty} \frac{1}{n}\frac{e^{-x}u_n}{\varphi(u_n)} = 1,$$

па је

$$\lim_{n\to\infty} (-\ln n - x + \ln u_n - \ln\varphi(u_n)) = 0,$$

односно, замјењујући густину нормалне расподјеле

$$\lim_{n\to\infty} \left(-\ln n - x + \ln u_n + \frac{1}{2}\ln(2\pi) + \frac{u_n^2}{2}\right) = 0. \tag{3.5.5}$$

Дијељењем са $\ln n$ слиједи да $u_n^2/(2\ln n) \to 1$, па

$$\lim_{n\to\infty}(2\ln u_n - \ln 2 - \ln\ln n) = 0,$$

односно,

$$\ln u_n = \frac{1}{2}(\ln 2 + \ln\ln n) + o(1).$$

Уврштавајући претходну једнакост у 3.5.5 добијамо

$$\frac{u_n^2}{2} = x + \ln n - \frac{1}{2}\ln(4\pi) - \frac{1}{2}\ln\ln n + o(1),$$

$$u_n^2 = 2\ln n\left\{1 + \frac{x - \frac{1}{2}\ln 4\pi - \frac{1}{2}\ln\ln n}{\ln n} + o\left(\frac{1}{\ln n}\right)\right\},$$

$$u_n = \sqrt{2\ln n}\left\{\frac{2\ln n + x - \frac{1}{2}\ln 4\pi - \frac{1}{2}\ln\ln n}{2\ln n} + o\left(\frac{1}{\ln n}\right)\right\},$$

$$u_n = \frac{x}{\sqrt{2\ln n}} + \sqrt{2\ln n} - \frac{\ln\ln n + \ln 4\pi}{2\sqrt{2\ln n}} + o\left(\frac{1}{\sqrt{2\ln n}}\right).$$

Дакле, имамо да важи

$$u_n = a_n x + b_n + o\left(\frac{1}{\sqrt{\ln n}}\right) = a_n x + b_n + o(a_n).$$

Како по 3.5.2 када је $\tau = e^{-x}$ имамо да је $P\{M_n \le u_n\} \to exp(-e^{-x})$ имамо да је

$$\lim_{x\to\infty} P\{M_n \le a_n x + b_n + o(a_n)\} = \exp(-e^{-x}),$$

односно

$$\lim_{x\to\infty} P\left\{\frac{M_n - b_n}{a_n} + o(1) \le x\right\} = \exp(-e^{-x}),$$

добијамо да слиједи 3.5.4 чиме је доказ завршен. ■

3.6 Област привлачења

Битно је знати која тачно од граничних расподјела одговара расподјели максимума датих независних, једнако расподијељених случајних промјенљивих односно, потребне и довољне услове при којима функција расподјеле припада одређеној области привлачења (Фрешеове, Вејбулове или Гумбелове расподјеле). Ове резултате до којих су дошли Фон Мизес (1936) и Гнеденко (1943) дајемо у наставку.

Теорема 3.6.1. *Нека је F апсолутно непрекидна функција расподјеле са густином расподјеле f. Тада су* ***довољни услови*** *при којима F припада свакој од могућих области привлачења дати испод.*

Област привлачења Гумбелове расподјеле:

(а) $f'(x) < 0$ *за све x из неког интервала (x_0, x_F) гдје је* $x_F = \sup\{t : F(t) < 1\} \leq +\infty$,

(б) $f(x) = 0$ *за* $x \geq x_F$,

(в) $\lim_{t \nearrow x_F} \frac{f'(t)(1-F(t))}{(f(t))^2} = -1$.

Област привлачења Фрешеове расподјеле:

(а) $f(x) > 0$ *за све* $x \geq x_0$ *и*

(б) $\lim_{t \to \infty} \frac{tf(t)}{1-F(t)} = \alpha > 0$.

Област привлачења Вејбулове расподјеле:

(а) $f(x) > 0$ *за свако x из неког интервала (x_0, x_F) гдје је* $x_F = \sup\{t : F(t) < 1\} < +\infty$,

(б) $f(x) = 0$ *за* $x > x_F$ *и*

(в) $\lim_{t \nearrow x_F} \frac{(x_F - t)f(t)}{1-F(t)} = \alpha > 0$.

Комплетан доказ за наведену теорему може да се пронађе у раду де Хан (1976), а овдје ће бити приказан само доказ за случај Фрешеове расподјеле, као примјер.

Доказ (случај области привлачења Фрешеове расподјеле). Претпоставимо да је

$$\lim_{t\to\infty} \frac{tf(t)}{1-F(t)} = \alpha > 0$$

гдје је $f(x) > 0$ за $x \geq x_0$. Ако је

$$\alpha(t) = \frac{tf(t)}{1-F(t)},$$

добијамо да за $x_2 \geq x_1 \geq x_0$ важи

$$\int_{x_1}^{x_2} \frac{\alpha(t)}{t} dt = -\ln(1-F(x_2)) + \ln(1-F(x_1)),$$

па је

$$1 - F(x_2) = (1-F(x_1)) e^{-\int_{x_1}^{x_2} \frac{\alpha(t)}{t} dt}.$$

Очито, постоји γ_n такво да је $1 - F(\gamma_n) = \frac{1}{n}$, и ако узмемо да је $x_1 = \gamma_n, x_2 = \gamma_n x$ (или обрнуто у зависности од тога да ли је $x \geq 1$ или $x < 1$) добијамо

$$n(1-F(\gamma_n x)) = exp\left(-\int_{\gamma_n}^{\gamma_n x} \frac{\alpha(t)}{t} dt\right) = exp\left(-\int_1^x \frac{\alpha(\gamma_n s)}{s} ds\right),$$

