

UNIVERZA V LJUBLJANI
EKONOMSKA FAKULTETA

MAGISTRSKO DELO

**APLIKACIJA TEORIJE EKSTREMNIH VREDNOSTI NA DONOSIH
DELNICE IZBRANEGA PODJETJA**

Ljubljana, november 2013

PETRA ČEPON

IZJAVA O AVTORSTVU

Spodaj podpisana Petra Čepon, študentka Ekonomske fakultete Univerze v Ljubljani, izjavljam, da sem avtorica magistrskega dela z naslovom Aplikacija teorije ekstremnih vrednosti na donosih delnice izbranega podjetja, pripravljene v sodelovanju s svetovalcem dr. Alešom Ahčanom.

Izrecno izjavljam, da v skladu z določili Zakona o avtorskih in sorodnih pravicah (Ur. l. RS, št. 21/1995 s spremembami) dovolim objavo magistrskega dela na fakultetnih spletnih straneh.

S svojim podpisom zagotavljam, da

- je predloženo besedilo rezultat izključno mojega lastnega raziskovalnega dela;
- je predloženo besedilo jezikovno korektno in tehnično pripravljeno v skladu z Navodili za izdelavo zaključnih nalog Ekonomske fakultete Univerze v Ljubljani, kar pomeni, da sem
 - poskrbela, da so dela in mnenja drugih avtorjev oziroma avtoric, ki jih uporabljam v magistrskem delu, citirana oziroma navedena v skladu z Navodili za izdelavo zaključnih nalog Ekonomske fakultete Univerze v Ljubljani, in
 - pridobila vsa dovoljenja za uporabo avtorskih del, ki so v celoti (v pisni ali grafični obliki) uporabljena v tekstu, in sem to v besedilu tudi jasno zapisala;
- se zavedam, da je plagiatstvo – predstavljanje tujih del (v pisni ali grafični obliki) kot mojih lastnih – kaznivo po Zakonu o avtorskih in sorodnih pravicah (Ur. l. RS, št. 21/1995 s spremembami);
- se zavedam posledic, ki bi jih na osnovi predloženega magistrskega dela dokazano plagiatstvo lahko predstavljalo za moj status na Ekonomski fakulteti Univerze v Ljubljani v skladu z relevantnim pravilnikom.

V Ljubljani, dne 12.11.2013

Podpis avtorice: _____

KAZALO

UVOD	1
1 ZGODOVINA	4
2 SPLOŠNA PRISTOPA K TEORIJI EKSTREMNIH VREDNOSTI	6
3 PREDSTAVITEV RAZLIČNIH PRISTOPOV K TEORIJI EKSTREMNIH VREDNOSTI	7
3.1 Prvi parametrični pristop	8
3.1.1 Fisher-Tipettov izrek	8
3.1.2 Posplošena porazdelitev ekstremnih vrednosti	9
3.1.3 Ocenjevanje GEV-porazdelitve po metodi največjega verjetja	10
3.2 Drugi parametrični pristop	11
3.2.1 Funkcija presežka.....	11
3.2.2 Posplošena Paretova porazdelitev	12
3.2.3 Izrek Pickands-Balkema-de Haan	13
3.2.4 Ocenjevanje GPD-porazdelitve po metodi največjega verjetja	13
3.2.5 Ocenjevanje porazdelitve repa	13
3.3 Neparometrični pristop in cenilka repnega indeksa	14
3.3.1 Hillova cenilka	14
3.3.2 Pickandsova cenilka	15
4 MERJENJE TVEGANJA	15
4.1 VaR	15
4.1.1 Uporaba VaR v EVT (GPD-porazdelitev)	17
4.2 ES	17
4.2.1 Uporaba ES v EVT (GPD-porazdelitev).....	18
4.3 Stopnja poplačila	18
4.3.1 Uporaba stopnje poplačila v EVT (GEV-porazdelitev)	19
4.3.2 Uporaba stopnje poplačila v EVT (GPD- porazdelitev)	19
5 GRAFIČNA ANALIZA PODATKOV	19
5.1 Graf kvantilov	19
5.2 MEF in ME-Plot	20
5.3 Hillov graf	20
6 PROCEDURA UPORABE TER PREDNOSTI IN SLABOSTI TEORIJE EKSTREMNIH VREDNOSTI TER NJENIH MODELOV	21
6.1 Procedura uporabe	21
6.2 Prednosti in slabosti EVT in njenih modelov	22
7 ANALIZA PODATKOV	23
7.1 Izbira nabora podatkov	23
7.2 Osnovni ali logaritmirani podatki	24
7.3 Podatki – delnica IBM	25
7.4 Razvoj rekordov	27
7.5 Graf kvantilov celotnega vzorca	28

7.6 Prvi parametrični pristop	29
7.6.1 GEV-porazdelitev in metoda maksimumov skupin podatkov.....	30
7.6.2 Graf kvantilov na podlagi prvega parametričnega pristopa.....	36
7.6.3 Stopnja poplačila na podlagi prvega parametričnega pristopa	37
7.7 Drugi parametrični pristop	40
7.7.1 Paretova porazdelitev	40
7.7.2 Graf kvantilov na podlagi drugega parametričnega pristopa.....	41
7.7.3 MEF na podlagi drugega parametričnega pristopa.....	42
7.7.4 Računski pristop določanja mejne vrednosti	43
7.7.5 Izbira mejne vrednosti	43
7.7.6 Ocenjevanje GPD-porazdelitve po metodi največjega verjetja	46
7.7.7 Merjenje tveganja na podlagi drugega parametričnega pristopa	47
7.8 Neparometrični pristop.....	49
7.8.1 Hillova neparometrična ocena repnega indeksa	49
SKLEP.....	51
LITERATURA IN VIRI.....	55
PRILOGE	
KAZALO SLIK	
Slika 1: Model maksimumov skupin podatkov (levo) in model presejanja mejne vrednosti (desno).....	7
Slika 2: Primer Fréchetove ($\alpha = 1,5$), Weibullove ($\alpha = -1,5$) in Gumbelove porazdelitvene funkcije ($\alpha = 0$)	9
Slika 3: Porazdelitvena funkcija F in porazdelitvena funkcija presežka Fu	12
Slika 4: Posplošena Paretova porazdelitev $H\xi, \sigma$ za negativne, pozitivne in ničelne vrednosti ξ , kjer je σ fiksiran na vrednost 1.....	13
Slika 5: Logaritemski dnevni donosi IBM-ove delnice v %, za obdobje 1962–2012.....	25
Slika 6: Logaritemski negativni (levo) in pozitivni (desno) dnevni donosi IBM-ove delnice v %, za obdobje 1962–2012	26
Slika 7: Histogram negativnih dnevnih logaritemskih donosov IBM-ove delnice.....	27
Slika 8: Grafični prikaz razvoja prvih desetih rekordov delnice IBM	28
Slika 9: Graf kvantilov celotnega opazovanega vzorca donosov delnice IBM, enote v %	29
Slika 10: Primer kumulativne porazdelitvene funkcije (levo) in verjetnostne gostote (desno) za vse tri GEV-porazdelitve	30
Slika 11: Maksimalne vrednosti negativnih donosov delnice IBM v % za različno izbrana obdobja: mesečno obdobje (levo zgoraj), četrletno (desno zgoraj), polletno (levo spodaj) in letno obdobje (desno spodaj)	31
Slika 12: Histogram maksimalnih izgub donosov delnice IBM na podlagi izbire mesečnih skupin podatkov	32
Slika 13: Razsevni diagram residualov in graf kvantilov, pridobljenih s prileganjem GEV-porazdelitve opazovanim podatkom.....	33

Slika 14: Razsevni diagrami residualov: mesečno obdobje (levo zgoraj), četrletno obdobje (levo spodaj), polletno obdobje (desno zgoraj) in letno obdobje (desno spodaj)	35
Slika 15: Graf kvantilov: mesečno obdobje (levo zgoraj), četrletno obdobje (levo spodaj), polletno obdobje (desno zgoraj) in letno obdobje (desno spodaj)	36
Slika 16: Graf kvantilov maksimalnih vrednosti, pridobljenih z izbiro mesečnih skupin podatkov, kjer je za teoretično porazdelitev izbrana Gumbelova porazdelitev	37
Slika 17: Letna stopnja izgube donosa za mesečne maksimume	38
Slika 18: 30-letna stopnja poplačila za mesečne (zgoraj levo), četrletne (spodaj levo), polletne (zgoraj desno) in letne (spodaj desno) maksimume	39
Slika 19: Kumulativna porazdelitvena funkcija (levo) in funkcija gostote verjetnosti (desno) za Paretovo porazdelitev tipa I in tipa II ter eksponentno porazdelitev	41
Slika 20: Graf kvantilov za izbrano mejno vrednost $u = 0,1$ (levo) in mejno vrednost $u = 1$ (desno)	42
Slika 21: Graf povprečnih odmikov za negativne donose delnice IBM	42
Slika 22: Prikaz izbire mejne vrednosti u , nad katero leži 5 % vseh negativnih donosov	43
Slika 23: Diagnostični grafi za GPD-porazdelitev, ki smo jo prilegali podatkom negativnih donosov IBM-ove delnice pri izbrani mejni vrednosti $u = 0,1$: porazdelitvena funkcija presežkov (levo zgoraj), rep porazdelitvene funkcije (levo spodaj), razsevni diagram residualov (desno zgoraj) in graf kvantilov (desno spodaj)	44
Slika 24: Diagnostični grafi za GPD-porazdelitev, ki smo jo prilegali podatkom negativnih donosov IBM-ove delnice pri izbrani mejni vrednosti $u = 1,03$: porazdelitvena funkcija presežkov (levo zgoraj), rep porazdelitvene funkcije (levo spodaj), razsevni diagram residualov (desno zgoraj) in graf kvantilov (desno spodaj)	45
Slika 25: Ocenjen oblikovni parameter ξ , ko se u in pripadajoče število presežkov spreminjata	46
Slika 26: Asimptotični 95 % interval zaupanja za Var_{99} in ES_{99} , ki temelji na GPD-porazdelitvi z mejno vrednostjo $u = 1,03$, katero smo prilegali negativnim donosom IBM-ove delnice	48
Slika 27: Ocene kvantilov reda 0,99 in 0,999 ter 0,9999 kot funkcija mejnih vrednosti ali kot funkcija presežkov	49
Slika 28: Hillova cenilka ξ_{Hill} za vsa opazovanja (levo) in za prvih 300 vrstilnih statistik (desno)	50
Slika 29: Hillova cenilka α_{Hill} za vsa opazovanja (levo) in za prvih 300 vrstilnih statistik (desno)	50

KAZALO TABEL

Tabela 1: Osnovne statistike za donos delnice IBM	26
Tabela 2: Osnovne statistike za negativen donos delnice IBM	27

<i>Tabela 3: Prikaz prvih desetih rekordov, njihova vrednost, število opazovanja, ko se je rekord zgodil ter matematično upanje in standardni odklon opazovanega rekorda v %.....</i>	<i>28</i>
<i>Tabela 4: Ocenjeni parametri GEV-porazdelitve za mesečne skupine podatkov</i>	<i>31</i>
<i>Tabela 5: Ocena parametrov za vsa obdobja (standardne napake v oklepajih).....</i>	<i>34</i>
<i>Tabela 6: Enoletna stopnja izgube donosa za mesečno izbrane podatke.....</i>	<i>38</i>
<i>Tabela 7: 30-letna stopnja izgube donosa za negativne donose IBM-ove delnice.....</i>	<i>39</i>
<i>Tabela 8: Ocenjeni parametri GPD-porazdelitve pri mejni vrednosti $u=1,03$.....</i>	<i>46</i>
<i>Tabela 9: Oceni VaR in ES, ki temeljita na GPD-aproksimaciji</i>	<i>47</i>
<i>Tabela 10: Oceni VaR in ES, ki temeljita na normalni porazdelitvi</i>	<i>47</i>
<i>Tabela 11: Rezultati Hilllove cenilke s standardnimi napakami za donose IBM-a</i>	<i>50</i>

UVOD

S teorijo ekstremnih vrednosti se srečujemo v vsakdanjem življenju, le da se morda tega ne zavedamo. Zagotovo smo bili vsi priča vsaj enemu ekstremnemu dogodku v življenju. Dovolj je, če spremljamo klimatske spremembe: obilne količine padavin, ki jih spremljajo poplave in plazovi, nenadno naraščanje morja, vedno večja frekvenca cunamijev po svetu in povečana tektonska aktivnost ter možnost potresov. Omenjeni pojavi imajo v primeru sklenitve zavarovanj proti takim dogodkom tudi neposredne posledice v zavarovalniškem sektorju, saj morajo zavarovalnice v zelo kratkem času kriti visoke odškodnine.

Teorija ekstremnih vrednosti ima visoko aplikativno vrednost na različnih področjih. Vpliv teorije ekstremnih vrednosti je mogoče zaznati tudi med športniki, saj želijo doseči izjemne športne dosežke in preseči pretekle rekorde. Trenutno najbolj pereče teme, ki so nas zaznamovale v zadnjem času, so precejšnje nestabilnosti in visoke izgube na finančnih trgih, finančne težave nekaterih članic EU ter splošna svetovna finančna kriza, ki so prav tako lahko predmet teorije ekstremnih vrednosti. Omenjeno področje - predvsem visoke izgube na finančnih trgih - bo osrednja tema, ki jo bomo podrobneje pogledali v empiričnem delu magistrske naloge.

Vsi zgoraj omenjeni dogodki so ekstremni dogodki, saj je po mnenju Rocca (2011) ekstremni dogodek tisti dogodek, ki se zgodi redko in z nizko verjetnostjo, vendar povzroči visoko škodo. Zgodi se, ko tveganje zavzame vrednost v repu porazdelitve. V klasični analizi podatkov so ekstremni dogodki po navadi uprizorjeni kot osamelci ali pa jih celo izključimo iz opazovanja. Torej podatke prilagajamo, da ustrezajo zelenemu modelu. Če iščemo le oceno vsakdanjih dogodkov, nam je izpuščeni ekstremni podatki ne bodo veliko spremenili. Če pa se vprašamo po dogodkih, ki se ne zgodijo zelo pogosto, in se obrnemo na teorijo ekstremnih vrednosti, potem bodo ti ekstremni podatki ključni faktor.

Večina podatkov iz vsakdanjega življenja ponazarja normalno porazdelitev oziroma njene sorodne porazdelitve, ki so natančno definirane in lahko izračunljive. Ko pa se sprašujemo po ekstremnih dogodkih, se le-ti nahajajo v repih porazdelitev. V splošnih raziskavah je prikazano, da so ti repi debelejši kot jih predvideva normalna porazdelitev. To je vodilo k številnim kritikam že obstoječih sistemov za upravljanje s tveganji in zato spodbudilo k razmišljanju in iskanju primernejših metodologij, ki bi se bile sposobne spopadati z ekstremnimi dogodki in njihovimi posledicami.

Zaradi pomembnosti ekstremnih dogodkov, katerih frekvenca je, kot že omenjeno, nizka, višina škode pa izredno visoka, je ključno, da se teoretična verjetnostna porazdelitev podatkom dobro prilega ravno v repu. Podatki, ki so porazdeljeni po debelem repu, imajo visoko stopnjo sploščenosti in asimetričnosti (Liang, 1999). Teorija ekstremnih vrednosti je po mnenju Bensalah (2000) le orodje, ki se ukvarja ravno z repi asimetričnih porazdelitev in nam omogoča najti najboljšo možno oceno repnega območja neke porazdelitve.

Struktura magistrskega dela je oblikovana tako, da omogoča jasen pregled obravnavanih tem. Delo je strukturirano v dva večja poglavja. V prvem, teoretičnem poglavju, obravnavamo teorijo ekstremnih vrednosti, pri čemer v uvodnem delu opredelimo definicijo ekstremnega dogodka in pojasnimo, kako je z njim povezana teorija ekstremnih vrednosti. Prav tako razložimo določene predpostavke, ki se uporabljajo v klasični statistični teoriji in so večkrat kršene. V zadnjih stotih letih je teorija ekstremnih vrednosti doživela hiter razvoj in uspešno aplikacijo na več področjih. Del tega poglavja je zato namenjen tudi zgodovini in razvoju te teori-

je. V nadaljevanju opišemo ključna pristopa in izreka v teoriji ekstremnih vrednosti. Prvi pristop se imenuje model maksimumov skupin podatkov in analizira ekstremne dogodke v podatkih, ti pa so združeni v skupine glede na pojavljanje v enakih časovnih obdobjih. Drugi pristop se imenuje model preseganja mejne vrednosti in analizira le opazovanja, ki presegajo vnaprej izbrano mejno vrednost. Iz teh dveh pristopov sta se razvila tudi glavna izreka in možne porazdelitve ekstremnih vrednosti, ki jih opišemo v tem poglavju. Zgoraj opisana pristopa sta parametrična pristopa k teoriji ekstremnih vrednosti, opišemo pa tudi neparametrični pristop k teoriji ekstremnih vrednosti. Pri slednjem je poudarek na Hillovi cenilki. Najbolj pogosto vprašanje, ki zadeva obvladovanje tveganja, je ocena ekstremnih kvantilov, saj nas zanima, kolikšnemu tveganju smo izpostavljeni oziroma koliko "premoženja" lahko izgubimo. Tipičen primer merjenj tveganja je angl. *Value at Risk* (VaR), drugi, manj pogosto uporabljeni metodi pa sta pričakovani izpad (angl. *Expected Shortfall* oziroma ES) in stopnja poplačila (angl. *Return Level*), ki ju prav tako opišemo. Poglavje končamo s predstavitvijo grafičnih analiz v teoriji ekstremnih vrednosti, saj lahko z njihovo pomočjo grobo ocenimo, kolikšna je mejna vrednost, ki jo iščemo, in posledično tudi, kolikšen je repni indeks, ki določa porazdelitev.