па како $\gamma_n \to \infty$, претходни израз конвергира ка $e^{-\alpha \ln x} = x^{-\alpha}$ када $n \to \infty$. Дакле, за $x > 0$ по теореми 3.5.1

$$\lim_{n\to\infty} P\{M_n \leq \gamma_n x\} = exp(-x^{-\alpha}).$$

За $x \leq 0$ како је $P\{M_n \leq \gamma_n x\} \leq P\{M_n \leq \gamma_n y\}$ за $y > 0$ и $P\{M_n \leq \gamma_n y\} \to exp(-y^{-\alpha})$, видимо да, када $y \to 0$ важи $\lim_{n\to\infty} P\{M_n \leq \gamma_n x\} = 0$. Тиме је доказ у случају области привлачења Фрешеове расподјеле завршен са нормирајућим константама $a_n = \gamma_n$ и $b_n = 0$. ■

Претходни доказ (за област привлачења Фрешеове расподјеле) се заснива на постојању низа (γ_n) таквог да је $1 - F(\gamma_n) \sim 1/n$, па су онда нормирајуће константе a_n, b_n из израза $P\{(M_n - b_n)/a_n \leq x\} \to G(x)$ представљене преко γ_n, у овом случају $a_n = \gamma_n, b_n = 0$. За произвољну функцију расподјеле F такво γ_n не мора увијек да постоји, али ако F припада области привлачења једне од екстремних функција расподјеле, онда се може наћи такво γ_n.
Ако

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

и $G(x)$ је непрекидна функција расподјеле, онда x можемо изабрати такво да је $G(x) = e^{-1}$, па имамо

$$\lim_{n\to\infty} P\{M_n \leq \gamma_n\} = e^{-1}$$

гдје је $\gamma_n = a_n x + b_n$. По теореми 3.5.1 тада је $1 - F(\gamma_n) \sim 1/n$.

Сада ћемо формулисати потребне и довољне услове при којима дата расподјела припада једној од три области привлачења у општем случају, дакле, не ради се више искључиво о апсолутно непрекидним расподјелама. Довољни услови се такође заснивају на постојању споменутог низа (γ_n).

Теорема 3.6.2. ***Потребни и довољни услови*** *да функција расподјеле F припада свакој од могуће три области привлачења су сљедећи:*

Област привлачења Гумбелове расподјеле:
Постоји строго позитивна функција $g(t)$ таква да важи

$$\lim_{t \nearrow x_F} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x}$$

за свако реално x, гдје је $x_F = \sup\{t : F(t) < 1\}$.

Област привлачења Фрешеове расподјеле:

(а) $x_F = \sup\{t : F(t) < 1\} = +\infty$;

(б) $\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha}$ *за свако $x > 0$ гдје је $\alpha > 0$.*

Област привлачења Вејбулове расподјеле:

(а) $x_F = \sup\{t : F(t) < 1\} < +\infty$; *и*

(б) $\lim_{h\searrow 0} \frac{1-F(x_F - xh)}{1-F(x_F - h)} = x^{\alpha}$ *за свако $x > 0$ гдје је $\alpha > 0$.*

Заправо, у случају када функција расподјеле F припада области привлачења Гумбелове расподјеле, може се показати да важи $\int_0^\infty (1 - F(u))du < +\infty$, а функцију g можемо одредити на сљедећи начин:

$$g(t) = \frac{\int_t^{x_F} (1 - F(u))du}{1 - F(t)}$$

за $t < x_F$.

Доказ (да су дати услови довољни). Претпоставимо да постоји низ (γ_n) такав да важи $1 - F(\gamma_n) \to \frac{1}{n}$ односно $n(1 - F(\gamma_n)) \to 1$ (што ће бити показано на крају овог доказа). Константе γ_n се разликују за сваки од три типа области привлачења. За довољно велико n важи, очито, $\gamma_n \to x_F$ и $\gamma_n < x_F$.

Област привлачења Фрешеове расподјеле:

Због једноставности, прво доказујемо теорему у случају области привлачења Фрешеове расподјеле. Ако F задовољава написане услове, означићемо t са γ_n, па за свако $x > 0$ имамо

$$n(1 - F(\gamma_n x)) \sim n(1 - F(\gamma_n))x^{-\alpha} \to x^{-\alpha},$$

при $n \to \infty$.
По теореми 3.5.1 из сљедећег израза за $x > 0$ важи

$$\lim_{n\to\infty} P\{M_n \leq \gamma_n x\} = exp\{-x^{-\alpha}\}.$$

Како је $\gamma_n > 0$ за довољно велико n и десна страна претходног израза тежи 0 када $x \searrow 0$, слиједи да је $P\{M_n \leq 0\} \to 0$, а за $x < 0$ да је $P\{M_n \leq \gamma_n x\} \leq P\{M_n \leq 0\} \to 0$. Дакле, имамо $P\{M_n \leq \gamma_n x\} \to G(x)$, гдје је G функција расподјеле другог типа из теореме 3.3.1. Ово може да се изложи и као

$$P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \to G(x), \tag{3.6.1}$$

гдје је $a_n = \gamma_n$ и $b_n = 0$, из чега слиједи да функција расподјеле F припада области привлачења Фрешеове расподјеле.

Област привлачења Вејбулове расподјеле:

Претпоставимо да важе услови за Вејбулову расподјелу из теореме. Ако узмемо да је $h_n = x_F - \gamma_n$, што тежи нули, имамо за свако $x > 0$ сљедећу једнакост

$$\lim_{n\to\infty} n(1 - F\{x_F - x(x_F - \gamma_n)\}) = x^{\alpha}.$$

Из те једнакости, ако замијенимо x са $-x$, за $x < 0$ добија се

$$\lim_{n\to\infty} n(1 - F\{x_F + x(x_F - \gamma_n)\}) = (-x)^{\alpha}.$$

По теореми 3.5.1 добијамо да сада функција расподјеле F припада области привлачења Вејбулове расподјеле ако за константе из 3.6.1 узмемо

$$a_n = x_F - \gamma_n \text{ и } b_n = x_F.$$

Област привлачења Гумбелове расподјеле:

Када функција расподјеле F задовољава одговарајуће критеријуме из теореме, имамо да за свако x важи, ако користимо $t = \gamma_n \nearrow x_F (\leq \infty)$, сљедеће

$$\lim_{n\to\infty} n(1 - F\{\gamma_n + xg(\gamma_n)\}) = e^{-x},$$

а из чега (такође по теореми 3.5.1) добијамо да дата функција расподјеле припада области привлачења Гумбелове расподјеле ако за константе узмемо

$$a_n = g(\gamma_n) \text{ и } b_n = \gamma_n.$$

Сада коначно доказујемо да неопадајући низ (γ_n) који задовољава једнакост $\lim_{n\to\infty} n(1 - F(\gamma_n)) = 1$ постоји. За γ_n можемо узети било који неопадајући низ такав да важи

$$F(\gamma_n - 0) \leq 1 - \frac{1}{n} \leq F(\gamma_n)$$

(као што је нпр. низ $\gamma_n = F^{-1}(1 - 1/n) = \inf\{x \mid F(x) \geq 1 - 1/n\})$). За такав низ имамо $n(1 - F(\gamma_n)) \leq 1$ из чега тривијално слиједи и $\limsup_{n\to\infty} n(1 - F(\gamma_n)) \leq 1$.

Преостаје још да се покаже да у сваком случају важи $\liminf_{n\to\infty} n(1 - F(\gamma_n)) \geq 1$. Како имамо $n(1-F(\gamma_n-0)) \geq 1$, довољно ће бити да покажемо да важи неједнакост

$$\liminf_{n\to\infty} \frac{1 - F(\gamma_n)}{1 - F(\gamma_n - 0)} \geq 1. \tag{3.6.2}$$

Поново ћемо почети од случаја за област привлачења **Фрешеове расподјеле**. Ако функција расподјеле F задовољава одговарајуће услове, лијева страна неједнакости 3.6.2 није мања од

$$\liminf_{n\to\infty} \frac{1 - F(\gamma_n)}{1 - F(\gamma_n x)} = x^{\alpha},$$

за било које $x < 1$, из чега 3.6.2 слиједи ако пустимо да $x \to 1$.

Слично се доказује и случај када функција расподјеле F задовољава услове за област привлачења **Вејбулове расподјеле**. Лијева страна израза 3.6.2 за $x > 1$ и $h_n = x_F - \gamma_n$ не може бити мања од

$$\liminf_{n\to\infty} \frac{1 - F(x_F - h_n)}{1 - F(x_F - xh_n)} = x^{-\alpha},$$

што, када $x \to 1$, такође тежи ка 1, из чега слиједи 3.6.2.

Коначно, када функција расподјеле F задовољава услове за област привлачења **Гумбелове расподјеле**, лијева страна израза 3.6.2 за $x < 0$ не може бити мања од

$$\liminf_{n\to\infty} \frac{1 - F(\gamma_n)}{1 - F(\gamma_n + xg(\gamma_n))} = e^{x},$$

што тежи 1 када $x \to 0$. Дакле, имамо да 3.6.2 важи и у овом случају. ■

Сумираћемо сада резултате до којих смо дошли у доказу претходне теореме у наредној посљедици.

Посљедица 3.6.1. Константе a_n и b_n у конвергенцији

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

могу, у зависности од 3 могућа типа граничних расподјела, да се изаберу на сљедећи начин:

1. тип (Гумбелова расподјела): $a_n = g(\gamma_n)$ и $b_n = \gamma_n$;
2. тип (Фрешеова расподјела): $a_n = \gamma_n$ и $b_n = 0$;
3. тип (Вејбулова расподјела): $a_n = x_F - \gamma_n$ и $b_n = x_F$,

гдје је

$$\gamma_n = F^{-1}\left(1 - \frac{1}{n}\right) = \inf\left\{x \mid F(x) \geq 1 - \frac{1}{n}\right\}.$$

Доказ. Погледати доказ теореме 3.6.2. ■

3.7 Генералисане Паретове расподјеле

Основа теорије екстремних вриједности је теорема о екстремалним типовима. Међутим, постоји још једна значајна теорема, коју називају и другом теоремом о екстремалним типовима, а тиче се тзв. генералисане Паретове расподјеле. Наредни резултати описују понашање вриједности које су веће од одређеног прага u, односно понашање репа расподјеле неке случајне промјенљиве, а они се и најчешће користе при моделовању одштета у осигурању. Дају нам одговор на питање: ако су вриједности екстремне, колико екстремне би могле бити? Расподјела до које се долази у овим резултатима је управо генералисана Паретова расподјела.