V drugem poglavju magistrskega dela se osredotočimo na finančne, predvsem delniške trge, ki so jih v preteklosti zaznamovale nagle spremembe stanj v gospodarstvu. Visoka volatilitnost delniških tečajev, ki označuje dovzetnost delnice za tečajna nihanja, predstavlja odličen primer teorije ekstremnih vrednosti. Glede na podkonstrukte zgoraj opisane teorije analiziramo donose delnice podjetja IBM (angl. *International Business Machines*) v obdobju od začetka leta 1962 do konca leta 2012. Na podlagi rezultatov, pridobljenih z grafičnimi in teoretičnimi metodami, upravičimo uporabo teorije ekstremnih vrednosti na teh podatkih. Dodano vrednost skušamo doseči s predstavitvijo obsežne analize v primeru, ko se podatki porazdeljujejo s porazdelitvijo debelega repa – donosi delnice IBM. Ker pa nas v praksi vedno zanima možna izguba "premoženja", je prava dodana vrednost v analizi merjenja tveganj. VaR je v zadnjih časih postal zelo uporabljena in razširjena mera tveganja, vendar predpostavlja normalne tržne pogoje. Na opazovanih podatkih pokažemo podcenjenost mer tveganja (VaR in ES) za predpostavljeno prileganje normalne porazdelitve v primerjavi s prileganjem porazdelitve debelega repa, ki je ustrezna porazdelitev za donose IBM-ove delnice.

Osnovni namen magistrskega dela je opozoriti bralca na pomen teorije ekstremnih vrednosti in njeno redko uporabo. Menimo, da kljub visoki aplikativni vrednosti na več področjih, ta teorija tako v svetu kot tudi v Sloveniji ni dovolj poznana. Na splošno teorijo ekstremnih vrednosti največ uporabljajo v meteorologiji in hidrologiji. Tudi naša, slovenska Agencija Republike Slovenije za okolje (v nadaljevanju ARSO) se v svojih raziskavah opira na zaključke teorije ekstremnih vrednosti, predvsem na metodo, ki ji rečemo metoda vrednosti nad izbranim pragom (angl. *Peak Over Threshold Method*). V Sloveniji sicer nismo zaznali raziskav s področja teorije ekstremnih vrednosti, razen tistih, ki jih opravlja ARSO. Menimo, da je ta teorija zelo uporabna tudi pri raziskovanju podatkov v finančnem in zavarovalniškem sektorju, zato se bomo v dani magistrski nalogi osredotočili ravno na finančni trg in izgubo delnic, ki smo ji bili priča predvsem v zadnjih letih oziroma desetletjih.

Temeljni cilj tega magistrskega dela je preko širšega teoretičnega pregleda teorije ekstremnih vrednosti, njenih zaključkov in ocenjenih mer tveganja uporabiti in analizirati zgoraj opisano na primeru donosov delnice podjetja IBM. S tem je preko praktičnega primera predstavljena teorija ekstremnih vrednosti in njegova uporaba v praksi.

Pomožni cilji, ki jih želimo doseči:

- predstaviti zgodovinski razvoj teorije ekstremnih vrednosti,
- opisati in predstaviti dva različna pristopa k teoriji ekstremnih vrednosti,
- predstaviti ključna izreka, ki se nanašata na teorijo ekstremnih vrednosti,
- opisati in definirati nekaj možnih porazdelitev v teoriji ekstremnih vrednosti, ki so hkrati tudi zaključki zgoraj omenjenih izrekov,
- oceniti glavni porazdelitvi teorije ekstremnih vrednosti z uporabo metode največjega verjetja,
- predstaviti tako parametrični kot neparametrični pristop k teoriji ekstremnih vrednosti,
- opisati glavne mere tveganja in njihove ocene,
- predstaviti nekaj grafičnih analiz podatkov, značilnih predvsem za ugotavljanje ekstremnih vrednosti,
- večino zgoraj omenjenih ciljev aplicirati na praktičnem primeru, zapisati in uprizoriti ugotovitve, njihovo povezavo z zgoraj opisano teorijo in podati zaključke na podlagi teh ugotovitev.

S tem delom želimo med drugim povzeti tudi izkušnje drugih avtorjev in raziskovalcev na tem področju. Današnji krizni časi zahtevajo, da nadgradimo obstoječe znanje o ekstremnih dogodkih in teoriji ekstremnih vrednosti ter naše izsledke predstavimo tudi širši javnosti. Upamo, da s tem pritegnemo k obravnavani tematiki še kakšnega raziskovalca z drugega področja, z namenom, da bolje oceni in predvidi stanje, ki se v najslabšem primeru lahko zgodi. Skupaj z raziskavami še drugih raziskovalcev bomo naredili korak naprej v razumevanju vedno bolj pogostih ekstremnih dogodkov in tako lažje predvideli prihajajoče krizno stanje ali vrednost.

V magistrski nalogi bomo **testirali hipotezo**, da je uporaba teorije ekstremnih vrednosti upravičena na donosih delnice izbranega podjetja, podjetja IBM. Zapisali bomo dobljene rezultate in podali ustrezne zaključke.

Metode dela, ki jih uporabljamo v magistrski nalogi, so teoretično-empirične, temeljijo predvsem na proučevanju teoretične podlage, ki se je razvijala skozi zgodovino, na uporabi tako teoretičnih kot praktičnih spoznanj ter spoznavanju in učenju uporabe teorije ekstremnih vrednosti v statističnem programskem jeziku R. R je statistično okolje za analizo in vizualizacijo podatkov, ki ga je razvil John Chambers s sodelavci (R-project, 2012). Pri pisanju in izdelavi magistrskega dela se sklicujemo na strokovno literaturo, prispevke in članke predvsem tujih avtorjev, ki so prispevali ključna spoznanja k obravnavani temi. Osnovno gradivo, na katerega se sklicujemo v teoretičnem delu magistrske naloge, je obsežna knjiga z naslovom *Modelling extremal events for insurance and finance*, ki so jo leta 1997 napisali Embrechts, Klüppelberg in Mikosch. V tem delu so avtorji strnili vso do tedaj znano teorijo o ekstremnih vrednostih tako z matematičnega področja, področja statistike in verjetnosti kot tudi s področja financ, zavarovalništva in drugih praktičnih primerov. V drugem, empiričnem delu magistrskega dela vso zgoraj navedeno teorijo apliciramo in analiziramo na praktičnem primeru IBM-ove delnice.

Kot že omenjeno, je teorija ekstremnih vrednosti področje, ki je v slovenskem prostoru dokaj novo in nepoznano, saj ga v praksi le redko zasledimo. Empirični del magistrskega dela je tako popolnoma samostojno in neodvisno delo, ki ni bilo še nikjer obravnavano.

V tej magistrski nalogi se soočamo le z manjšimi **omejitvami**. Te so predvsem matematične narave. Skoraj na vsakem matematičnem področju se soočamo z določenimi predpostavkami in tudi teorija ekstremnih vrednosti ni izjema. Te predpostavke se nahajajo v obeh parametričnih modelih, zato se pojavljajo tako v teoretičnem kot tudi empiričnem delu magistrske naloge. Gre za predpostavko o neodvisnosti in enaki porazdeljenosti podatkov, ki v praksi ne velja vedno. Vendar je to predpostavko v primeru kršenja možno s poznavanjem področja ekonometrije odpraviti. Druga matematična predpostavka pa se nanaša na obstoj limite (podrobneje razloženo v Poglavju 6). Določene omejitve, ki se nanašajo na neustrezen zajem obsega ekstremnih vrednosti v prvem parametričnem pristopu, so odpravljene z razvojem drugega parametričnega pristopa, kjer so zajete vse ekstremne vrednosti nad vnaprej izbrano mejno vrednostjo. Pri velikosti zajema ekstremnih podatkov in izbiri mejne vrednosti pa se soočamo z omejitvijo subjektivne ocene slednjih.

1 ZGODOVINA

Začetki teorije ekstremnih vrednosti (angl. *EVT – Extreme Value Theory*) segajo v leto 1928, začetnik te teorije pa je bil Leonard Henry Caleb Tippett (1902–1985). Tippett je bil zaposlen v združenju raziskovalcev bombažne industrije v Veliki Britaniji (angl. *British Cotton Industry Research Association*) kot raziskovalni delavec, ki je poskušal dognati, kako bi naredil niti bombaža močnejše. Bil je angleški statistik, ki je večino časa svoje zaposlitve posvetil povezovanju tehnologije in statistike. V procesu raziskovanja je kmalu ugotovil, da so bombažne niti tako močne, kot je močen njihov najšibkejši člen oziroma vlakno. Tako je bilo treba nujno razviti novo verjetnostno teorijo, saj takratna klasična statistična teorija (angl. *Classical Statistical Theory*) ni znala odgovoriti na enostavna vprašanja, če so se pojavili ekstremni dogodki. S pomočjo Ronalda Aylmerja Fisherja (1890–1962) je Tippett leta 1928 postavil temelje sedaj imenovane teorije ekstremnih vrednosti in našel tri asimptotične limite, ki opisujejo porazdelitev ekstremov. Te tri porazdelitve so dobile ime po znanih matematikih in znanstvenikih.

Prva izmed teh treh porazdelitev se imenuje Gumbelova porazdelitev. Ime je dobila po nemškem matematiku Emilu Juliusu Gumbelu (1891–1966). Gumbel je uporabil EVT v realnih svetovnih problematikah, predvsem v inženirstvu, meteoroloških in hidroloških fenomenih. Zanimal ga je na primer največji letni pretok rek: "It seems that the rivers know the theory. It only remains to convince the engineers of the validity of this analysis." Tako je Gumbel prvič opisal svojo teorijo in tudi porazdelitev leta 1958 v knjigi *Statistics of Extremes*. Gumbelova porazdelitev se uporablja za prikaz porazdelitve ekstremnih vrednosti in nakazuje, da bi morali zgodovinski opazovani podatki imeti normalno ali eksponentno porazdelitev, če bi želeli pokazati, da se njihove maksimalne vrednosti porazdeljujejo po Gumbelovi porazdelitvi. Do sedaj ima ta porazdelitev največjo vlogo v modeliranju ekstremnih vrednosti na področjih, ki se ukvarjajo s poplavami, količino dežja in potresi, uporablja pa se tudi v gradbeništvu.

Druga izmed teh treh porazdelitev se imenuje Fréchetova porazdelitev, ime pa je dobila po francoskem matematiku Mauricu Renéju Fréchetu (1878–1973). Fréchet je leta 1927 napisal članek, povezan s porazdelitvami debelih repov. Podobno kot Gumbelova porazdelitev se tudi Fréchetova porazdelitev uporablja za prikaz porazdelitve ekstremnih vrednosti, kjer morajo imeti zgodovinski opazovani podatki debel rep, torej morajo imeti na primer Cauchyjevo ali Studentovo t-porazdelitev, njene maksimalne vrednosti pa se bodo potem porazdeljevale po Fréchetovi porazdelitvi. Danes se Fréchetova porazdelitev uporablja predvsem v modeliranju finančnih porazdelitev, saj je splošno znano, da tržni donosi največkrat opisujejo porazdelitev debelega repa.

Zadnja, tretja porazdelitev pa se imenuje Weibullova porazdelitev, ki je dobila ime po švedskem inženirju in znanstveniku Waloddiju Weibullu (1887–1979). Weibull se je ukvarjal predvsem z močjo in natezno trdnostjo materialov ter tako leta 1939 objavil članek, kjer je prvič opisal "Weibullovo" porazdelitev v verjetnostni teoriji in statistiki. V svojem življenju je Weibull napisal veliko člankov, med katerimi je najbolj znan članek iz leta 1951, ki ga je predstavil ASME (angl. *American Society of Mechanical Engineers*), v njem pa je izpostavil sedem različnih primerov, v katerih je uporabil omenjeno porazdelitev, za katero je trdil, da je uporabna pri različnih problemih: od testiranja moči jekla do višine odraslih moških. Kot v zgornjih dveh primerih se Weibullova porazdelitev uporablja za prikaz porazdelitve ekstremnih vrednosti, kjer morajo imeti opazovani podatki suh rep, na primer enakomerno ali beta porazdelitev, njihov maksimum pa se bo potem porazdeljeval po Weibullovi porazdelitvi. Danes se Weibullova porazdelitev uporablja predvsem za analize podatkov o življenjski dobi produktov v inženirstvu.

Odločilno dejanje pa je leta 1943 storil ruski matematik Boris Vladimirovich Gnedenko (1912–1995), ki je združil vse dotodanje znanje o ekstremnih dogodkih. Skupaj z zelo znanim matematikom Andrejem Nikolajevičem Kolmogorovom je tako zapisal prvi izrek v teoriji ekstremnih vrednosti oziroma Izrek Fisher-Tippett-Gnedenko (krajše: Fisher-Tippetov izrek).

Potem ko sta Fisher in Tippett opisala osnovne tri tipe porazdelitev ekstremnih vrednosti, je Richard von Mises (1883–1953) leta 1936 prvi predlagal uprizoritev teh treh vrst porazdelitev EVT v eno samo, ki se imenuje posplošena porazdelitev ekstremnih vrednosti (angl. *Generalized Extreme Value Distribution*). Kasneje, leta 1955, je tudi Jenkinson predstavil to enostavnejšo matematično obliko – GEV-porazdelitev, zato se posplošeni porazdelitvi ekstremnih vrednosti (GEV-porazdelitvi) večkrat reče tudi Jenkinson-von Mises parametrizacija.

Prvemu izreku v teoriji ekstremnih vrednosti je kmalu sledil tudi drugi. Imenoval se je Izrek Pickands-Balkema-de Haan, ki je za visoke mejne vrednosti povezal limitno obliko funkcije presežka s posplošeno Paretovo porazdelitvijo (angl. *Generalized Pareto Distribution*). Prve ideje tega izreka so se pojavile v delih avtorjev Balkeme in de Haana (1974) ter Pickandsa (1975).

Prvi, ki so se začeli ukvarjati s področjem ekstremnih vrednosti, so bili statistiki. Kasneje, v zgodnjih 50. letih 20. stoletja, so se jim pridružili še hidrologi, ki so se soočali z obsežnimi poplavami, zemeljskimi plazovi in nenadnim naraščanjem morja. V 80. letih 20. stoletja so jim sledili aktuarji, ki so se ukvarjali z modeliranjem debelih repov neživljenjskih zavarovalniških zahtevkov. Sredi 90. let 20. stoletja so se EVT posvetili še analitiki časovnih finančnih serij, z uveljavitvijo modela Value-at-Risk pa so proti koncu 90. let 20. stoletja sledili še kvantitativni upravljalci s tveganji (angl. *risk managers*).

V začetku 20. stoletja je Value-at-Risk (VaR) postala standardna mera za ocenjevanje tveganosti, saj jo je kot mero tveganja predlagal tudi Baselski komite za nadzor bank (angl. *Basel Committee on Banking Supervision*). Obstoječe standardne metode za izračun VaR-a predpostavljajo normalno porazdeljene podatke, ti normalno porazdeljeni podatki pa podcenjujejo tveganja visokih kvantilov in verjetnosti dogodkov v repih porazdelitev. Ta predpostavka je v primeru časovnih finančnih vrst največkrat kršena, saj imajo podatki porazdelitev debelega repa. Tako se je pričela uporabljati teorija ekstremnih vrednosti, ki je upoštevala porazdelitve repov in ocenila VaR bolj realistično.

2 SPLOŠNA PRISTOPA K TEORIJI EKSTREMNIH VREDNOSTI

Statistiki so se v 18. stoletju med drugim preživljali tudi s svetovanjem igralcem na srečo. Pri iskanju odgovora na vprašanje "Če pravičen kovanec vržemo 100-krat, kolikšna je verjetnost, da pade glava več kot 60-krat?" je francoski matematik Abraham de Moivre (1667–1754) odkril tako imenovano "normalno krivuljo" oziroma "normalno porazdelitev". Neodvisno od njega je francoski matematik Pierre-Simon Laplace (1749–1827) izpeljal centralni limitni izrek, kjer je normalna porazdelitev limita za porazdelitev opazovanega vzorca. Slednji je postal osnovna ideja pri razvoju teorije ekstremnih vrednosti.

Dandanes se statistiki občasno ukvarjajo s svetovanjem ekonomistom, katerim normalna porazdelitev ne ustreza vedno pri izbranem modelu. Če nas na primer zanima kolikšna je verjetnost, da se zgodi novo krizno obdobje v prihajajočih 10 letih, potem normalna porazdelitev modelu ne ustreza več, saj nas zanima verjetnost redkega dogodka, ki pa se nahaja v repu porazdelitve. Razvoj teorije ekstremnih vrednosti se ukvarja predvsem s porazdelitvijo v repu in nam pokaže, da lahko odgovorimo tudi na takšno vprašanje.

Tako sta se do danes razvila in nastala dva pristopa, ki imata za podlago dva temeljna izreka v teoriji ekstremnih vrednosti:

- Prvi pristop ali model maksimumov skupin podatkov, ki temelji na Fisher-Tippetovem izreku (Fisher & Tippett, 1928),
- Drugi pristop ali model preseganja mejne vrednosti, ki temelji na Izreku Pickands-Balkema-de Haan (Balkema & de Haan, 1974; Pickands, 1975).

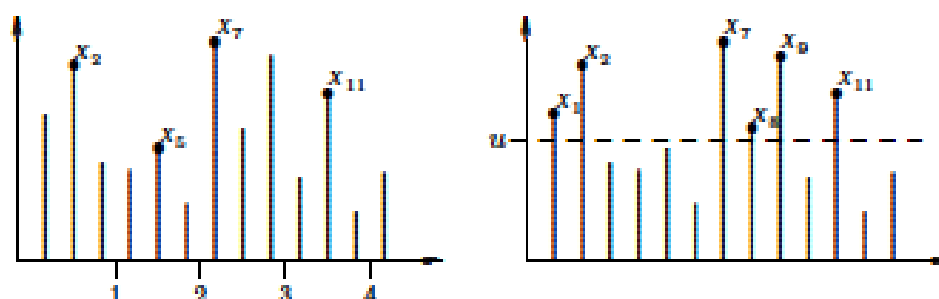
Prvi pristop se imenuje **model maksimumov skupin podatkov** (angl. *Block Maxima Model*) in je najstarejši model analize ekstremnih vrednosti. Prvi ga je opisal in predstavil Gumbel (1958). Ta metoda analizira najbolj ekstremne dogodke v podatkih, ti pa so združeni v skupine glede na pojavljanje v enakih časovnih obdobjih, npr. v mesecih ali letih. Model maksimumov skupin podatkov torej v isto skupino združi dogodke, ki so se zgodili v istem časovnem obdobju, nato pa poišče najvišjo vrednost v posamezni skupini in analizira tako pridobljene podatke. Pri tem modelu je uporabljen Fisher-Tippetov izrek (1928), ki pravi, da je ena izmed treh porazdelitev ekstremnih vrednosti v limiti prilegana opazovanim podatkom. Te tri porazdelitve bomo spoznali v naslednjih poglavjih. Glede na pogostost pojavljanja maksimumov lahko izračunamo opisne statistike in izračunamo cenilko po metodi največjega verjetja za porazdelitev teh maksimumov skupin podatkov. S tem modelom lahko določamo tudi sezone: mesečne, četrtletne ali letne podatke; glede na razpoložljivost podatkov pa lahko opazujemo celo dnevna ali nekaj-urna obdobja. Za ekstremne dogodke se porazdelitvena funkcija tako normaliziranih maksimumov imenuje splošena porazdelitev ekstremnih vrednosti.