Из датог узорка екстремне вриједности се могу издвојити на више начина, од којих је најчешћи метод прекорачења датог прага (енгл. *peaks-over-threshold method)*.

Нека је дат случајан узорак $(X_1, X_2, \dots, X_N)$, гдје су $X_1, X_2, \dots, X_N$ независне случајне величине са истом функцијом расподјеле F и нека је одређен праг u. Прекорачења прага су сљедеће вриједности веће од u (којих има одређен број n)

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

Број прекорачења N_u је случајна величина дефинисана као

$$N_u = \sum_{i \le N} I\{X_i > u\}.$$

N_u има биномну расподјелу са параметрима n и p, гдје је $p = 1 - F(u)$. Средња вриједност броја прекорачења датог прага u је онда

$$E(N_u) = n(1 - F(u)).$$

Нека је X случајна величина са функцијом расподјеле F. Прекорачења се појављују под условом да је вриједност случајне величине X већа од прага u, па је функција расподјеле прекорачења прага F_u једнака:

$$F_u'(x-u) = P\{X \le x \mid X > u\} = \frac{P\{X \le x, X > u\}}{P\{X > u\}} = \frac{F(x) - F(u)}{1 - F(u)}, \qquad (3.7.1)$$

за $x \ge u$, или

$$F_u(x) = P\{X \le x + u \mid X > u\} = \frac{F(x+u) - F(u)}{1 - F(u)},$$

за $x \ge 0$.

Математичко очекивање величине прекорачења датог прага u, за дату функцију расподјеле F, назива се средње прекорачење:

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

Испоставља се да су све могуће непрекидне функције расподјеле које се могу појавити као граничне расподјеле прекорачења прага u када u тежи десном крају носача расподјеле F, истог типа као једна од наредних функција расподјеле.
Експоненцијална расподјела:

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

Паретова расподјела с параметром α:

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

Бета расподјела с параметром α:

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

Ове три расподјеле се називају генералисане Паретове расподјеле. Ако се уведе смјена $\gamma = \frac{1}{\alpha}$ код Паретове расподјеле и $\gamma = -\frac{1}{\alpha}$ код бета расподјеле добије се јединствен, неприкадан модел $F_\gamma(x)$, при чему важи

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

Тада је:
Експоненцијална расподјела:

$$W_0(x) = 1 - e^{-x}, \gamma = 0, x \geq 0.$$

Паретова расподјела с параметром γ:

$$W_{1,\gamma}(x) = 1 - (1 + \gamma x)^{-\frac{1}{\gamma}}, x \geq 0, \gamma > 0.$$

Бета расподјела с параметром γ:

$$W_{2,\gamma} = 1 - (1 + \gamma x)^{-\frac{1}{\gamma}}, x \in \left[0, -\frac{1}{\gamma}\right], \gamma < 0.$$

Очита је веза између расподјела екстремних вриједности и генералисаних Паретових расподјела:

$$W_0(x) = 1 + \ln G_0(x), \text{ у случају када је } \ln G_0(x) > -1$$

$$W_{1,\alpha}(x) = 1 + \ln G_{1,\alpha}(x), \text{ у случају када је } \ln G_{1,\alpha}(x) > -1$$

$$W_{2,\alpha}(x) = 1 + \ln G_{2,\alpha}(x), \text{ у случају када је } \ln G_{2,\alpha}(x) > -1.$$

Генералисана Парето расподјела је у литератури најчешће представљена као двопараметарска расподјела чија је функција расподјеле:

$$W_{\gamma,\beta}(x) = \begin{cases} 1 - \left(1 + \gamma \frac{x}{\beta}\right)^{-\frac{1}{\gamma}}, & \text{за } \gamma \neq 0, \\ 1 - e^{-\frac{x}{\beta}}, & \text{за } \gamma = 0, \end{cases}$$

гдје је $\beta > 0$ параметар скалирања, а носач је $x \geq 0$ за параметар облика $\gamma \geq 0$ и $0 \leq x \leq -\beta/\gamma$ за $\gamma < 0$.

Дакле, тзв. друга теорема о екстремалним типовима, до које су дошли Пикандс (1975), Балкема и де Хан[10] (1974) нам управо говори да су све непрекидне расподјеле које се могу појавити као граничне расподјеле прекорачења прага u истог типа као генералисана Паретова расподјела:

Теорема 3.7.1 (Пикандс-Балкема-де Хан теорема)**.** *Нека је* $(X_1, X_2, ..., X_3)$ *низ независних случајних промјенљивих са функцијом расподјеле* F *и нека је* F_u *њена функција расподјеле прекорачења прага* u*. Тада важи:*

$$\lim_{u \to \infty} F_u(x) = W_{\gamma,\beta}(x), \tag{3.7.2}$$

гдје је $W_{\gamma,\beta}(x)$ *генералисана Паретова расподјела.*

Након изабраног прага u процјењују се γ и β најчешће преко метода максималне вјеродостојности, док је сами избор прага прекорачења u компромис између избора прага довољно високог да се теорема може сматрати валидном, а довољно ниског да имамо довољно материјала за процјену потребих параметара.