V primeru, ko bi se v eni skupini podatkov pojavilo več visokih oziroma maksimalnih vrednosti, medtem ko bi bila v drugih skupinah podatkov maksimalna vrednost precej nizka (glede na ostale maksimalne vrednosti v drugih skupinah), modeliranje samo maksimumov skupin podatkov ne bi bilo več najbolj učinkovito. V opisanem primeru bi obstajalo veliko ekstremnih vrednosti, ki bi se nahajale v desnem repu porazdelitve, vendar ne bi bile vključene v analizo. Tako se je razvil drugi pristop, ki se imenuje **model preseganja mejne vrednosti** (angl. *Peak Over Threshold Model*) in se osredotoča le na opazovanja, ki presegajo vnaprej predpisano mejno vrednost. Včasih POT-model imenujemo tudi model presežkov čez določeno mejno vrednost (angl. *Exceedances Over Thresholds Model*). Ta model uporablja porazdelitev pre-

sežkov čez mejno vrednost, torej se ukvarja le s porazdelitvijo repnega območja. Glavni rezultat v POT-modelu sta pokazala Pickands (1975) ter Balkema in de Haan (1974), ki sta dokazala, da se porazdelitvena funkcija presežka v limiti približuje posplošeni Paretovi porazdelitvi.

Za boljšo predstavbo si oglejmo spodnjo Sliko 1, kjer je na levi strani uprizorjen model maksimumov skupin podatkov in na desni strani model preseganja mejne vrednosti. Za model maksimumov skupin podatkov smo vsa opazovanja razdelili na štiri enako dolga obdobja, kjer opazovanja X_2, X_5, X_7 in X_{11} predstavljajo maksimum vsakega izmed štirih obdobj. Opazimo, da tako nismo zajeli vseh maksimalnih vrednosti, saj je na primer v tretjem obdobju več opazovanih vrednosti višjih od vrednosti X_5 , ki pa predstavlja maksimalno vrednost v drugem obdobju. To omejitev lahko rešimo z izbiro drugega modela, POT-modela, kjer vsa opazovanja X_1, X_2, X_7, X_8, X_9 in X_{11} presegajo neko dano mejo u in tako zajamemo celoten obseg ekstremnih dogodkov.

Slika 1: Model maksimumov skupin podatkov (levo) in model preseganja mejne vrednosti (desno)



Vir: M. Gilli & E. Kellezi, *An Application of Extreme Value Theory for Measuring Financial Risk*, 2006, str.4.

3 PREDSTAVITEV RAZLIČNIH PRISTOPOV K TEORIJ EKSTREMNIH VREDNOSTI

Teorija ekstremnih vrednosti je ena izmed vej v statistiki, ki jo uporabljamo, ko se soočamo s problemom modeliranja vrednosti, ki so visoke in se redko zgodijo. Parametrični pristop k teoriji ekstremnih vrednosti poskuša prilegati porazdelitveno funkcijo opazovanim dogodkom. Pri tem daje večji poudarek na repe porazdelitev kot na osrednja območja porazdelitve. Danes obstajata dva različna parametrična pristopa k teoriji ekstremnih vrednosti. V prvem parametričnem modelu se EVT ukvarja z lastnostmi maksimalnih vrednosti in ugotavlja, kakšna je oblika repa porazdelitve, katerega oblikujejo te maksimalne vrednosti. V drugem parametričnem pristopu pa se EVT ukvarja z lastnostmi vrednosti, ki presegajo mejno vrednost, in ugotavlja, kakšna je ustrezna oblika repa porazdelitve, ki jo te vrednosti opišejo. V teoriji ekstremnih vrednosti se je razvil tudi neparametrični pristop. V magistrski nalogi se bomo med različnimi neparametričnimi pristopi osredotočili predvsem na Hillovo cenilko. Predstavili bomo tudi Pickandsovo cenilko.

3.1 Prvi parametrični pristop

3.1.1 Fisher-Tippetov izrek

Za modeliranje ekstremnih vrednosti je treba pogledati porazdelitvene funkcije ekstremov. Uporabimo enako idejo kot pri centralnem limitnem problemu, opisanem v Prilogi 2. Slednji se sicer osredotoča na lastnosti normaliziranih in centraliziranih kumulativnih vsot, pri EVT pa nas bodo zanimale lastnosti normaliziranih in centraliziranih kumulativnih maksimumov. Torej iščemo karakterizacijo vseh limitnih zakonov H (H je porazdelitvena funkcija za Y) ter zaporedji (a_n) , (b_n) v shemi

$$\frac{\max(X_1, \dots, X_n) - b_n}{a_n} \Rightarrow Y \quad (1)$$

Označimo: $M_n = \max(X_1, \dots, X_n)$.

Rešitev enačbe (1) nam da t. i. **FISHER-TIPPETTOV IZREK**, ki je tudi prvotni pristop k reševanju problema ekstremnih vrednosti.

Izrek 1 (Fisher-Tippetov izrek)¹: Predpostavimo, da so X_1, \dots, X_n neodvisne enako porazdeljene slučajne spremenljivke s porazdelitveno funkcijo F in da obstajata taki normirni konstanti $a_n > 0$ in $b_n \in \mathbb{R}$, da velja

$$F^n(a_n x + b_n) = P \left[\frac{M_n - b_n}{a_n} \leq x \right] \Rightarrow H(x), \quad (2)$$

kjer je H neizrojena limitna porazdelitvena funkcija². Potem ima H eno izmed sledečih porazdelitvenih funkcij ekstremne vrednosti:

$$\begin{aligned} \Phi_\alpha(x) &= \begin{cases} 0, & (x < 0) \\ \exp(-x^{-\alpha}), & (x > 0) \end{cases} && \text{FRÉCHET} \\ \Psi_\alpha(x) &= \begin{cases} 1, & (x < 0) \\ \exp(-(-x)^\alpha), & (x > 0) \end{cases} && \text{WEIBULL} \\ \Lambda(x) &= \exp\{-e^{-x}\} && x \in \mathbb{R} && \text{GUMBEL} \end{aligned} \quad (3)$$

Porazdelitev ekstremnih vrednosti predstavlja možno limitno porazdelitev normaliziranih in centraliziranih maksimumov neodvisnih in enako porazdeljenih (n.e.p.) slučajnih spremenljivk. To je po teoriji ekstremnih vrednosti največja vrednost iz vzorca n.e.p. slučajnih spremenljivk, ki stremi k asimptotični porazdelitvi, ta pa je odvisna le od porazdelitve repa tega vzorca opazovanih slučajnih spremenljivk.

- **Gumbelova porazdelitev ali porazdelitev ekstremnih vrednosti tipa I**

Če je rep opazovane porazdelitve neomejen in pada vsaj tako naglo kot eksponentna funkcija, potem je asimptotična porazdelitev ekstremnih vrednosti Gumbelova porazdelitev. Gumbelova porazdelitvena funkcija je znana tudi kot Fisher-Tippetova porazdelitev tipa I in je največkrat označena kot EVO , F_{EVO} ali Λ . Porazdelitev ekstremnih vrednosti tipa I vključuje primere porazdelitev, kot so: eksponentna, gama, normalna, lognormalna ali

¹ Izrek velja tudi, če namesto x uporabimo njegovo transformacijo $\frac{x-\mu}{\sigma}$, za vsak $\mu \in \mathbb{R}$ in vsak $\sigma > 0$.

² Neizrojena porazdelitvena funkcija je taka porazdelitvena funkcija, ki nima vse gostote v eni točki.

Gumbelova porazdelitev. Ker se rep porazdelitve obnaša kot eksponentna funkcija, vsi momenti Gumbelove porazdelitve obstajajo.

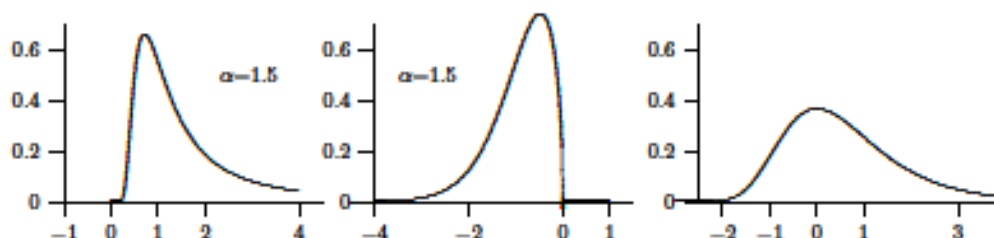
- **Fréchetova porazdelitev ali porazdelitev ekstremnih vrednosti tipa II**

Če je zgornji del repa opazovane porazdelitve neomejen, vendar vsi njeni momenti niso končni, ter hkrati pada po potenčnem zakonu oziroma polinomsko, potem je asimptotična porazdelitev ekstremnih vrednosti Fréchetova porazdelitev. Njeni momenti obstajajo le do stopnje $\alpha = \frac{1}{\xi}$. Ta porazdelitev je znana tudi pod imenom Fisher-Tippettova porazdelitev tipa II in je največkrat označena z oznakami EVI , F_{EV1} ali Φ_α . Primeri porazdelitev, ki spadajo pod porazdelitev ekstremnih vrednosti tipa II, so: Paretova, Burrova, Studentova t, Cauchyjeva, log-gama ali Fréchetova porazdelitev. Porazdelitve, ki imajo tako obliko repa, se z drugim imenom imenujejo tudi porazdelitve debelega repa. Z drugimi besedami to pomeni, da če ima opazovana porazdelitev debel rep, je Fréchetova družina porazdelitev edina družina porazdelitev, kateri lahko ta opazovana porazdelitev v limiti pripada. Mnoge raziskave so pokazale, da je Fréchetova porazdelitev najprimernejša oblika porazdelitve za večino finančnih časovnih serij podatkov.

- **Weibullova porazdelitev ali porazdelitev ekstremnih vrednosti tipa III**

Če ima opazovana porazdelitvena funkcija omejen oziroma končen rep (t. j. obstaja končna točka te porazdelitve) in so hkrati vsi njeni momenti končni, potem je asimptotična porazdelitev ekstremnih vrednosti Weibullova porazdelitev. Poznamo jo tudi pod imenom Fisher-Tippettova porazdelitev tipa III in jo označujemo z EVI , F_{EV2} ali Ψ_α . Primeri porazdelitev ekstremnih vrednosti tipa III so: nekatere potenčne porazdelitve, enakomerna, beta in Weibullova porazdelitev.

Slika 2: Primer Fréchetove ($\alpha = 1,5$), Weibullove ($\alpha = -1,5$) in Gumbelove porazdelitvene funkcije ($\alpha = 0$)



Vir: M. Gilli & E. Kellezi, *An Application of Extreme Value Theory for Measuring Financial Risk*, 2006, str.5.

Kot je razvidno s Slike 2, repni indeks α nadzoruje sploščenost in asimetrijo porazdelitvene funkcije. Za pozitivne α bosta sploščenost in asimetrija hitro narasli.

3.1.2 Posplošena porazdelitev ekstremnih vrednosti

Za enostavnejšo uporabo sta Jenkinson (1955) in von Mises (1936) predlagala, da se uporablja predstavitev porazdelitvene funkcije z enim samim parametrom. Ta predstavlja vse tri primere porazdelitvenih funkcij v enem prikazu. Tako je nastala posplošena porazdelitev ekstremnih vrednosti (angl. *GEV – Generalised Extreme Value Distribution*):

Definicija 1 (Posplošena porazdelitev ekstremnih vrednosti (GEV-porazdelitev))³:

Definirajmo

$$H_{\xi}(x) = \begin{cases} \exp\{-(1 + \xi x)^{-\frac{1}{\xi}}\}, & \text{če } \xi \neq 0 \\ \exp\{-\exp\{-x\}\}, & \text{če } \xi = 0 \end{cases}, \quad (4)$$

kjer je $1 + \xi x > 0$, H_{ξ} pa imenujemo posplošena porazdelitvena funkcija ekstremnih vrednosti s parametrom ξ .

Osnovna ideja je, da parameter ξ obstaja kot limita. Če je:

- $\xi = \frac{1}{\alpha} > 0$, se ujema s Fréchetovo porazdelitvijo Φ_{α} ,
- $\xi = 0$, se ujema z Gumbelovo porazdelitvijo Λ ,
- $\xi = -\frac{1}{\alpha} < 0$, se ujema z Weibullovo porazdelitvijo Ψ_{α} .

Zgornjo parametrično družino $(H_{\xi})_{\xi \in \mathbb{R}}$ lahko zapišemo tudi tako:

$$H_{\xi} = \begin{cases} \Phi_{\frac{1}{\xi}}, & \text{če je } \xi > 0, \\ \Lambda, & \text{če je } \xi = 0, \\ \Psi_{-\frac{1}{\xi}}, & \text{če je } \xi < 0. \end{cases} \quad (5)$$

Ker je tudi za GEV-porazdelitev možna transformacija $x \rightarrow \frac{x-\mu}{\sigma}$, ima ta porazdelitev v resnici lahko tri parametre. Po mnenju Reissa in Thomasa (2000) so ti sledeči :

- μ , ki sta ga poimenovala lokacijski parameter (angl. *location parameter*),
- α oziroma ξ , ki sta ga predstavila kot parameter oblike oziroma Pareto indeks (angl. *shape parameter*),
- σ , ki je definiran kot parameter višine (angl. *scale parameter*).

3.1.3 Ocenjevanje GEV-porazdelitve po metodi največjega verjetja

Vrednosti funkcije verjetnostne gostote pri danih vrednostih odvisnih spremenljivk in parametrov rečemo funkcija verjetja (angl. *likelihood function*). Vrednost parametra, ki maksimira to verjetnost, pa imenujemo cenilka po metodi največjega verjetja oziroma krajše, cenilka največjega verjetja (angl. *maximum likelihood estimator – MLE*). Podrobnejša razlaga funkcije in cenilke največjega verjetja je zapisana v Prilogi 2.

Obstajajo različne metode ocenjevanja modelov ekstremnih vrednosti, kot so različne grafične metode, metode momentov, metoda največjega verjetja, metoda L-momentov... Med vsemi naštetimi je metoda največjega verjetja pri analizi ekstremnih vrednosti najbolj uporabljena, kljub temu da je verjetnost ustrezna le za $\xi > -\frac{1}{2}$ (Zhao, 2010). V nadaljevanju bomo analizirali problem ekstremnih vrednosti na IBM-ovih podatkih, kjer bomo videli, da je $\xi > 0$, kar je

³ Tudi za GEV je možna uporaba transformacije $\frac{x-\mu}{\sigma}$, za vsak $\mu \in \mathbb{R}$ in vsak $\sigma > 0$.

splošno značilno za finančne serije podatkov. Uporaba cenilke največjega verjetja na teh podatkih bo torej upravičena.

Logaritmirana funkcija verjetja L za vzorec n.e.p. slučajnih spremenljivk X_1, \dots, X_n , ki imajo GEV-porazdelitev z uporabljenimi transformacijami $x \rightarrow \frac{x-\mu}{\sigma}$, je:

$$\ln L(\xi, \sigma, \mu) = -n \ln \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^n \ln \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^n \ln \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \quad (6)$$

za $\xi \neq 0$ in kjer je $1 + \xi \left(\frac{x_i - \mu}{\sigma}\right) > 0$,

in

$$\ln L(\sigma, \mu) = -n \ln \sigma - \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^n \exp \left[-\left(\frac{x_i - \mu}{\sigma}\right)\right] \quad (7)$$

za $\xi = 0$.

3.2 Drugi parametrični pristop

Pogosto se nam zgodi, da imamo malce drugačno situacijo od tiste, ki smo jo opisali pri prvem parametričnem pristopu. Odličen primer je zavarovalništvo. Tu nas po navadi ne zanimajo maksimalne izgube v obdobjih, ampak nas zanimajo vse izgube, ki presežejo neko določeno mejno vrednost. Ker zavarovalnice konstantno in dlje časa pobirajo manjše zneske premij, si lahko privoščijo izplačila škod do vrednosti zbranih premij, od te vrednosti naprej pa imajo izgubo zavarovalnice. Zato jih zanimajo predvsem izgube oziroma škodni dogodki, ki presežejo to mejno vrednost. To je seveda le enostaven primer razlage v zavarovalništvu, na podlagi katerega se je razvil drugi pristop k teoriji ekstremnih vrednosti.

V tem poglavju se bomo torej ukvarjali le z verjetnostno porazdelitvijo presežkov čez neko mejno vrednost. Ta pristop imenujemo model preseganja mejne vrednosti. Tako bomo lahko odgovorili na vprašanje: "Če vemo, da so naši podatki ekstremni, kako visok ekstrem lahko dosežejo?" Odgovor se nahaja v funkciji presežka, posplošeni Paretovi porazdelitvi (GPD-porazdelitev) in Izreku Pickands-Balkema-de Haan, ki poveže funkcijo presežka in GPD-porazdelitev.

3.2.1 Funkcija presežka

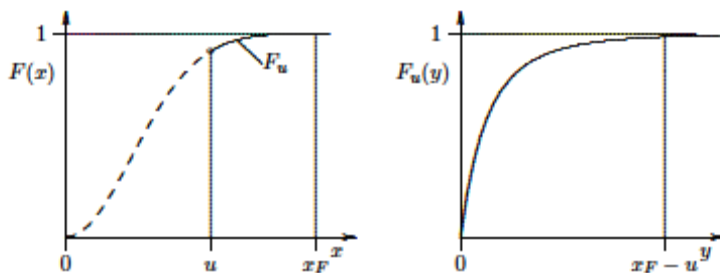
Kot pri analizi modela maksimumov skupin podatkov naj bo X_1, X_2, \dots, X_n zaporedje neodvisnih enako porazdeljenih slučajnih spremenljivk s kumulativno porazdelitveno funkcijo F , ki je neznana. Naj bo maksimum teh slučajnih spremenljivk definiran kot $M_n = \max\{X_1, X_2, \dots, X_n\}$. Ekstremni dogodki so predstavljeni kot vrednosti X , ki presežejo visoko mejno vrednost u . To nam opiše t. i. porazdelitvena funkcija presežka čez določeno mejno vrednost u , ki je definirana kot pogojna verjetnost (Chernobai, Rachev, Fabozzi, 2007):

$$F_u(y) = P\{X - u \leq y | X > u\} = \frac{F(y+u) - F(u)}{1 - F(u)}, \text{ kjer je } y \geq 0. \quad (8)$$

Porazdelitvena funkcija presežka predstavlja verjetnost, da vrednost X preseže mejo u pri vrednosti največ $y \geq 0$, kjer so y presežki, definirani kot $y = X - u$.

Funkcija $e(u) = E(X - u|X > u)$ se imenuje povprečna funkcija presežka.

Slika 3: Porazdelitvena funkcija F in porazdelitvena funkcija presežka F_u



Vir: M. Gilli & E. Kellezi, *An Application of Extreme Value Theory for Measuring Financial Risk*, 2006, str.7.

3.2.2 Posplošena Paretova porazdelitev

Izrek Pickands-Balkema-de Haan, ki ga bomo spoznali v Podpoglavju 3.2.3, pravi, da lahko za dovolj visoko mejo u porazdelitveno funkcijo presežka aproksimiramo s posplošeno Paretovo porazdelitvijo (GPD-porazdelitvijo).

Povedano drugače (Embrechts et al., 1997): porazdelitvena funkcija presežka $F_u(y)$ konvergira h GPD-porazdelitvi, ko se meja u povečuje.

Paretova porazdelitev ima lahko izjemno debel (desni) rep, ki lahko vodi tudi do situacije, ko so zelo visoke izgube že zanemarljivo verjetne. Pri $\alpha < 1$ oziroma $\xi > 1$ govorimo o zelo odebeljenem repu porazdelitve, kjer sta matematično upanje in varianca neskončni. Pri modeliranju ekstremno visokih škod oziroma vrednosti lahko v izogib tej situaciji zaradi praktičnosti izračunov uporabimo nekoliko modificirano dvoparametrično (ali posplošeno) Paretovo porazdelitev, pri kateri se upošteva le prilagajanje empiričnih podatkov v repih porazdelitve.

Definicija 3 (Posplošena Paretova porazdelitev (GPD-porazdelitev)): Definirajmo porazdelitveno funkcijo

$$H_{\xi, \sigma}(x) = \begin{cases} 1 - (1 + \frac{\xi}{\sigma} x)^{-\frac{1}{\xi}}, & \text{če } \xi \neq 0 \\ 1 - e^{-\frac{x}{\sigma}}, & \text{če } \xi = 0 \end{cases}, \quad (9)$$

kjer je $x \geq 0$, če je $\xi \geq 0$ in $0 \leq x \leq -\frac{\sigma}{\xi}$, če je $\xi < 0$. ξ imenujemo parameter oblike, indeks repa ali pa Pareto indeks, pove pa nam debelino repa. Večji kot je ξ , debelejši je rep in obratno.

Med modelom maksimumov skupin podatkov in modelom preseganja mejne vrednosti obstaja tesna povezava. Za izbrano vrednost u parametri ξ , μ in σ GEV-porazdelitve določajo parametre ξ in σ GPD-porazdelitve. Pri GPD-porazdelitvi smo en parameter izgubili, ko smo izbrali velikost mejne vrednosti u . Natančneje, oblikovni parameter GEV porazdelitve ξ je natančno oblikovni parameter ξ v GPD-porazdelitvi in je neodvisen od mejne vrednosti u . Posledično, če je $\xi < 0$, potem je F porazdelitvena funkcija v Weibullovi družini porazdelitev in $H_{\xi, \sigma}$ je

Paretova porazdelitev tipa II; če je $\xi = 0$, potem je F v Gumbelovi porazdelitveni družini in $H_{\xi,\sigma}$ je eksponentna porazdelitev; če pa je $\xi > 0$, potem je F v Fréchetovi porazdelitveni družini in $H_{\xi,\sigma}$ je Paretova porazdelitev tipa I.

3.2.3 Izrek Pickands-Balkema-de Haan

Izrek 2 (Izrek Pickands-Balkema-de Haan): Naj bodo X_1, \dots, X_n neodvisne enako porazdeljene slučajne spremenljivke s porazdelitveno funkcijo F in F_u funkcija presežka. Potem lahko za velik razred porazdelitvenih funkcij F funkcijo presežka aproksimiramo s

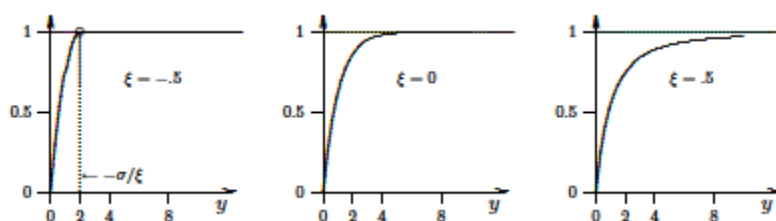
$$F_u(y) \approx H_{\xi,\sigma}(y), \quad u \rightarrow \infty, \quad (10)$$

kjer je

$$H_{\xi,\sigma}(y) = \begin{cases} 1 - \left(1 + \frac{\xi}{\sigma}y\right)^{-\frac{1}{\xi}}, & \text{če } \xi \neq 0 \\ 1 - e^{-\frac{y}{\sigma}}, & \text{če } \xi = 0 \end{cases}, \quad (11)$$

za $y \geq 0$, če je $\xi \geq 0$ in $0 \leq y \leq -\frac{\sigma}{\xi}$, če je $\xi < 0$.

Slika 4: Posplošena Paretova porazdelitev $H_{\xi,\sigma}$ za negativne, pozitivne in ničelne vrednosti ξ , kjer je σ fiksiran na vrednost 1



Vir: M. Gilli & E. Kellezi, *An Application of Extreme Value Theory for Measuring Financial Risk*, 2006, str.8.

3.2.4 Ocenjevanje GPD-porazdelitve po metodi največjega verjetja

Tako kot pri GEV-porazdelitvi je tudi pri GPD-porazdelitvi najprimernejša izbira ocenjevanja parametrov metoda največjega verjetja za $\xi > -\frac{1}{2}$, ki ustreza porazdelitvi debelega repa (Hosking & Wallis, 1987). Pri GPD-porazdelitvi moramo vnaprej izbrati mejno vrednost u , ki nam tvori nov vzorec. Prvotni vzorec obsega slučajne spremenljivke X_1, \dots, X_n , po izbrani u pa nam ta mejna vrednost določi nov vzorec slučajnih spremenljivk X_1, \dots, X_m ($m < n$), ki so prav tako n.e.p. po GPD-porazdelitvi. Logaritmirana funkcija verjetja je enaka:

$$\ln L(\xi, \sigma | u) = -m \ln \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^m \ln \left[1 + \xi \left(\frac{x_i - u}{\sigma}\right)\right]. \quad (12)$$

3.2.5 Ocenjevanje porazdelitve repa

Ugotovili smo, da nas pri drugem pristopu k EVT ne zanima celotna porazdelitev, vendar le porazdelitev in debelost repa. Po Izreku 2 vemo, da $F_u(y)$ konvergira h GPD-porazdelitvi za

zelo velike u , iz katerega lahko ocenimo, kakšna je porazdelitev repa. To oceno imenujemo repna cenilka oziroma repni indeks ξ . Kot že omenjeno, večji kot je ξ , debelejši je rep.

V tem poglavju bomo izpeljali izračun repne cenilke, v Poglavju 4 pa bomo s pomočjo te enačbe izpeljali še kvantilne ocene, natančnejše VaR, pričakovani izpad in stopnjo poplačila.

Izračun repne cenilke na podlagi Izreka Pickands-Balkema-de Haan :

- Vzemimo $x = u + y$ in izrazimo $F(x)$ iz porazdelitvene funkcije presežka $F_u(x - u) = \frac{F(x) - F(u)}{1 - F(u)}$:

$$F(x) = (1 - F(u))F_u(x - u) + F(u). \quad (13)$$

- Nadomestimo $F_u(y)$ z oceno GPD-porazdelitve $H_{\xi, \sigma, u}(x - u)$ (kot pravi Izrek 2) in $F(u)$ z oceno $\frac{n - N_u}{n}$ (Gilli & Kellezi, 2006). Oznaka n predstavlja končno število opazovanj, N_u pa število opazovanj, ki presegajo mejo u . Tako dobimo:

$$\hat{F}(x) = \frac{N_u}{n} \left(1 - \left(1 + \frac{\hat{\xi}}{\hat{\sigma}}(x - u) \right)^{-\frac{1}{\hat{\xi}}} \right) + \left(1 - \frac{N_u}{n} \right). \quad (14)$$

- To se poenostavi v :

$$\hat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \frac{\hat{\xi}}{\hat{\sigma}}(x - u) \right)^{-\frac{1}{\hat{\xi}}}, \quad (15)$$

kjer $\hat{\xi}$ in $\hat{\sigma}$ predstavljata cenilki za ξ in σ po metodi največjega verjetja.

- Za izračun cenilki po metodi največjega verjetja moramo najprej vnaprej določiti mejno vrednost u . Ta mejna vrednost mora biti dovolj visoka, da bo GPD-aproksimacija ustrezna in hkrati dovolj nizka, da bo na voljo zadostna količina opazovanj za natančno prileganje porazdelitve opazovanim podatkom.

3.3 Neparometrični pristop in cenilka repnega indeksa

Najbolj znan neparometrični pristop k teoriji ekstremnih vrednosti je Hillova cenilka, ki ocenjuje repni indeks porazdelitve z debelim repom. Ocena repnega indeksa nam določa, kolikšna je verjetnost, da se zgodi ekstremni dogodek. Repni indeks porazdelitve je indikator, da opazovana porazdelitev močno odstopa od povprečne vrednosti v osrednjem delu porazdelitve. Debelina in oblika repa sta neposredno povezana z repnim indeksom porazdelitve. Označimo ga s $\xi = \frac{1}{\alpha}$ in predstavlja najvišji red, do katerega vsi momenti še obstajajo.

3.3.1 Hillova cenilka

Do sedaj je bilo opravljenih že veliko raziskav o cenilkah repnega indeksa oziroma o cenilkah parametra oblike, vendar do danes ostaja Hillova cenilka še vedno najbolj znana in največkrat uporabljena (Rocco, 2011). Kot je razvidno že iz poimenovanja, jo je leta 1975 opisal Bruce M. Hill.

Hillova cenilka prekaša druge cenilke predvsem po tem, da se osredotoča na ocene repnega indeksa. Povedano drugače, osredotoča se na primere, ko je kumulativna porazdelitvena funkcija v maksimalnem območju privlačnosti Fréchetove porazdelitve. V drugih primerih Hillova cenilka ne deluje, zato je treba uporabiti katero drugo cenilko.

Definicija 4 (Hillova cenilka): Naj bodo X_1, \dots, X_n neodvisne enako porazdeljene slučajne spremenljivke, $X_{(1)}, \dots, X_{(n)}$ vrstilne statistike. Če uporabljamo $k + 1$ vrstilnih statistik, je Hillova cenilka repnega indeksa ξ enaka:

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \ln \left(\frac{X_{(i)}}{X_{(k+1)}} \right) = \tilde{\xi}^{Hill}(k). \quad (16)$$

Vidimo, da je cenilka odvisna od parametra $k \in \{1, \dots, n - 1\}$. Ta predstavlja neko mejno vrednost, ki ločuje osrednja opazovanja od opazovanj v repu porazdelitve. Povedano drugače, Hillova cenilka meri razdaljo, s katero ekstremni dogodki v povprečju presegajo mejno vrednost u .

Kot že večkrat omenjeno, velja $\xi = \frac{1}{\alpha}$, kar velja tudi za Hillovo cenilko: $\xi^{Hill} = \frac{1}{\alpha^{Hill}}$.

3.3.2 Pickandsova cenilka

James Pickands III je prav tako leta 1975 v svojem delu opisal cenilko, ki se po Hillovi cenilki razlikuje po tem, da lahko oceni parameter oblike oziroma repni indeks za katerokoli od nam že znanih treh porazdelitev ekstremnih vrednosti (Gumbelovo, Fréchetovo in Weibullovo porazdelitev).

Definicija 5 (Pickandsova cenilka): Naj bodo X_1, \dots, X_n neodvisne enako porazdeljene slučajne spremenljivke, $X_{1,n}, \dots, X_{n,n}$ vrstilne statistike. Pickandsova cenilka repnega indeksa ξ je:

$$\tilde{\xi}^{Pickands}(k, n) = \frac{1}{\ln 2} \ln \frac{X_{k,n} - X_{2k,n}}{X_{2k,n} - X_{4k,n}}. \quad (17)$$

4 MERJENJE TVEGANJA

Najbolj pogosto vprašanje, ki zadeva obvladovanje tveganja v financah, je ocena ekstremnih kvantilov, saj nas zanima, kolikšnemu tveganju smo izpostavljeni oziroma koliko "premoženja" lahko izgubimo. V nadaljevanju bomo govorili o tržnem tveganju, ki predstavlja tveganje v spremembah cen ali vrednostih, kjer nas zanima le vrednost izgub in ne vrednost dobičkov. Tipičen in popularen primer merjenj tveganja je po mnenju Joriona (2001) Value at Risk (VaR). Drugi, manj pogosto uporabljeni metodi pa sta pričakovani izpad (angl. *Expected Shortfall* oziroma ES) in stopnja poplačila (angl. *Return Level*).

4.1 VaR

Splošna definicija predstavlja VaR kot mero tveganja, ki meri največjo potencialno izgubo, ki naj bi jo utrpel imetnik določenega portfelja ob neki dani stopnji zaupanja v določenem investicijskem obdobju.

Predpostavimo, da slučajna spremenljivka X z zvezno porazdelitveno funkcijo F modelira izgube oziroma negativne rezultate za nek določen finančni instrument v nekem določenem časovnem obdobju. VaR_p lahko definiramo kot p -ti kvantil porazdelitve F

$$VaR_p = F^{-1}(p), \quad 0 < p < 1, \quad (18)$$

kjer F^{-1} imenujemo kvantilna funkcija (opisana v Prilogi 2) in je definirana kot inverz porazdelitvene funkcije F (opisan v Prilogi 2). Po navadi VaR predstavlja 95. ali 99. percentil oziroma 95 % ali 99 % kvantil. Zapišemo, da je $p = 0,95$ ali $p = 0,99$. Oznaka p predstavlja vnaprej izbrano stopnjo zaupanja, ki jo še vzamemo za sprejemljivo.

Če nas, na primer, zanima 95 % kvantil, potem $VaR_{0,95}$ predstavlja vrednost, ki jo preseže le 5 % observacij (ta so urejena od najnižje do najvišje vrednosti) naključne spremenljivke. To pomeni, da bo izguba preseгла $VaR_{0,95}$ s 5 % verjetnostjo.

Med prednosti uporabe VaR lahko štejemo možnost primerjave med različnimi vrstami tveganja, kot tudi med različnimi subjekti ali institucijami, ki jih opazujemo (Wilson, 1998). Kot predlaga Baselski kapitalski sporazum ali Basel II (Basel committee on Banking Supervision, 2001), ga lahko uporabimo neposredno kot instrument za določanje zahtevanega minimalnega kapitala za varovanje pred tveganji. Rezultati metode VaR pomagajo tudi pri managementu, saj pomagajo določiti izpostavljenost tveganju za vsako posamezno organizacijsko enoto (Down, 2002).

Leta 2001 je Baselski komite za nadzor bank predlagal VaR za primerno mero tveganja, kar nakazuje, da je VaR postal zelo uporabljena in prikladna mera za merjenje in ocenjevanje tveganj. Prvi, ki je uporabil VaR v svojem poslovnem poročilu, je bil v devetdesetih letih J. P. Morgan. Vendar pa je treba VaR, kot pravi Szegö (2002), previdno obravnavati, saj je bil večkrat napačno privzet kot mera tveganja.

Pristop z VaR metodo se osredotoča na rezultate v normalnih tržnih pogojih (Jorion, 2001). Torej se ta metoda uporablja za aproksimacijo z normalno porazdelitvijo, kar pa podcenjuje tveganje visokih kvantilov, še posebej v primeru porazdelitev debelega repa, ki so zelo pogosti v finančnih podatkih. Torej je VaR omejen zgolj na prikaz meje velikih izgub, ne pa tudi na višine vseh velikih izgub, ki se nahajajo desno od vnaprej izbrane meje v porazdelitvi (Yamai in Yoshida, 2002).

Na še eno slabost VaR že dlje časa opozarjajo med drugimi tudi Artzner, Delbaen, Eber in Heath (1999), Mina in Xiao (2001), Embrechts, McNeil in Straumann (2002) ter Chavez-Demoulin, Embrechts in Nešlehová (2006). Ta slabost je neustrežanje pogoju subaditivnosti. Ta se pojavi, če podatki ne ustrezajo normalni ali Studentovi t -porazdelitvi s končnimi variancami. Zato VaR v primeru porazdelitev ekstremnih vrednosti ne zadostuje vsem pogojem mere tveganja. Najprej si poglejmo definicijo mere tveganja (Szegö, 2002), po kateri sledi razlaga:

Definicija 6: Da bi bila mera tveganja ρ , ki naboru podatkov o škodah

$$L = X_1, X_2, \dots, X_n \quad (19)$$

priredi mero tveganja, ustrežna kot metoda merjenja tveganja, morajo veljati štiri pogoji:

- translacijska invarianca: $\rho(x + \alpha r_0) = \rho(x) - \alpha$ za vsak $x \in \mathbb{R}$, $\alpha \in \mathbb{R}$, r_0 pa je netvegana stopnja,
- subaditivnost: $\rho(x + y) \leq \rho(x) + \rho(y)$, za vsak $x, y \in \mathbb{R}$,
- pozitivna homogenost: $\rho(\lambda x) = \lambda \rho(x)$, za vsak $x \in \mathbb{R}$, $\alpha \in \mathbb{R}$,
- monotonost: če velja $x \leq y$, potem velja tudi $\rho(x) \leq \rho(y)$, za vsak $x, y \in \mathbb{R}$.

Če subaditivnost ne bi veljala, bi to pomenilo, da bi bilo namesto razpršitve tveganj preko različnih naložb v portfelju bolj primerno naložiti vse premoženje v eno samo naložbo, kar pa vemo, da ni res.