Да би се процијенио реп расподјеле користимо 3.7.1 и за $x \geq u$ имамо

$$F(x) = (1 - F(u))F_u'(x - u) + F(u).$$

Сада, по теореми 3.7.1 знамо да можемо да процијенимо $F_u'(x-u)$ са $W_{\gamma,\beta}(x-u)$ за велико u. $F(u)$ можемо процијенити из података помоћу $F_n(u)$, емпиријске функције расподјеле у тачки u. То значи да за $x \geq u$ можемо да користимо сљедећу процјену репа

$$\widehat{F(x)} = (1 - F_n(u))W_{\gamma,\beta}(x) + F_n(u).$$

[10]Laurens de Haan (1937 -) - холандски економиста

4 Примјери

У овом поглављу ће бити приказано неколико примјера разних расподјела које припадају једној од три области привлачења екстремних расподјела. Као и раније, константе a_n и b_n су константе из израза

$$\lim_{n \to +\infty} P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = G(x). \tag{4.0.3}$$

Затим ће методологија из седмог одјељка претходне главе бити примијењена (мада би потпуна анализа била далеко опширнија) на реалне податке - одштетне захтјеве у области осигурања од пожара.

4.1 Област привлачења Гумбелове расподјеле

Примјер 4.1.1 (*Гумбелова расподјела*)**.** Већ је речено да свака расподјела екстремног типа и сама припада својој области привлачења. Дакле, када је ријеч о Гумбеловој расподјели имамо

$$F(x) = e^{-e^{-x}},$$

$$F^n(x) = e^{-e^{-x+\ln n}},$$

$$F^n(x + \ln n) = F(x).$$

Тако је 4.0.3 задовољено за коефицијенте $a_n = 1$ и $b_n = \ln n$.

Примјер 4.1.2 (*Нормална расподјела*)**.** У теореми 3.5.2 видјели смо да нормална расподјела припада области привлачења Гумбелове расподјеле са константама a_n и b_n

$$a_n = \frac{1}{\sqrt{2\ln n}},$$

$$b_n = \sqrt{2\ln n} - \frac{\ln\ln n + \ln 4\pi}{2\sqrt{2\ln n}}.$$

Примјер 4.1.3 (*Експоненцијална расподјела*)**.** Нека $x \in \varepsilon(1)$, односно, x има функцију расподјеле

$$F(x) = 1 - e^{-x}.$$

Када је τ из теореме 3.5.1 позитиван број, можемо да изаберемо u_n такво да је $1 - F(u_n) = \tau/n$ тако што ћемо изабрати

$$u_n = -\ln\frac{\tau}{n} = -\ln\tau + \ln n,$$

па важи (по теореми 3.5.1)

$$\lim_{n\to\infty} P\{M_n \leq -\ln\tau + \ln n\} = e^{-\tau}.$$

Означавајући $\tau = e^{-x}$ долазимо до 4.0.3 за

$$a_n = 1, b_n = \ln n \text{ и } G(x) = e^{-e^{-x}}.$$

Када је ријеч о експоненцијалној расподјели са параметром λ добијамо да су нормирајуће константе

$$a_n = \frac{1}{\lambda} \text{ и } b_n = \frac{1}{\lambda}\ln n.$$

4.2 Област привлачења Фрешеове расподјеле

Примјер 4.2.1 (*Фрешеова расподјела*)**.** Наравно, и Фрешеова расподјела припада својој области привлачења (теорема 3.2.1). С обзиром да важи

$$F^n(n^{\frac{1}{\alpha}}x) = F(x)$$

долазимо до нормирајућих константи $a_n = n^{\frac{1}{\alpha}}$ и $b_n = 0$.

Примјер 4.2.2 (*Паретова расподјела*)**.** Нека је

$$F(x) = 1 - kx^{-\alpha}, \text{гдје је } \alpha > 0, k > 0, x \geq k^{\frac{1}{\alpha}}, \text{и}$$

$$F(x) = 0 \text{ за } x < k^{\frac{1}{\alpha}}.$$

Тада је

$$\lim_{n\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha},$$

те, на основу теореме 3.6.2 закључујемо да Паретова расподјела припада Фрешеовој области привлачења. Тада имамо

$$\lim_{n\to\infty} P\left\{M_n \leq \left(\frac{kn}{\tau}\right)^{\frac{1}{\alpha}}\right\} = e^{-\tau}.$$

Ако ставимо $\tau = x^{-\alpha}$ за $x \geq 0$, добијамо

$$\lim_{n\to\infty} P\left\{\frac{M_n}{(kn)^{1/\alpha}} \leq x\right\} = e^{-x^{-\alpha}},$$

односно, нормирајуће константе су дате са $a_n = (kn)^{\frac{1}{\alpha}}$ и $b_n = 0$.