Ravno merilo subaditivnosti pri VaR ne velja vedno, še posebej v primerih škod oziroma padcev, ki ustrezajo porazdelitvam z debelimi repi, zato v takih primerih namesto VaR raje uporabimo pogojni VaR oziroma pričakovani izpad (ES). To predlagajo tudi številni avtorji, npr. Artzner et al. (1999), Jorion (2001), Mina in Xiao (2001), Rockafellar in Uryasev (2002).

4.1.1 Uporaba VaR v EVT (GPD-porazdelitev)

Kljub tem pomankljivostim pa se VaR ocena v praksi še vedno uporablja. Preuredimo jo lahko tudi za porazdelitve z debelim repom, kar prinaša določene prednosti. Prednost VaR, ki temelji na EVT, pred VaR, ki predpostavlja normalno porazdelitev, je ta, da prva ne predpostavlja, kakšna je porazdelitev dogodka oziroma vrste, ki jo opazujemo, ter da vemo, kolikšna je potencialna možna velikost izgube.

Za lažjo predstavo iz Podpoglavja 3.2.5 prepíšimo izpeljano formulo za repno cenilko:

$$\widehat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \frac{\xi}{\widehat{\sigma}} (x - u) \right)^{-\frac{1}{\xi}}. \quad (20)$$

Za podano stopnjo verjetnosti p je kvantil \widehat{x}_p na repu ocenjen z obratom repne cenilke:

$$VaR_p = F^{-1}(p), \quad (21)$$

$$\widehat{VaR}_p = \widehat{x}_p = u + \frac{\widehat{\sigma}}{\xi} \left(\left(\frac{n}{N_u} (1 - p) \right)^{-\xi} - 1 \right), \quad (22)$$

kjer je n končno število opazovanj in N_u število opazovanj, ki presegajo mejo u . V statistiki je to kvantilna ocena, v financah pa jo poimenujemo VaR ocena.

4.2 ES

Pričakovani izpad (angl. *ES – Expected Shortfall*), ki mu večkrat rečemo tudi pogojni VaR (angl. *CVaR – Conditional Value at Risk*), je repno pogojno matematično upanje, ki ocenjuje potencialno velikost izgube, ki presega VaR:

$$ES_p = E(X|X > VaR_p) \quad (23)$$

Z drugimi besedami, ES nam pove, kolikšna je pričakovana pogojna izguba, ko se zgodi dogodek, ki leži v odebeljenem desnem repu porazdelitve in je večji od VaR.

Razpišimo ES_p kot:

$$\widehat{ES}_p = \widehat{VaR}_p + E(X - \widehat{VaR}_p | X > \widehat{VaR}_p), \quad (24)$$

kjer je pogojno matematično upanje ravno pričakovana vrednost funkcije presežka nad mejno vrednostjo \widehat{VaR}_p . Znana je kot povprečna funkcija presežka za posplošeno Paretovo porazdelitev s parametrom $\xi < 1$:

$$e(z) = E(X - z | X > z) = \frac{\sigma + \xi z}{1 - \xi}, \quad \sigma + \xi z > 0, \quad z = VaR_p - u. \quad (25)$$

Ta funkcija vrne povprečje presežkov X -a nad spreminjajočo vrednostjo meje z .

4.2.1 Uporaba ES v EVT (GPD-porazdelitev)

Če poznamo definicijo pričakovanega izpada

$$ES_p = E(X | X > VaR_p) \quad (26)$$

in povprečno funkcijo presežka za GPD-porazdelitev

$$e(z) = E(X - z | X > z) = \frac{\sigma + \xi z}{1 - \xi}, \quad \sigma + \xi z > 0, \quad (27)$$

za $z = VaR_p - u$ in X (ki predstavlja presežek y čez u), znamo razpisati:

$$\widehat{ES}_p = \widehat{VaR}_p + \frac{\hat{\sigma} + \hat{\xi}(VaR_p - u)}{1 - \hat{\xi}} = \frac{VaR_p}{1 - \hat{\xi}} + \frac{\hat{\sigma} + \hat{\xi}u}{1 - \hat{\xi}}. \quad (28)$$

Za razliko od VaR pogojni VaR oziroma ES zadostuje merilu subaditivnosti, zato je njegova uporaba bolj primerna pri modeliranju tveganih dogodkov, ki imajo lastnosti ekstremnih vrednosti.

4.3 Stopnja poplačila

Stopnja poplačila predstavlja pričakovano višino plačila na primer (po)zavarovalnice, ki bo v povprečju presežena enkrat v k ($k = \frac{1}{p}$) obdobjih dolžine n .

Če je F porazdelitev opazovanih maksimalnih vrednosti, ki se nahajajo nad zaporednimi, neprekrivajočimi in enako dolgimi obdobji, potem je stopnja poplačila enaka

$$R_{n,k} = F^{-1}\left(1 - \frac{1}{k}\right). \quad (29)$$

Vidimo, da tudi stopnja poplačila, enako kot VaR, predstavlja p -ti kvantil porazdelitve F :

$$R_{n, \frac{1}{p}} = F^{-1}(1 - p). \quad (30)$$

4.3.1 Uporaba stopnje poplačila v EVT (GEV-porazdelitev)

V primeru, ko opazovani podatki ustrezajo GEV-porazdelitvi in imamo k obdobji dolžine n , je stopnja poplačila enaka:

$$R_{n,k} = \mu_n - \frac{\sigma_n}{\xi_n} \left(1 - \left(-\ln \left(1 - \frac{1}{k} \right) \right)^{-\xi_n} \right). \quad (31)$$

4.3.2 Uporaba stopnje poplačila v EVT (GPD-porazdelitev)

V primeru, ko opazovani podatki ustrezajo GPD-porazdelitvi in imamo k obdobji dolžine n ter N_u predstavlja število opazovanj, ki presegajo mejo u , potem je stopnja poplačila enaka:

$$R_{n,k} = u + \frac{\sigma_n}{\xi_n} \left(\left(\frac{N_u}{n} k \right)^{\xi_n} - 1 \right). \quad (32)$$

5 GRAFIČNA ANALIZA PODATKOV

S pomočjo spodnjih treh grafičnih analiz podatkov si lahko pomagamo pri izboru mejne vrednosti u in posledično tudi pri oceni repnega indeksa ξ . Mejna vrednost u mora biti dovolj visoka, da porazdelitvena funkcija presežka $F_u(y)$ v limiti konvergira h GPD-porazdelitvi. Le tako bo uporaba Izreka Pickands-Balkema-de Haan upravičena.

Interpretacija prvih dveh primerov grafične analize je predstavljena v delih avtorjev, kot so Beirlant, Teugels in Vyncker (1996), Embrechts et al. (1997) ter v Hogg in Klugman (1984). Zadnji primer, Hillovo cenilko oziroma Hillov graf, pa je prvi predstavil Hill (1975), po katerem sta tudi dobila ime.

5.1 Graf kvantilov

Definicija 7 (Graf kvantilov): Naj bodo X_1, \dots, X_n neodvisne enako porazdeljene slučajne spremenljivke, $X_{(1)}, \dots, X_{(n)}$ vrstilne statistike. Naj bo F hipotetična porazdelitev podatkov. Graf kvantilov je definiran z nizom točk

$$\left\{ X_{(k)}, F^{-1} \left(\frac{n-k+1}{n} \right), k = 1, \dots, n \right\}. \quad (33)$$

V statistiki je graf kvantilov zelo uporabno orodje za raziskovanje, ali je določen vzorec res porazdeljen, kot želimo. Graf kvantilov ima na osi x ponazorjene kvantile raziskovane porazdelitve, na osi y pa kvantile hipotetične porazdelitve (pri EVT je to po navadi Gumbelova porazdelitev). Če vzorec prihaja iz hipotetične porazdelitve (Gumbelove), potem je graf kvantilov linearne oblike. Kot že rečeno, je za graf kvantilov v EVT tipično uprizorjena Gumbelova porazdelitev, tj. porazdelitev suhega repa. Če so podatki iz Gumbelovega maksimalnega območja privlačnosti (opisanega v Prilogi 2), potem točke na grafu ležijo vzdolž premice. Če nastane konkavni vzorec, potem imamo porazdelitev debelega repa, če pa je konveksni, imamo porazdelitev suhega repa.

5.2 MEF in ME-Plot

Funkcijo povprečnih odmikov oziroma MEF (angl. *Mean Excess Function*) uporabljamo za testiranje porazdelitve repa in za določanje višine meje u . Zanima nas porazdelitev MEF.

Funkcija povprečnih odmikov je definirana kot kvocient med vsoto presežkov nad mejno vrednostjo u in številom podatkov, ki so presegli to mejo u , označeno z N_u :

$$e_n(u) = \frac{1}{N_u} \sum_{i=1}^{N_u} (X_i - u)^+. \quad (34)$$

Spomnimo se povprečne funkcije presežka za GPD-porazdelitev:

$$e(u) = E(X - u | X > u) = \frac{\sigma + \xi u}{1 - \xi}, \quad \sigma + \xi u > 0. \quad (35)$$

Cenilka za funkcijo $e(u)$ je povprečje opazovanj, ki presežejo mejno vrednost u , torej ravno zgoraj zapisan MEF - $\widehat{e}_n(u)$.

MEF ocenjuje funkcijo presežka, ki opisuje pričakovan presežek, ko se presežek že pojavi. Če so podatki iz GPD-porazdelitve s pozitivnim ξ (porazdelitev debelega repa), potem je MEF uprizorjen kot premica s pozitivnim naklonom nad neko določeno mejno vrednostjo u . Če podatki prihajajo iz porazdelitve, ki je iz Gumbelovega maksimalnega območja privlačnosti (porazdelitev suhega repa), dobimo premico z negativnim naklonom.

Graf povprečnih odmikov (angl. *ME-Plot*) je grafično orodje, s katerim si po mnenju McNeila in Freya (2000) lahko pomagamo izbrati pravilno mejno vrednost u , ki jo iščemo pri drugem parametričnem pristopu. Definiran je kot:

$$\{(u, e_n(u)), X_{(n)} < u < X_{(1)}\}. \quad (36)$$

5.3 Hillov graf

Definicija 8 (Hillov graf): Naj bodo X_1, \dots, X_n neodvisne enako porazdeljene slučajne spremenljivke, $X_{(1)}, \dots, X_{(n)}$ vrstilne statistike. Če uporabljamo $k + 1$ vrstilnih statistik, je Hillova cenilka repnega indeksa ξ enaka:

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \ln \left(\frac{X_{(i)}}{X_{(k+1)}} \right) = \widehat{\xi}^{Hill}(k). \quad (37)$$

Hillov graf je definiran z nizom točk: $\{(k, H_{k,n}^{-1}), 1 \leq k \leq n - 1\}$.

S tega grafa razberemo presežek u , da določimo stabilno območje repnega indeksa. Hillova cenilka je cenilka največjega verjetja za GPD-porazdelitev. Ker ekstremne porazdelitve konvergirajo h GPD-porazdelitvi nad visoko mejno vrednostjo u (Izrek Pickands-Balkema-de Haan), je uporaba Hillovega grafa upravičena, če je u dovolj velik.

6 PROCEDURA UPORABE TER PREDNOSTI IN SLABOSTI TEORIJE EKSTREMNIH VREDNOSTI TER NJENIH MODELOV

Teorija ekstremnih vrednosti ima sicer visoko aplikativno vrednost, ko se soočamo s podatki, ki opisujejo porazdelitev z debelim repom, vendar njena uporaba ni vedno enostavna. Za lažje razumevanje in enostavnejšo uporabo, je v nadaljevanju predstavljena procedura uporabe EVT po korakih in možne prednosti in slabosti njene uporabe.

6.1 Procedura uporabe

Prvi pristop k teoriji ekstremnih vrednosti:

- Razdelimo vsa opazovanja v enako dolge neprekrivajoče skupine (obdobja) podatkov. Dolžina teh obdobij je lahko poljubna, vendar daljše kot je obdobje, manj podatkov bo na voljo. Najpogostejša izbira so mesečna, četrtna ali letna obdobja.
- Iz vsakega obdobja izberemo maksimalno vrednost.
- Sedaj GEV-porazdelitev prilegamo tem maksimalnim vrednostim. Tako dobimo ocene vseh treh parametrov te porazdelitve, ξ , μ in σ . Parametri so ponavadi ocenjeni po metodi največjega verjetja, lahko pa uporabimo tudi druge cenilke.
- Če je ocenjen oblikovni parameter $\xi > 0$, potem podatki kažejo na porazdelitev debelega repa, če je $\xi < 0$, potem podatki kažejo na porazdelitev suhega repa, če pa je $\xi = 0$, potem se podatki porazdeljujejo eksponentno.
- Grafično preverimo, če se izbrana GEV-porazdelitev dobro prilega podatkom. Navečkrat uporabljeno grafično orodje pri EVT je graf kvantilov.
- Ko je ustreznost GEV-aproksimacije potrjena in so parametri porazdelitve ocenjeni, lahko izračunamo različne mere tveganja: stopnja poplačila, VaR ali ES.

Drugi pristop k teoriji ekstremnih vrednosti:

- Določiti moramo višino mejne vrednosti u . To lahko storimo s pomočjo grafičnih orodij (graf kvantilov, ME-Plot) ali s pomočjo enostavne funkcije, ki izračuna kolikšna je mejna vrednost, če nas zanima določen delež najvišjih vrednosti (npr. kolikšna je mejna vrednost, da nad njo leži le največjih 5% opazovanj). Delež ali mejno vrednost lahko določamo sami, vendar moramo paziti, da bo izbrana mejna vrednost dovolj visoka, da bo GPD-aproksimacija ustrezna in hkrati dovolj nizka, da bo na voljo zadostna količina opazovanj za natančno prileganje porazdelitve opazovanim podatkom.
- Z izbrano mejno vrednostjo določimo, katere so tiste vrednosti presežkov (opazovanja), ki jih bomo uporabljali v nadaljni analizi.
- Ko je mejna vrednost določena, lahko npr. z metodo največjega verjetja ocenimo parametra ξ in σ GPD-porazdelitve.
- Če je ocenjen oblikovni parameter $\xi > 0$, potem podatki kažejo na porazdelitev debelega repa, če je $\xi < 0$, potem podatki kažejo na porazdelitev suhega repa, če pa je $\xi = 0$, potem se podatki porazdeljujejo eksponentno.
- Grafično preverimo (npr. z grafom kvantilov) ali je mejna vrednost pravilno izbrana in ali se GPD-porazdelitev dobro prilega podatkom.
- Ko je mejna vrednost ustrezno določena in so parametri porazdelitve ocenjeni, lahko izračunamo različne mere tveganja: stopnja poplačila, VaR ali ES.

6.2 Prednosti in slabosti EVT in njenih modelov

Poleg znanstveno potrjenih prednosti EVT obstaja tudi nekaj skupnih problemov v obeh parametričnih metodah EVT. Tako prednosti kot slabosti EVT so opisane v nadaljevanju tega poglavja.

Po mnenju Embrechthsa et al. (1997), Bensalaha (2000), Colesa (2001), Fernandez (2003b) so prednosti uporabe modela ekstremnih vrednosti sledeče:

- Teorija ekstremnih vrednosti vsebuje matematično in statistično podprte teoretične lastnosti porazdelitev ekstremnih vrednosti in tako predstavlja dobro orodje za obdelavo dogodkov z nizko frekvenco in visokimi vrednostmi.
- Teorija ekstremnih vrednosti se neposredno osredotoča na ekstremne vrednosti. Te zavzemajo repna območja porazdelitve.
- POT-model je zelo uporaben za analizo katastrofalnih izgub, ki ležijo preko neke visoke mejne vrednosti. Podobno analizo predstavlja tudi metoda VaR.
- Na voljo so tako računke, grafične kot tudi teoretične metode določanja lastnosti porazdelitve ekstremnih vrednosti. To velja tudi za rep porazdelitve, kar v običajnih analizah po navadi ni na voljo.
- Na voljo so neparametrični kazalniki oblikovnega parametra porazdelitve ekstremnih vrednosti. Zelo znan in uporabljen je Hillov kazalnik oziroma Hillova cenilka.
- Teorija ekstremnih vrednosti določa trdno metodološko osnovo za izračun ocene VaR za repna območja porazdelitve. Določa tudi izračun pričakovane izgube, ki je višja od VaR, tj. določa izračun ES.

Tako v matematiki kot tudi v statistiki se skoraj pri vsaki tematiki soočamo z določenimi predpostavkami in omejitvami. To velja tudi za teorijo ekstremnih vrednosti, kjer so se določene slabosti odpravile z razvojem novega pristopa, kljub temu pa manjše omejitve še vedno obstajajo. Ustreznost modela EVT je v veliki meri odvisna od subjektivne ocene zajema ekstremnih podatkov in poznavanja celotne analize ekstremnih vrednosti. S ciljem minimiziranja možnosti subjektivne napake, smo opravili analizo tako obeh možnih parametričnih pristopov, kot tudi neparametričnih. Poleg tega smo naredili tudi več grafičnih pristopov k teoriji ekstremnih vrednosti in s tem še dodatno potrdili pravilnost. V nadaljevanju navajamo možne slabosti teorije ekstremnih vrednosti.

Kot smo že ugotovili v Poglavju 2, modeliranje samo z metodo maksimumov skupin podatkov ni najbolj učinkovito, če imamo na voljo še ostale podatke o ekstremnih vrednostih. Predpostavimo, da nas zanimajo mesečne skupine podatkov in se v prvem mesecu zgodi veliko ekstremnih vrednosti, v naslednjem mesecu pa je najvišja ekstremna vrednost nižja od večine ekstremnih vrednosti v prvem mesecu. Metoda maksimumov skupin podatkov bo uporabila le eno najvišjo vrednost v prvem mesecu in eno najvišjo vrednost v drugem mesecu. Tako bomo zanemarili veliko ekstremnih vrednosti, ki so se zgodile v prvem mesecu in so lahko celo višje od ekstrema v drugem mesecu. Da bi se izognili takemu primeru, se je razvil še drugi parametrični pristop k EVT, model preseganja mejnih vrednosti ali POT-metoda. Ta upošteva vse ekstremne vrednosti, ki presežejo neko vnaprej določeno mejno vrednost. Tako ne zanemarimo ekstremnih dogodkov, ki se zgodijo v istem opazovanem obdobju, saj je obseg ekstremnih vrednosti odvisen le od izbire mejne vrednosti.

Opazimo, da se v obeh parametričnih pristopih soočamo s problemom zajemanja ustreznega obsega ekstremnih podatkov. Osredotočimo se najprej na prvi parametrični pristop – na meto-

do maksimumov skupin podatkov. Če je število opazovanj v eni skupini podatkov oziroma v obdobju premajhno, imamo problem s pristranskostjo ocenjenih parametrov. Če pa imamo veliko opazovanj v obdobju in malo obdobj, imamo opravka z napako vzorčenja. Podobno velja tudi pri izbiri mejne vrednosti. Če je izbrana mejna vrednost prenizka, imamo problem s pristransko cenilko parametrov. Če pa je izbrana mejna vrednost previsoka, imamo opravka z visoko varianco, saj imamo na voljo premalo podatkov.

Embrecths et al. (1997), Smith (1999), Bensalah (2000), Kellezi in Gilli (2000), Coles (2001) ter Fernandez (2003b) omenjajo še naslednje slabosti teorije ekstremnih vrednosti:

- Predpostavka v teoriji ekstremnih vrednosti je, da so vsi podatki neodvisni in enako porazdeljeni. Neodvisnost je lahko kršena npr. s serijsko korelacijo, enakost porazdelitve pa ne velja v primeru sezonskih značilnosti ali trendov v podatkih. Vendar je te kršitve možno s poznavanjem področja ekonometrije odpraviti. V empiričnem delu magistrske naloge, kjer smo aplicirali uporabo EVT na podatke o donosih delnice podjetja IBM, se z zgornjimi problemi nismo soočali, zato takšna analiza ni eksplicitno narejena.
- Limitni zakon GEV-porazdelitve za model maksimumov skupin podatkov trdi, da bo, ko ima standariziran maksimum limito, ta limita pripadala eni izmed GEV-porazdelitev. Ta zakon sam po sebi ne zagotavlja obstoja limite in v realnosti obstaja širok razred modelov, za katere ta limita sploh ne obstaja. Če pa obstaja, velja, da bo pripadala eni izmed GEV-porazdelitev. Enako drži za limitni zakon GPD-porazdelitve.
- Hitrost konvergiranja podatkov k limitnim porazdelitvam EVT je lahko zelo počasna. To pomeni, da potrebujemo za dobro aproksimacijo porazdelitve zelo veliko podatkov.
- Izbira mejne vrednosti, na kateri temelji POT-metoda, temelji predvsem na grafičnih metodah (ME-Plot, graf kvantilov). Med negrafične metode uvrščamo na primer funkcijo "findThreshold", ki se uporablja v statističnem programu R. Ta funkcija najde takšno mejno vrednost, da določeno število (npr. 5 % najvišjih opazovanj) leži nad to mejno vrednostjo. Kljub zelo dobrim rezultatom, ki jih zagotavljajo zgoraj našteje možnosti, je treba najti še natančnejše metode izbire mejne vrednosti.
- Analiza dogodkov, ki temelji na EVT, poudarja ekstremne dogodke in tako zanemara vpliv nizko in srednje porazdeljenih dogodkov.

7 ANALIZA PODATKOV

Proti koncu magistrskega dela prehajamo na empirični del magistrske naloge.

7.1 Izbira nabora podatkov

Za predmet naše raziskave smo izbrali delnice podjetja IBM – International Business Machines, ki se ukvarja z razvojem in izdelavo proizvodov ter storitev informacijske tehnologije. Podjetje IBM je bilo ustanovljeno leta 1888, uradno pa je bilo združeno sredi julija 1911. Sedež podjetja se nahaja v New Yorku, v mestu Armonk. Podjetje je znano predvsem po tem, da je leta 1981 na trg plasiralo prvi osebni računalnik, ki je bil dostopen širši množici ljudi, sicer pa je bilo ustanovljeno pred skoraj stotimi leti. Danes podjetje zaposluje več kot 330.000 ljudi v več kot 170 državah in velja za vodilno podjetje na področju sistemov in programske opreme. Svojim strankam ponuja celovite rešitve, tako svetovanje na področju tehnične arhitekture sistemov kot različne strateške analize in implementacije sistemov informacijske tehnologije. Podjetje IBM ima tudi v Sloveniji svojo hčerinsko podjetje in je na slovenskem trgu prisotno že več kot 70 let, uradno pa je bilo ustanovljeno leta 1992. V IBM Slovenija je danes zaposlenih več kot 270 ljudi, ki se ukvarjajo s tehnološkimi rešitvami na različnih področjih

poslovanja, kot so bančništvo in zavarovalništvo, zdravstvo, telekomunikacije, vladne in javne ustanove ter proizvodnja in distribucija.

Podatke za vrednosti IBM-ove delnice smo pridobili iz spletne strani "Yahoo Finance" (2012). Opazovano obdobje obsega 51 let, od začetka leta 1962 do konca leta 2012. Za ceno smo izbrali prilagojeni zaključni tečaj "Adjusted Close Price". To je zaključni tečaj "Close Price", prilagojen za izplačilo dividend in za cepitve delnic (angl. *stock split*).

Definicija cepitve delnic je bila pridobljena iz spletne strani Ljubljanske borze (2012) in se glasi: "Cepitev delnic je metoda, s katero se vsako delnico razdeli na več delnic (glede na izbrano razmerje), pri čemer se struktura kapitala ne spremeni, namen pa je povečati likvidnost delnic oziroma povečati dostopnosti delnic širšemu krogu vlagateljev." V našem primeru IBM-ove delnice se je denimo leta 1999 zgodila cepitev delnic 2 : 1, kar pomeni, da se je tečaj delnice razpolovil, število delnic na trgu pa se je povečalo za dvakratnik prvotnega števila delnic. Tržna kapitalizacija podjetja je torej ostala enaka.

Mednarodno podjetje IBM je v našem opazovanem obdobju naredilo 7 cepitev delnic, zadnja se je zgodila leta 1999. Najbolj pogosta cepitev je bila 2 : 1, sledi ji dvakratna cepitev 5 : 4 ter cepitvi 4 : 1 in 3 : 2. Že iz zgoraj omenjene definicije cepitve delnic lahko sklepamo, da cepitev delnic povzroči nenaden skok tečaja, ki pa ni rezultat tržnih reakcij, ampak akcije podjetja. Zato smo kot pravo zaključno vrednost tečaja, kot že zgoraj omenjeno, izbrali prilagojeni zaključni tečaj ("Adjusted Close Price"), ki upošteva izplačilo dividend ter cepitev delnic.

7.2 Osnovni ali logaritmirani podatki

Alexander (2008) podrobneje predstavlja nekaj že splošno znanih in nekaj malo manj znanih dejstev o donosih, ki so opisani v tem podglavju.

Najprej povzemimo definicijo logaritemskega donosa.

Logaritemski donos je logaritem iz kvocienta vrednosti na koncu opazovanega obdobja in vrednosti na začetku:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right), \quad (38)$$

kjer je r_t dnevni donos, P_t vrednost današnjega donosa in P_{t-1} vrednost včerajšnjega donosa.

Vemo, da iz razvoja funkcije $\ln(1 + x)$ v Taylorjevo vrsto sledi: $\ln(1 + x) \approx x$ za majhne x . Če to apliciramo na donose, dobimo:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln\left(\frac{P_t - P_{t-1}}{P_{t-1}} + 1\right) \approx \frac{P_t - P_{t-1}}{P_{t-1}} = R_t, \quad (39)$$

kjer je R_t logaritemski donos.

S pomočjo te aproksimacije lahko sklepamo, da sta za majhne donose navaden donos in logaritemski donos enaka. Logaritemski donos predstavlja stopnjo rasti v odstotkih. V praksi to pomeni, da se zgornja aproksimacija uporablja predvsem za dnevne donose.

Velja tudi, da je k -dnevni logaritemski donos vsota k zaporednih dnevnih donosov:

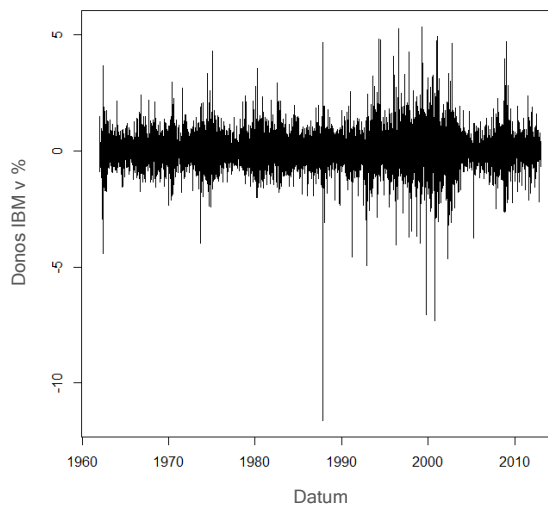
$$r_{kt} = \sum_{i=0}^{k-1} r_{t-i}. \quad (40)$$

Poleg teh lastnosti obstaja še več argumentov, ki utemeljujejo uporabo logaritemskih donosov pred navadnimi donosi, vendar to ni predmet obravnave naše magistrske naloge. Po mnenju Carol Alexander (2008) in vrste ostalih avtorjev, ki to potrjujejo, je tudi v našem primeru izbira logaritemskih donosov pravilna. Zato teorijo ekstremnih vrednosti uporabljamo na logaritemskih dnevni donosih IBM-ove delnice. Obdobje opazovanja je od 2. 1. 1962 do 31. 12. 2012. Naš vzorec tako obsega 12.836 opazovanj. Izključeni so netrgovalni dnevi – vikendi in prazniki. Vrednosti so z izbiro logaritemskih donosov merjene v odstotkih.

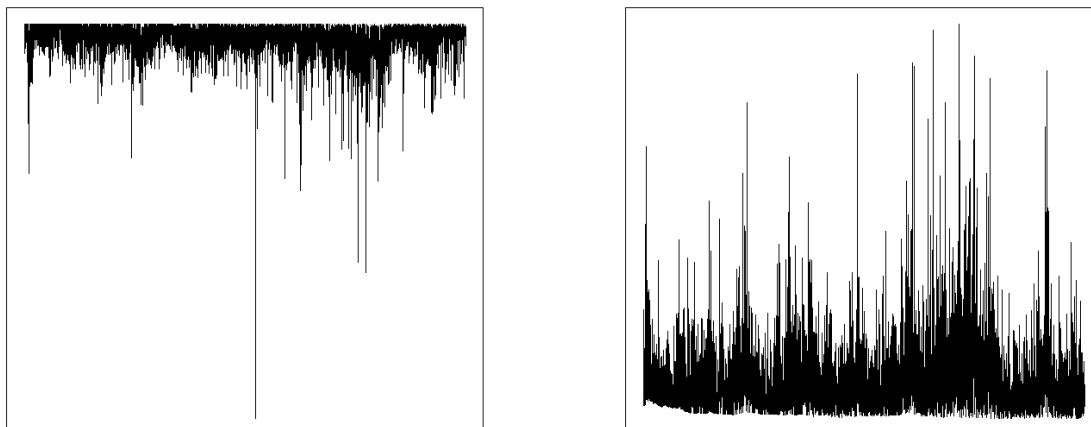
7.3 Podatki – delnica IBM

Najprej si bomo pogledali gibanje donosov IBM-ove delnice skozi čas in si tako ustvarili nekakšno širšo predstavo o naših podatkih. S Slike 5 je jasno razviden t. i. črni ponedeljek oziroma zlom finančnega trga oktobra 1987, ko je dnevni padec delnice IBM-a znašal 11,6 %. Če izključimo ta zlom, so se dnevni logaritemski donosi gibali med $-7,3\%$ in $5,4\%$. Na Sliki 5 so uprizorjeni vsi donosi IBM-ove delnice skozi opazovano obdobje, na Sliki 6 pa so prikazani ločeni donosi, negativni (levo) in pozitivni (desno).

Slika 5: Logaritemski dnevni donosi IBM-ove delnice v %, za obdobje 1962–2012



Slika 6: Logaritemski negativni (levo) in pozitivni (desno) dnevni donosi IBM-ove delnice v %, za obdobje 1962–2012



Kot prvo informacijo si bomo pogledali nekaj osnovnih statistik ter histogram dnevnih logaritemskih donosov IBM-ove delnice.

Tabela 1: Osnovne statistike za donos delnice IBM

	Št. opazovanj	Min v %	Max v %	Mat. upanje v %	Std. odklon v odst.točk	Asim.	Splošč.
IBM	12.836	-11,634	5,369	0,015	0,701	-0,281	12,348

Osnovne statistike za podatke o dnevnih donosih delnice IBM nam povejo, da imamo 12.836 opazovanj, minimalna opazovana vrednost znaša -11,6 %, največja pa 5,4 %. Vidimo, da so podatki asimetrični v levo, saj je asimetričnost (angl. *skewness*) negativna. Asimetričnost v levo pomeni, da je levi rep daljši od desnega. Za podatke, ki imajo asimetričnost enako nič, rečemo, da so normalno porazdeljeni. Naši podatki za asimetričnost so sicer negativni, vendar blizu ničle, torej bi lahko predpostavili, da imajo dolg levi rep in so skoraj normalno porazdeljeni. Ko pa pogledamo podatke za sploščenost (angl. *kurtosis*), predpostavko o normalni porazdelitvi ovržemo, saj bi bila sploščenost za normalno porazdelitev enaka 3, pri naših podatkih pa je dobrih 12. Visoka sploščenost nam pove, da imamo opravka z debelimi repi. Torej, če povzamemo: naši podatki kažejo, da niso porazdeljeni normalno, ampak da imajo dolg in debel levi rep.

V statističnem programu R smo uporabili več dodatnih paketov, ki so prilagojeni ravno za računanje ekstremnih vrednosti. Ti paketi pri računanju ekstremnih vrednosti uporabljajo pozitivne maksimalne vrednosti oziroma vrednosti v desnem repu porazdelitve. Torej, če imajo naši podatki debel levi rep, bomo za ohranjanje dolge finančne pozicije (tj. nakup vrednostnega papirja v pričakovanju, da bo njegova cena zrasla) in za zadostitev zgornjega pogoja uporabili negativne dnevne logaritemske donose:

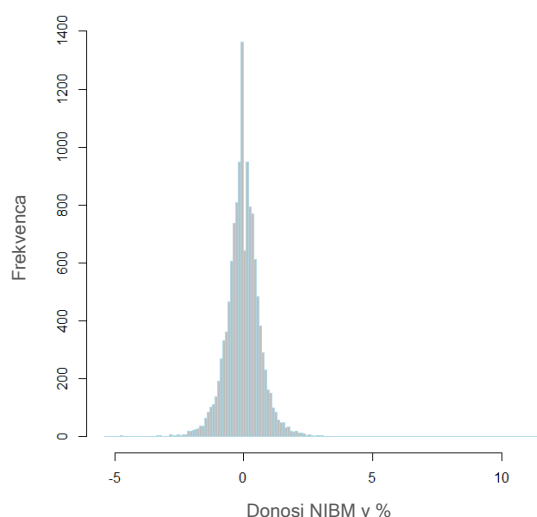
$$nibm = -ibm \quad (41)$$

Tabela 2: Osnovne statistike za negativen donos delnice IBM

	Št. opazovanj	Min v %	Max v %	Mat. upanje v %	Std. odklon v odst.točk	Asim.	Splošč.
NIBM	12.836	-5,369	11,634	-0,015	0,701	0,281	12,348

Iz Tabele 2 osnovnih statistik za "nibm" sedaj vidimo, da sta se minimalni in maksimalni vrednosti obrnili, povprečje oziroma matematično upanje je dobilo negativen predznak, asimetričnost pa pozitivnega, kar je natanko tisto, kar smo hoteli – asimetričnost v desno. Tudi v nadaljevanju raziskave bomo uporabljali le negativne dnevne logaritemske donose, zato tega ne bomo vsakič znova eksplicitno navajali. Te negativne logaritemske donose bomo obravnavali kot pozitivna števila in bodo nosili oznako "nibm". V primeru, ko bomo obravnavali ekstremne dogodke, le-ti zavzemajo vrednosti v desnem repu porazdelitve, zato je izbira negativnih logaritemskih donosov upravičena.

Slika 7: Histogram negativnih dnevnih logaritemskih donosov IBM-ove delnice



Če pogledamo histogram, porazdelitev z desnim debelim repom ni tako očitna. Zaradi asimetričnosti, ki je blizu 0, na prvi pogled zgleda precej podobno normalni porazdelitvi, vendar bodo nadaljnje raziskave pokazale nasprotno.

7.4 Razvoj rekordov

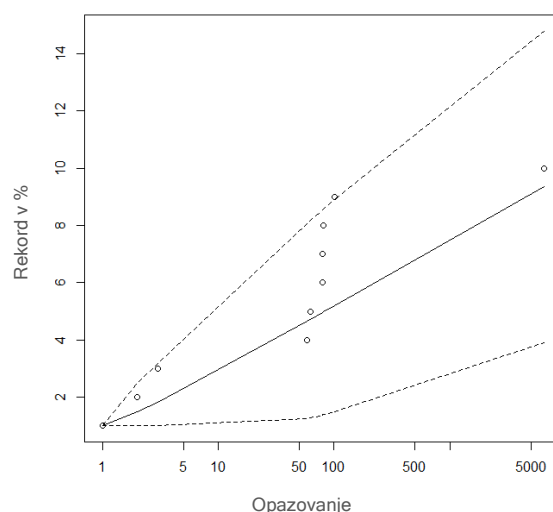
Ker nas zanima predvsem najslabši možen izid (angl. *worst case scenario*), si bomo pogledali podatkovni in grafični okvir razvoja rekordov negativnih dnevnih logaritemskih donosov IBM-a in njihova pričakovana obnašanja. Predpostavili bomo, da so opazovani podatki n.e.p. V Tabeli 3 je prikazanih 10 največjih rekordov, njihova vrednost (vrednost negativnega donosa IBM-ove delnice), opazovanje, v katerem se je rekord zgodil ter pričakovano matematično upanje in standardni odklon v primeru n.e.p.-podatkov. Opazimo, da je zadnji, deseti rekord največji in znaša 11,6 %, zgodil pa se je v 6.483. opazovanju.