Примјер 4.2.3 (*Кошијева расподјела*)**.** Нека је

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x.$$

Тада важи

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)} = \lim_{t\to\infty}\frac{\frac{1}{2}-\frac{1}{\pi}\operatorname{arctg}(tx)}{\frac{1}{2}-\frac{1}{\pi}\operatorname{arctg} t}$$

$$\lim_{t\to\infty}\frac{\frac{\pi}{2}-\operatorname{arctg}(tx)}{\frac{\pi}{2}-\operatorname{arctg} t} = \frac{1}{x}.$$

Када уведемо смјену $\operatorname{arctg} t = \pi/2 - s$ имамо $t = \operatorname{tg}(\pi/2 - s) = \operatorname{ctg} s$, па је

$$\lim_{t\to\infty}\left(\frac{\pi}{2}-\operatorname{arctg} t\right)t = \lim_{s\to 0} s\operatorname{ctg} s = 1, \text{ и}$$

$$\lim_{t\to\infty}\left(\frac{\pi}{2}-\operatorname{arctg}(tx)\right)t = 1.$$

Зато, по теореми 3.6.2, добијамо да Кошијева расподјела заиста припада области привлачења Фрешеове расподјеле. Нормирајуће константе су a_n и $b_n = 0$ при чему a_n одређујемо из услова

$$1 - F(a_n) = \frac{1}{2} - \frac{1}{\pi}\operatorname{arctg} a_n = \frac{1}{n}.$$

Односно, $a_n = \operatorname{tg}(\frac{\pi}{2} - \frac{\pi}{n}) = \operatorname{ctg}\frac{\pi}{n}$.

4.3 Област привлачења Вејбулове расподјеле

Примјер 4.3.1 (*Вејбулова расподјела*)**.** Као и у претходна два случаја, и Вејбулова расподјела припада својој области привлачења. За сваки природан број n и за сваки реалан број x важи

$$F^n(n^{-\frac{1}{\alpha}}x) = F(x)$$

па су нормирајуће константе дате са

$$a_n = n^{-\frac{1}{\alpha}} \text{ и } b_n = 0.$$

Примјер 4.3.2 (*Равномјерна расподјела на интервалу* (0, l))**.** Нека је функција расподјеле дата са

$$F(x) = \frac{x}{l} \text{ за } 0 \le x \le l, l > 0,$$

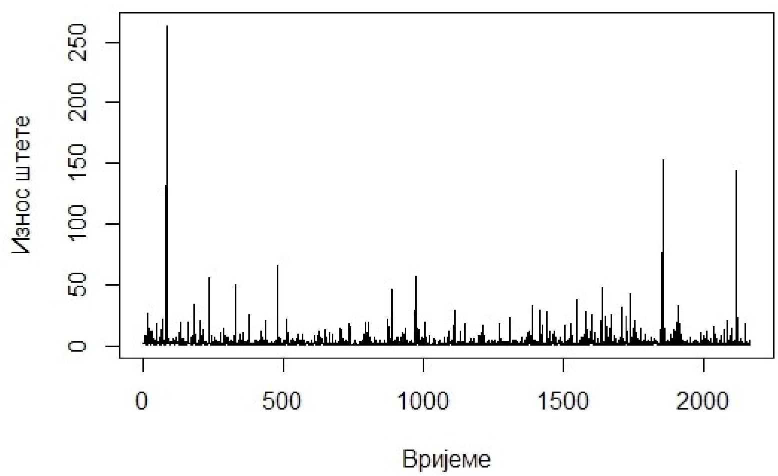
и $F(x) = 0$ за $x < 0, F(x) = 1$ за $x > l$. Тада, за свако $x > 0$ важи

$$\lim_{t\downarrow 0}\frac{1-F(x_F - tx)}{1-F(x_F - t)} = \lim_{t\downarrow 0}\frac{1-F(l-tx)/l}{1-F(l-t)/l} = x,$$

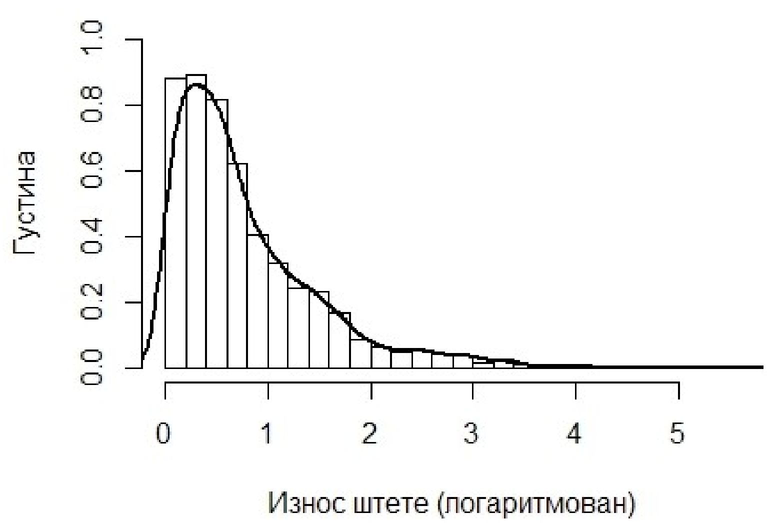
из чега закључујемо да равномјерна расподјела заиста припада области привлачења Вејбулове расподјеле.