Tabela 3: Prikaz prvih desetih rekordov, njihova vrednost, število opazovanja, ko se je rekord zgodil ter matematično upanje in standardni odklon opazovanega rekorda v %

Število	Vrednost Rekorda v %	Opazovanje	Matematično upanje v %	Std. odklon v odst. točk.
1	0,343	1	1	0
2	0,870	2	1,500	0,500
3	0,888	3	1,833	0,687
4	0,911	58	4,646	1,737
5	1,138	63	4,728	1,760
6	1,169	79	4,953	1,822
7	1,201	80	4,965	1,826
8	2,938	81	4,978	1,829
9	4,412	101	5,197	1,887
10	11,6344	6483	9,354	2,777

Podobno opazimo tudi na Sliki 8, kjer so podatki iz zgornje tabele uprizorjeni še grafično. Na x-osi so predstavljena opazovanja, na y-osi pa višina, ki jo je posamezni rekord dosegel. Črtkana črta prikazuje 95 % interval zaupanja za razvoj rekordov, če predpostavimo, da so podatki n.e.p.

Slika 8: Grafični prikaz razvoja prvih desetih rekordov delnice IBM



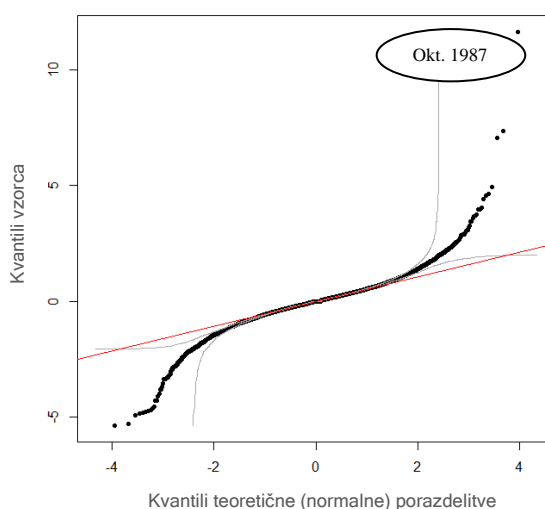
7.5 Graf kvantilov celotnega vzorca

V teoretičnem delu magistrske naloge smo definirali kaj je graf kvantilov in čemu je namenjen. Gre za zelo uporabno grafično orodje, ki nam pove, ali je določen vzorec res porazdeljen, tako kot želimo. To bomo preverili tudi na naših opazovanih podatkih. Zanimalo nas bo, ali se podatki porazdeljujejo po normalni porazdelitvi, kot na prvi pogled nakazuje histogram. Na to vprašanje si lahko odgovorimo, če pogledamo Sliko 9. Na x-osi⁴ so uprizorjeni kvantili teoretične porazdelitve (normalne porazdelitve), na y-osi pa kvantili opazovanega vzorca. Če bi

⁴ V statističnem programu R je lahko definicija grafa kvantilov malo drugačna, kot smo jo zapisali v podglavju **Error! Reference source not found.** - lahko ima zamenjani osi x in y, vendar to ne spremeni namembnosti tega grafa.

bili opazovani podatki porazdeljeni normalno, bi se točke prilegale rdeči premici, tako pa opazimo, da se na repih porazdelitev odcepi. To nakazuje na porazdelitev debelih ali suhih repov. Če nastane konkavni vzorec, potem ima porazdelitev suh rep, če pa nastane konveksni vzorec, ima porazdelitev debel rep. Sklepamo lahko, da ima naša porazdelitev debel tako desni kot tudi levi rep. Osredotočili se bomo na desni rep, saj nas zanimajo največji padci donosov delnice IBM.

Slika 9: Graf kvantilov celotnega opazovanega vzorca donosov delnice IBM, enote v %



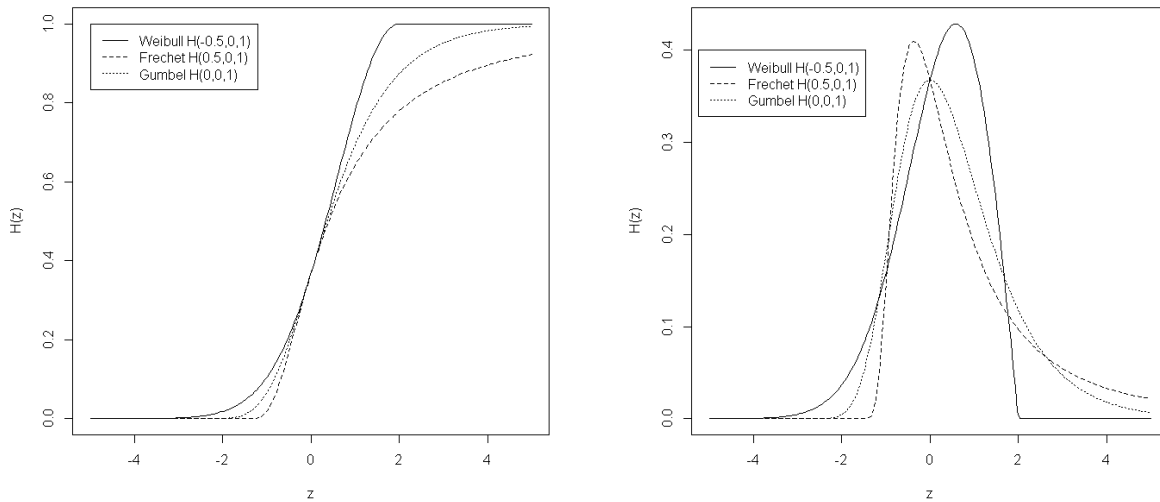
Očitno je, da imajo donosi IBM-ove delnice občutno debelejši rep od normalne porazdelitve, kar nakazuje na Fréchetovo družino GEV-porazdelitev s $\xi > 0$. Z grafa lahko razberemo grobo oceno, da se konveksni vzorec oziroma debel desni rep prične nekje okoli 1 %.

7.6 Prvi parametrični pristop

Ugotovili smo, da se opazovani donosi ne porazdeljujejo po normalni porazdelitvi, zato je treba ugotoviti, katera je tista porazdelitev, ki ji ustrezajo. Treba je najti ustrezno prileganje porazdelitve v repu. Najprej bomo to prileganje poiskali s prvim parametričnim pristopom v EVT in kasneje še z drugim. Če se bodo podatki prilegali katerikoli izmed porazdelitev EVT, bomo z metodo največjega verjetja ocenili tudi parametre porazdelitev.

Na spodnji Sliki 10 si lahko ogledamo primer kumulativne porazdelitvene funkcije H_ξ (levo) in funkcije gostote verjetnosti h_ξ (desno) za Fréchetovo ($\xi = 0,5$), Weibullovo ($\xi = -0,5$) in Gumbelovo ($\xi = 0$) porazdelitev. Uprizorjen je naključni vzorec "z", ki vsebuje 200 opazovanj, katerega vrednosti se gibljejo med -5 in 5 . Fréchetova porazdelitev je definirana le za vrednosti $z > -2$ ter Weibullova le za $z < 2$. Prva vrednost katerekoli porazdelitve predstavlja parameter oblike ξ , druga lokacijski parameter μ in tretja parameter višine σ .

Slika 10: Primer kumulativne porazdelitvene funkcije (levo) in verjetnostne gostote (desno) za vse tri GEV-porazdelitve



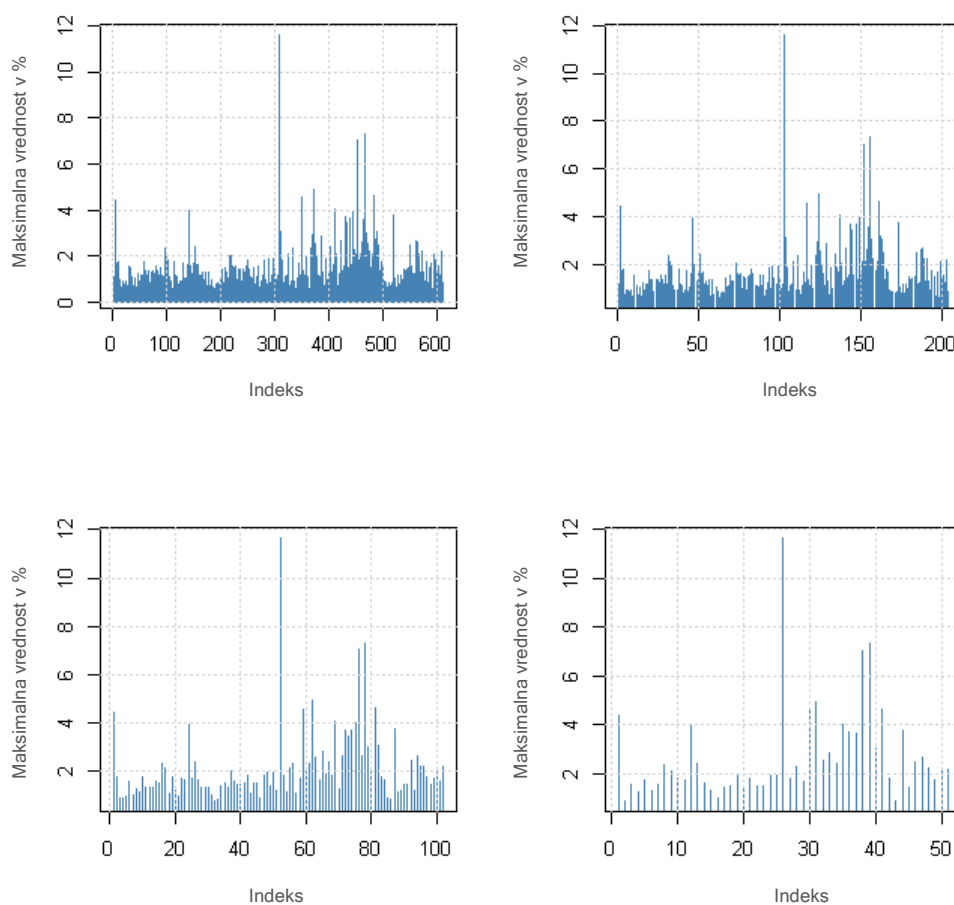
7.6.1 GEV-porazdelitev in metoda maksimumov skupin podatkov

Iz Fisher-Tippettovega izreka vemo, da podatki ustrezajo eni izmed GEV-porazdelitev. Ugotoviti moramo, kateri. Za ocenitev parametrov GEV-porazdelitve bomo uporabili metodo največjega verjetja na negativnih dnevni donosih IBM-ove delnice.

Na voljo imamo $T = 12.836$ opazovanj donosov IBM-ove delnice. Ta opazovanja razdelimo na k neprekrivajočih se obdobjih oziroma skupin podatkov dolžine n . Veljati mora enačba $T = n * k$.

Predpostavimo, da nas zanimajo maksimumi letnih skupin podatkov. Število opazovanj v enem obdobju je potem enako 252 - število delovnih dni v letu. Ker vemo, da imamo na voljo 12.836 opazovanj in da eno obdobje obsega 252 opazovanj, potem je število letnih obdobjih enako 51. Enako lahko storimo, če nas zanima drugačna dolžina obdobja. Če želimo maksimume mesečnih skupin podatkov, potem je število opazovanj v mesecu enako 21, število mesecev pa 612. Za četrtnetne skupine podatkov imamo 63 opazovanj v enem četrtnetju in 204 četrtnetij. Za pol-letne skupine podatkov pa je 126 opazovanj v enem polletju in 102 različnih polletij. Seveda lahko uporabimo katerokoli dolžino obdobja, vendar krajše kot je obdobje, več podatkov imamo na voljo za nadaljnjo analizo. To je dobro razvidno s Slike 11, kjer vidimo, kako se zmanjšuje število podatkov, ko povečujemo dolžino obdobja. Te izbrane dolžine obdobjih bomo opazovali tudi v nadaljevanju analize metode maksimumov skupin podatkov oziroma prve parametrične metode.

Slika 11: Maksimalne vrednosti negativnih donosov delnice IBM v % za različno izbrana obdobja: mesečno obdobje (levo zgoraj), četrtno (desno zgoraj), polletno (levo spodaj) in letno obdobje (desno spodaj)



Najprej se bomo osredotočili na maksimalne vrednosti negativnih donosov za mesečne skupine podatkov. Kot že omenjeno, je število opazovanj v enem mesecu enako 21. Uporabljena je najvišja vrednost v vsakem mesecu in ker imamo na voljo 12.836 opazovanj, dobimo skupaj 612 maksimalnih vrednosti. Prvi pristop teorije ekstremnih vrednosti predlaga, da GEV-porazdelitev prilegamo tem 612 maksimalnim vrednostim. Tako dobimo ocene vseh treh parametrov ξ , μ in σ , ki so ocenjeni po metodi največjega verjetja.

Tabela 4: Ocenjeni parametri GEV-porazdelitve za mesečne skupine podatkov

Število vseh opazovanj	12.836		
Število maksimalnih vrednosti	612		
Število opazovanj v izbranem obdobju	21		
Ocene parametrov	$\hat{\xi} = 0,239$	$\hat{\sigma} = 0,403$	$\hat{\mu} = 0,842$
Asimp. standardna napaka ocen parametrov	$SE(\hat{\xi}) = 0,032$	$SE(\hat{\sigma}) = 0,015$	$SE(\hat{\mu}) = 0,018$

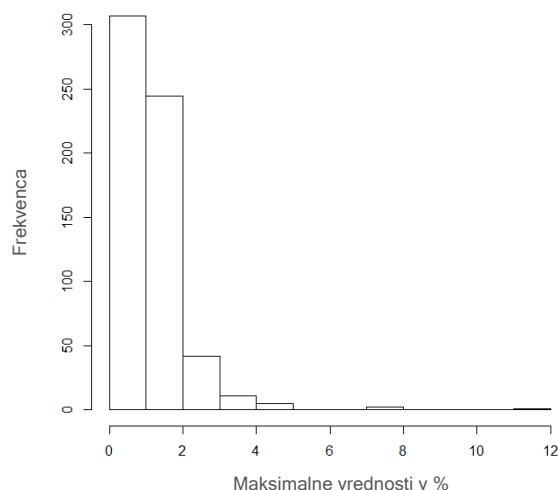
Rezultat nam pokaže število vseh možnih opazovanj (12.836), število maksimalnih opazovanj, če nas zanima mesečno obdobje (612), število opazovanj v vsakem obdobju (21), ocene

vseh treh parametrov (ξ, σ, μ) po metodi največjega verjetja ter njihovo asimptotično standardno napako.

Vidimo, da je ocena za ξ po metodi največjega verjetja enaka 0,239 z asimptotično standardno napako $SE(\hat{\xi}) = 0,032$. Tako lahko izračunamo tudi asimptotični 95 % interval zaupanja za $\hat{\xi}$. Ta zavzema vrednosti med $[0,175; 0,303]$. Ker je $\hat{\xi}$ in celoten interval zaupanja za ξ pozitiven, to nedvomno kaže na porazdelitev debelega repa.

Za dodatno informacijo lahko pogledamo histogram teh 612 maksimalnih vrednosti, ki smo jih dobili z izbiro mesečnega obdobja. Histogram je uprizorjen na Sliki 12 in prav tako prikazuje porazdelitev z debelim repom.

Slika 12: Histogram maksimalnih izgub donosov delnice IBM na podlagi izbire mesečnih skupin podatkov



Sedaj, ko smo ocenili parametre GEV-porazdelitve, lahko s pomočjo spodnjih dveh slik preverimo, ali je GEV-porazdelitev res pravilna izbira porazdelitve za naše opazovane podatke.

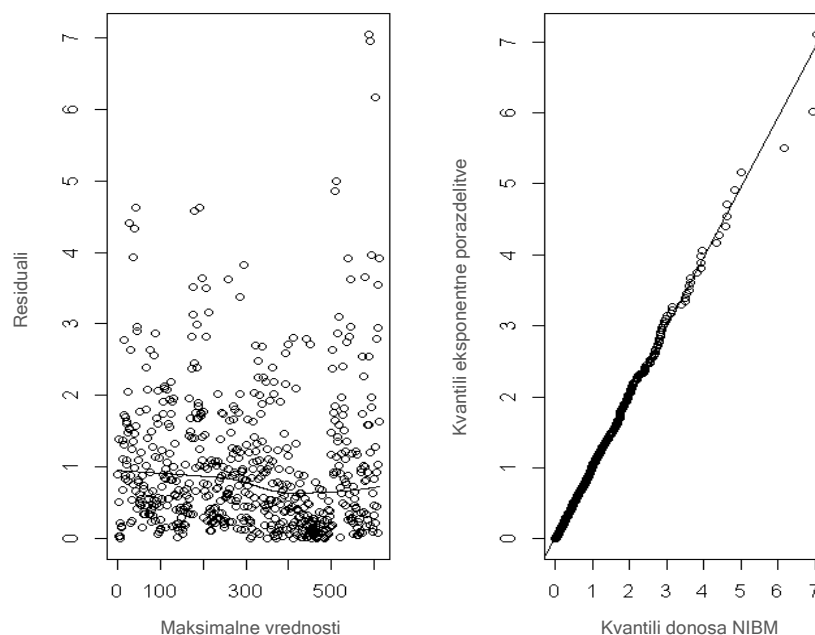
Prvi graf na Sliki 13 prikazuje razsevni diagram residualov⁵, ki smo jih dobili s prileganjem GEV-porazdelitve opazovanim podatkom. Vidimo, da ne odražajo avtokorelacije ali heteroskedastičnosti v podatkih. Drugi graf na Sliki 13 pa prikazuje graf kvantilov s teoretično eksponentno porazdelitvijo. Residuali GEV-porazdelitve, katero smo prilegali opazovanim podatkom mesečnih skupin podatkov, so definirani kot

$$W_i = \left(1 + \xi \frac{M_n^{(i)} - \hat{\mu}}{\hat{\sigma}} \right)^{-\frac{1}{\xi}}. \quad (42)$$

Podatki so sedaj preurejeni v n.e.p.-eksponentno porazdeljene slučajne spremenljivke z ničelno hipotezo, da se prilegajo GEV-porazdelitvi. Podatki se dobro prilegajo premici, torej je naša predpostavka o izbiri GEV-porazdelitve upravičena.

⁵ Residual predstavlja razliko med opazovano in ocenjeno vrednostjo.

Slika 13: Razsevni diagram residualov in graf kvantilov, pridobljenih s prileganjem GEV-porazdelitve opazovanim podatkom



Z uporabo ocen parametrov po metodi največjega verjetja, ki smo jih dobili v zgornjem prileganju GEV-porazdelitve mesečnim skupinam podatkov, lahko sedaj odgovorimo na spodnje vprašanje:

- Kolikšna je verjetnost, da bo naslednji mesec največji negativni donos presegel vse prejšnje negativne donose? Z drugimi besedami, kolikšna je verjetnost, da bo naslednji mesec dosežen nov rekord (največji negativni donos)?