4.4 Осигурање од пожара

Примјер 4.4.1. У статистичком програму *R*, постоји база података *danish* која обухвата дневне износе штета насталих усљед пожара у Данској у периоду од јануара 1980. до децембра 1990. године. Величина узорка је 2157. Износи су мјерени у милионима Данских круна (тада је однос валута био 1.000.000 DKK = 128.980,74 EUR). Примијенићемо теорију екстремних вриједности на ове податке. Односно, користећи историјске податке о штетама, предвиђамо величину будућих штета, уз претпоставку да су штете независне случајне величине са истом расподјелом.



Слика 4: Подаци о штетама - danish



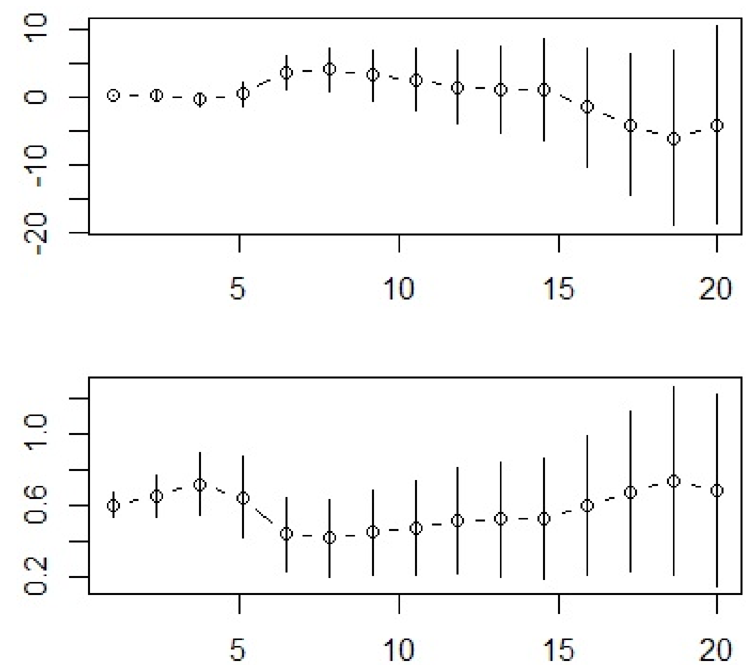
Слика 5: Хистограм логаритмованих података о штетама

На слици 4 имамо приказ висина штета, гдје видимо да се појављивало неколико екстрема и приближно вријеме тих догађаја, док на слици 5 имамо хистограм логаритмованих података гдје можемо да примијетимо да подаци имају тешки десни реп.

Ако желимо да моделујемо стохастичко понашање максимума ових штета, користимо основне расподјеле екстремних вриједности (Фрешеову, Гумбелову или Вејбулову), а ако желимо да моделујемо штете које имају услов прекорачења високих вриједности (што ће у овом примјеру и бити урађено), користимо генералисану Паретову расподјелу.

Просјечан износ штете је 3,38 милиона Данских круна. За процјену расподјеле екстрема кориштен је пакет *extRemes*.

Прије свега, потребно је изабрати одговарајући праг u, а слика 6 нам може помоћи у томе. На горњој слици су приказане вриједности параметра скалирања β у зависности од прага, а на доњој вриједности параметра облика γ, оба параметра са 95% интервалом повјерења. Потребно је изабрати праг такав да су вриједности параметара стабилне, а да остане довољно података, односно вриједности изнад прага.



Слика 6: Оцјена параметара генералисане Паретове расподјеле у зависности од прага

Већ за вриједности изнад 10 интервали повјерења се повећавају, тако да ћемо изабрати $u = 10$, односно штету у износу од 1.289.807,4 евра. Постоји $N_u = 109$ штета (од њих 2157) које су веће од изабраног прага, што јесте оптималан број података.

Функција *fevd* из поменутог пакета користи методу максималне вјеродостојности и оцјењује параметар облика γ и параметар скалирања β из генералисане Паретове расподјеле. Добијају се оцјене $\gamma = 0,5$ са стандардном грешком $0,14$ и $\beta = 6,98$. са стандардном грешком $1,11$. Тј. добијамо сљедећи модел:

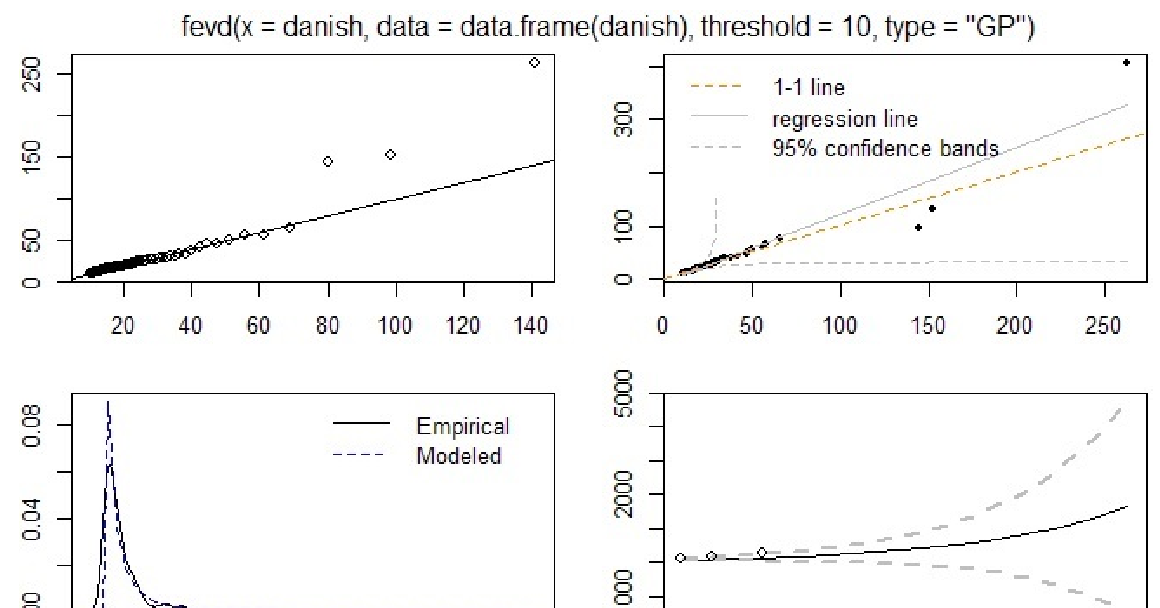
$$W(x) = 1 - \left(1 + \frac{0,5(x-10)}{6,98}\right)^{-1/0,5},$$

односно

$$W(x) = 1 - (0,07x + 0,28)^{-2}.$$

Примјера ради, сада је могуће помоћу овог модела израчунати вјероватноће да штета од пожара пређе вриједности од 20, 40, 60 или 100 милиона Данских круна, ако претпостављамо да је штета екстремна, односно да је већа од прага од 10 милиона. Добијају се, редом, вјероватноће од око $34,9\%, 10\%, 4,7\%$ и $1,8\%$.

Дакле, ово је модел који одговара искључиво репу расподјеле штета. На слици испод имамо дијагностички приказ модела, сви подаци који су приказани су подаци већи од прага u.



Слика 7: Дијагностички приказ модела који описује штете од пожара. Имамо график емпиријских квантила у зависности од квантила модела (горе лијево), квантили узорка из модела - емпиријски квантили са 95%−тним интервалом повјерења (горе десно), график густине емпиријских података и модела (доле лијево) и период понављања у годинама, тј. на x оси је процјена очекиваног броја година који прође до избијања пожара који направи штету y (доле десно).

5 Закључак

Циљ овог рада био је упознавање и разматрање основних резултата теорије екстремних вриједности, што је веома значајно, између осталог, и у области осигурања, с обзиром на утицај који неочекивано велике штете имају на портфолио осигуравајућег друштва. Како очекивана вриједност укупне суме износа штета директно утиче на премију коју треба одредити при склапању уговора о осигурању, актуари морају да познају како могу да се понашају вриједности одштетних захтјева.

Акценат је свакако на три типа расподјела (Гумбелова, Фрешеова и Вејбулова расподјела) које су могуће као граничне расподјеле линеарно нормираног максимума једнако расподијељених независних случајних величина. Улога теореме о екстремалним типовима (Фишер-Типет-Гнеденко теореме) за максимуме је слична улози коју централна гранична теорема (ЦГТ) има када је ријеч о просјечним вриједностима, осим што ЦГТ може да се примијени на просјечну вриједност узорка из било које расподјеле са коначном дисперзијом, док теорема о екстремалним типовима важи само при условима да расподјела нормализованог максимума конвергира, што не мора да буде (а ако конвергира, та гранична вриједност је управо једна од три типа расподјела).

Осим споменуте три расподјеле, велики значај има и генералисана Паретова расподјела, посебно у (ре)осигурању, с обзиром да на основу Пикандс-Балкема-де Хан теореме (тзв. друге теореме о екстремалним типовима) знамо да је то асимптотска расподјела репа непознате расподјеле неке случајне промјенљиве. За разлику од прве теореме, у овом случају су од интереса вриједности изнад изабраног прага. Од великог је значаја информација о периоду понављања екстремних догађаја у години коју нам обезбјеђује ова теорија, јер може да се користи при рачунању просјечних годишњих трошкова осигуравајућег друштва.

Ти главни резултати теорије екстремних вриједности су основа за све обимнији развој ове области статистике, а исход тога је у великој мјери олакшано предвиђање будућих прилика и креирање планова и стратегија.

Литература

[1] M.R. Leadbetter, G. Lindgren, H. Rootzen (2015): *Extremes and related properties*; (2015);

[2] A.A. Balkema Laurens de Haan (1972): *On R. von Mises' condition for the domain of attraction of* $exp(-e^{-x})$

[3] Др Павле Младеновић (2014): *Елементи актуарске математике*, Математички факултет, Београд

[4] Paul Embrechts, Sidney I. Resnick, Genady Samorodnitsky (1999): *Extreme Value Theory as a Risk Management Tool*, Cornell University

[5] R. -D. Reiss, M. Thomas (2007): *Statistical Analysis of Extreme Values*

[6] Laurens de Haan, Ana Ferreira (2006): *Extreme Value Theory, An Introduction*

[7] Alecander J. McNeil (1996): *Estimating the Tails of Loss Severity Distributions using Extreme Value Theory*, Department Mathematik, ET H Zentrum, Zürich

[8] Joan del Castillo, Maria Padilla (2015): *Modeling Extreme Values by the Residual Coefficient of Variation* (`https://arxiv.org/pdf/1510.00179.pdf`)

[9] Eric Gilleland, Richard W. Katz (2016): *extRemes 2.0: An Extreme Value Analysis Package in R* (`https://arxiv.org/pdf/1510.00179.pdf`)