Na to vprašanje lahko odgovorimo matematično. Število maksimumov, če opazujemo mesečno obdobje, znaša 612. Zanima nas, ali bo naslednji, 613 maksimum večji od vseh ostalih, ob dejstvu, da vemo, da naš največji maksimum znaša 11,6 %. To lahko zapišemo kot:

$$Pr\left(M_{21}^{(613)} > \max\left(M_{21}^{(1)}, \dots, M_{21}^{(612)}\right)\right) = 1 - H_{\hat{\xi}, \hat{\mu}, \hat{\sigma}}(11,6). \quad (43)$$

Odgovor na zgornje vprašanje se glasi: obstaja 0,023 % verjetnost, da se bo naslednji nov rekord zgodil v naslednjem mesecu. Ta verjetnost je sicer zelo majhna, vendar z daljšanjem obdobja dobimo tudi večjo verjetnost, kar bomo videli v nadaljevanju.

Zgoraj smo opisali različno dolga obdobja, ki nas bodo zanimala. Zato bomo prilegali GEV-porazdelitev še maksimalnim vrednostim v četrletnih, polletnih in letnih skupinah podatkov. Tudi tu smo uporabili metodo največjega verjetja. V spodnji tabeli so združeni podatki in ocene parametrov za različno dolga izbrana obdobja.

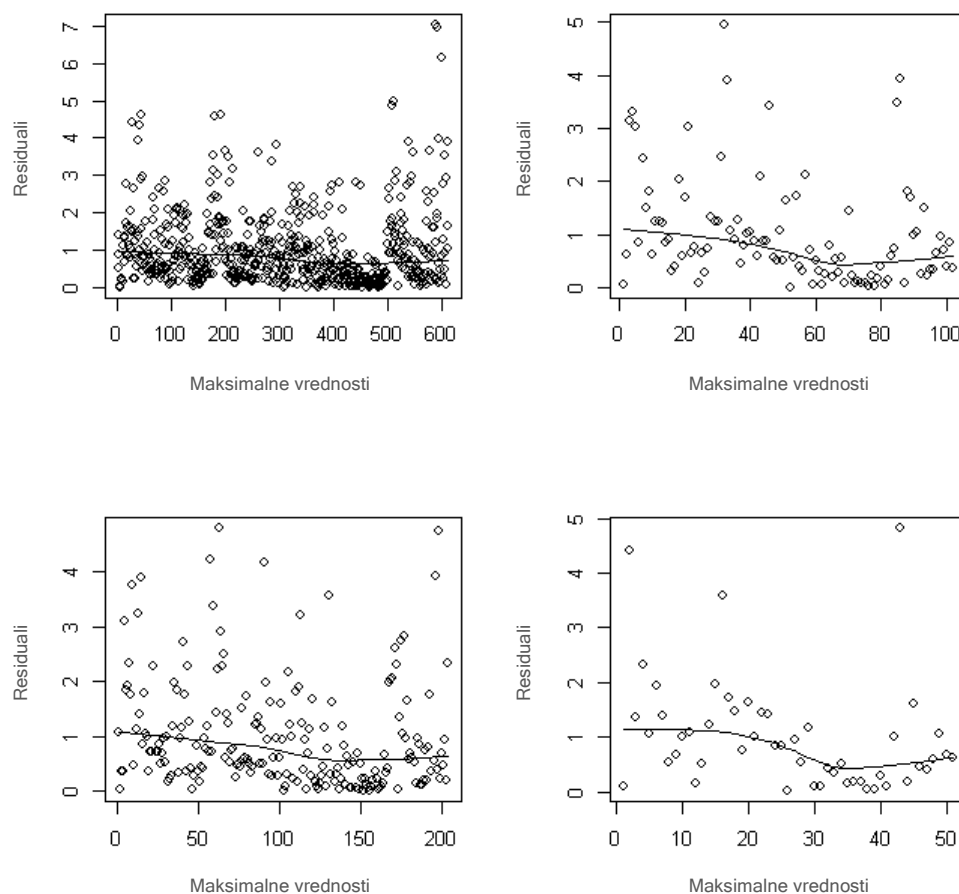
Tabela 5: Ocena parametrov za vsa obdobja (standardne napake v oklepajih)

Obdobje	Št. opazovanj v izbranem obdobju – n	Št. maksimalnih vrednosti – k	Parameter oblike – $\hat{\xi}_n$	Parameter višine – $\hat{\sigma}_n$	Lokacijski parameter – $\hat{\mu}_n$
Mesec	21	612	0,239 (0,032)	0,403 (0,015)	0,842 (0,018)
Četrtletje	63	204	0,372 (0,068)	0,469 (0,033)	1,168 (0,038)
Polletje	126	102	0,378 (0,093)	0,584 (0,058)	1,456 (0,066)
Leto	252	51	0,373 (0,121)	0,772 (0,108)	1,794 (0,122)

Iz Tabele 5 je razvidno, da parameter višine in lokacijski parameter postopoma naraščata, ko se n povečuje. Ta rezultat je pričakovan, saj je število maksimalnih vrednosti v opazovanih obdobjih nepadajoča funkcija v odvisnosti od n . S povečevanjem števila n se zmanjšuje k in posledično se povečujeta parameter višine in lokacijski parameter. Ocene parametra oblike oziroma repnega indeksa so stabilne za $n \geq 63$, torej za četrtletna, polletna in letna obdobja, vrednosti pa znašajo okoli 0,37. Rezultati, kjer za opazovano obdobje vzamemo leto ($n = 252$), imajo višjo volatinito oziroma standardno napako, saj se število obdobjev zmanjša na 51 let, prej pa smo imeli na primer 204 četrtletnih obdobjev. Asimptotični 95 % interval zaupanja za $\hat{\xi}$ znaša za mesečno obdobje [0,175; 0,303], za četrtletno obdobje [0,236; 0,508], za polletno obdobje [0,192; 0,564] in za letno obdobje [0,131; 0,615]. Vsi intervali zaupanja vsebujejo samo pozitivne vrednosti za $\hat{\xi}$, kar kaže na porazdelitev debelega repa ne glede na izbrano dolžino obdobja.

Prav tako lahko pogledamo razsevni diagram residualov in graf kvantilov za različno dolžino obdobjev (kot smo to storili pri mesečnih obdobjih) in tako preverimo ali je GEV-porazdelitev pravilna izbira porazdelitve. Vemo, da daljše kot je obdobje, manj podatkov imamo na voljo.

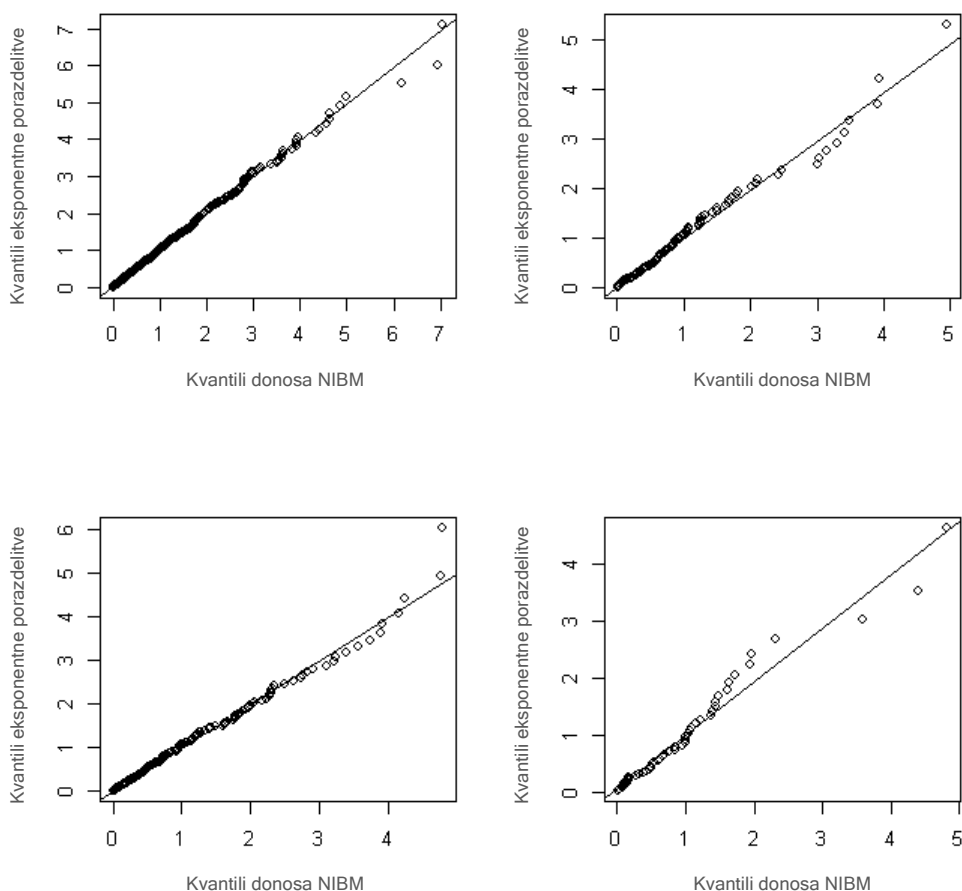
Slika 14: Razsevni diagrami residualov: mesečno obdobje (levo zgoraj), četrtno obdobje (levo spodaj), polletno obdobje (desno zgoraj) in letno obdobje (desno spodaj)



Slika 14 nam pokaže, da porazdelitve residualov ne kažejo avtokorelacije ali heteroskedastičnosti v podatkih kljub različni dolžini obdobj.

Na Sliki 15 opazimo, da skoraj vsa opazovanja ležijo blizu linearne črte ne glede na izbrano obdobje. To nam pove, da se GEV-porazdelitev dobro prilega opazovanim podatkom za katerokoli izbrano dolžino obdobj.

Slika 15: Graf kvantilov: mesečno obdobje (levo zgoraj), četrletno obdobje (levo spodaj), polletno obdobje (desno zgoraj) in letno obdobje (desno spodaj)



Sedaj, ko imamo vse parametre ocenjene in smo pokazali, da je uporaba GEV-porazdelitve upravičena, nas najbolj zanima odgovor na vprašanje, ali bo v naslednjem opazovanem obdobju dosežen nov rekord. Spomnimo se, da je bila verjetnost, da bo v naslednjem mesecu dosežen nov rekord, enaka 0,023 %. Tako verjetnostno napoved lahko z metodo maksimumov skupin podatkov naredimo tudi za druga obdobja. Za podatke negativnih donosov IBM-ove delnice je rezultat sledeč:

- obstaja 0,25 % verjetnost, da se bo naslednji nov rekord zgodil v naslednjem četrletju.
- obstaja 0,47 % verjetnost, da se bo naslednji nov rekord zgodil v naslednjega pol leta.
- obstaja 0,91 % verjetnost, da se bo naslednji nov rekord zgodil v naslednjem letu.

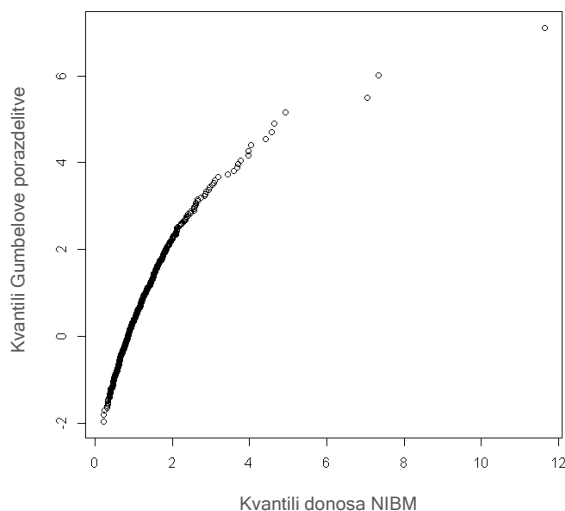
Kot smo pričakovali, daljše kot je obdobje, večja je verjetnost, da se bo naslednji rekord zgodil v naslednjem tako dolgem obdobju.

7.6.2 Graf kvantilov na podlagi prvega parametričnega pristopa

Na Sliki 15 smo videli, da je izbira GEV-porazdelitve upravičena za katerokoli izbrano dolžino obdobja. Sedaj pa nas zanima katera izmed družin GEV-porazdelitve je prava. Poglejmo si graf kvantilov, ki ima za teoretično porazdelitev izbrano Gumbelovo porazdelitev, za primer

dolžine obdobja pa vzemimo mesečne maksimalne vrednosti. Če podatki mesečnih maksimumov ustrezajo Gumbelovi porazdelitvi, bo graf izgledal kot premica, sicer izbrana porazdelitev ni prava.

Slika 16: Graf kvantilov maksimalnih vrednosti, pridobljenih z izbiro mesečnih skupin podatkov, kjer je za teoretično porazdelitev izbrana Gumbelova porazdelitev



Kot že omenjeno, zgornji graf kvantilov uporablja za teoretično porazdelitev Gumbelovo porazdelitev (porazdelitev suhega repa). Ta je definirana kot $\Lambda(x) = \exp\{-e^{-x}\}$, $x \in \mathbb{R}$. Inverz te funkcije, ki se drugače imenuje tudi kvantilna funkcija, je enak: $\Lambda^{-1}(p) = -\ln(-\ln(p))$. Količina $x_p = \Lambda^{-1}(p)$ definira p -ti kvantil funkcije Λ . Na x -osi so uprizorjeni kvantili maksimalnih vrednosti, pridobljeni z izbiro mesečnih skupin podatkov, na y -osi pa kvantili Gumbelove porazdelitve. Opazimo, da graf ni linearen, ampak konkaven, kar ponovno kaže na porazdelitev debelega repa s $\xi > 0$. Ustrezna družina porazdelitev je torej Fréchetjeva porazdelitev.

7.6.3 Stopnja poplačila na podlagi prvega parametričnega pristopa

V zgornjih podpoglavjih smo pokazali, da opazovani podatki opisujejo porazdelitev z debelim repom, za katero smo po metodi največjega verjetja tudi ocenili vse tri parametre porazdelitve. Sedaj lahko na podlagi teh parametrov ocenimo, s kakšno izgubo se lahko v prihodnosti soočamo. Ena izmed možnih metod za merjenje tveganja je stopnja poplačila. To je zelo uporabna mera tveganja, ko nas zanimajo maksimalne vrednosti v skupinah podatkov, in je povezana z visokim kvantilom. Stopnja poplačila $R_{n,k}$ je definirana kot stopnja, ki je presežena enkrat na vsake k obdobja dolžine n :

$$P\left(M_n^{(i)} > R_{n,k}\right) = \frac{1}{k}, \quad (44)$$

kjer je $M_n^{(i)}$ maksimum v skupini podatkov oziroma obdobju. Stopnja poplačila torej predstavlja pričakovano višino donosa, ki bo v povprečju presežena enkrat v k ($k = \frac{1}{p}$) obdobjih dolžine n . Obdobje, v katerem je stopnja poplačila presežena, se imenuje stresni interval (angl. *stress period*). V našem primeru se lahko stopnja poplačila imenuje tudi stopnja izgube donosa.

Na podlagi zgoraj ocenjenih parametrov za GEV-porazdelitev, ki ustrezajo mesečnim skupinam podatkov, lahko na primer odgovorimo na vprašanje:

- Kolikšna je 12-mesečna oziroma letna stopnja poplačila negativnih donosov IBM-ove delnice? Povedano drugače, kako visok bo negativni donos, ki bo v povprečju presežen enkrat letno?

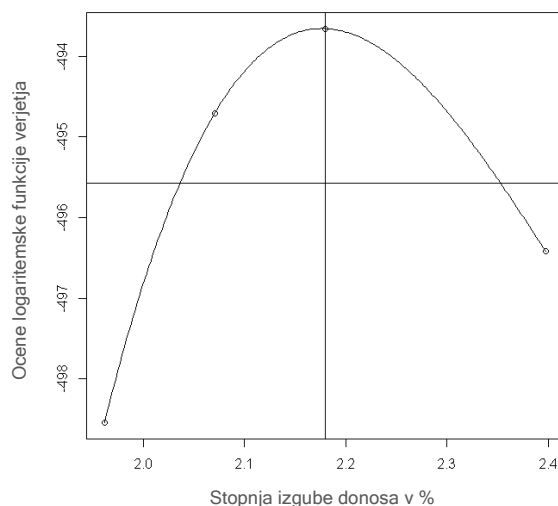
Že iz Podpoglavja 7.6.1 vemo, da velja formula $T = k * n$. Če opazujemo mesečne skupine podatkov, imamo v vsakem mesecu 21 opazovanj. Predpostavimo še, da nas zanima letna stopnja poplačila, torej opazujemo 12 mesecev in skupaj porabimo 252 opazovanj za vsako leto ($k = 252$). Vemo, da je število vseh opazovanj enako $T = 12.836$. Ker nas zanima letna stopnja poplačila (ta vsebuje 252 opazovanj), dobimo dolžino opazovanja enako 50 let: $\frac{T}{k} = \frac{12.836}{252} = 50,9$. Torej, v našem primeru je to stopnja izgube donosa, ki bo v povprečju presežena enkrat letno v dolžini naslednjih 50 let.

Tabela 6: Enoletna stopnja izgube donosa za mesečno izbrane podatke

Spodnji IZ	Ocenjena stopnja poplačila	Zgornji IZ
2,038	2,180	2,352

V zgornji Tabeli 6 je zapisana stopnja poplačila oziroma stopnja izgube donosa in njen 95 % interval zaupanja za mesečne maksimume, porazdeljene po GEV-porazdelitvi. Številka v prvem stolpcu predstavlja spodnji interval zaupanja, številka v drugem stolpcu ocenjeno stopnjo poplačila, številka v tretjem stolpcu pa zgornji interval zaupanja te ocenjene stopnje poplačila. Vidimo, da bo naša stopnja izgube donosa, ki bo presežena enkrat letno v dolžini naslednjih 50 let, znašala 2,18 %. Na Sliki 17 je grafično prikazana ocenjena stopnja poplačila ter njen levi in desni asimptotični 95 % interval zaupanja.

Slika 17: Letna stopnja izgube donosa za mesečne maksimume



Izračunamo lahko tudi letno stopnjo poplačila, ki jo dobimo iz ocenjenih parametrov GEV-porazdelitve, ki smo ji prilegali podatke za različna obdobja. Izračunali smo letni stopnji poplačila za četrtnete in polletne maksimume negativnih donosov IBM-ove delnice in znašata 1,9 % oziroma 1,7 %.

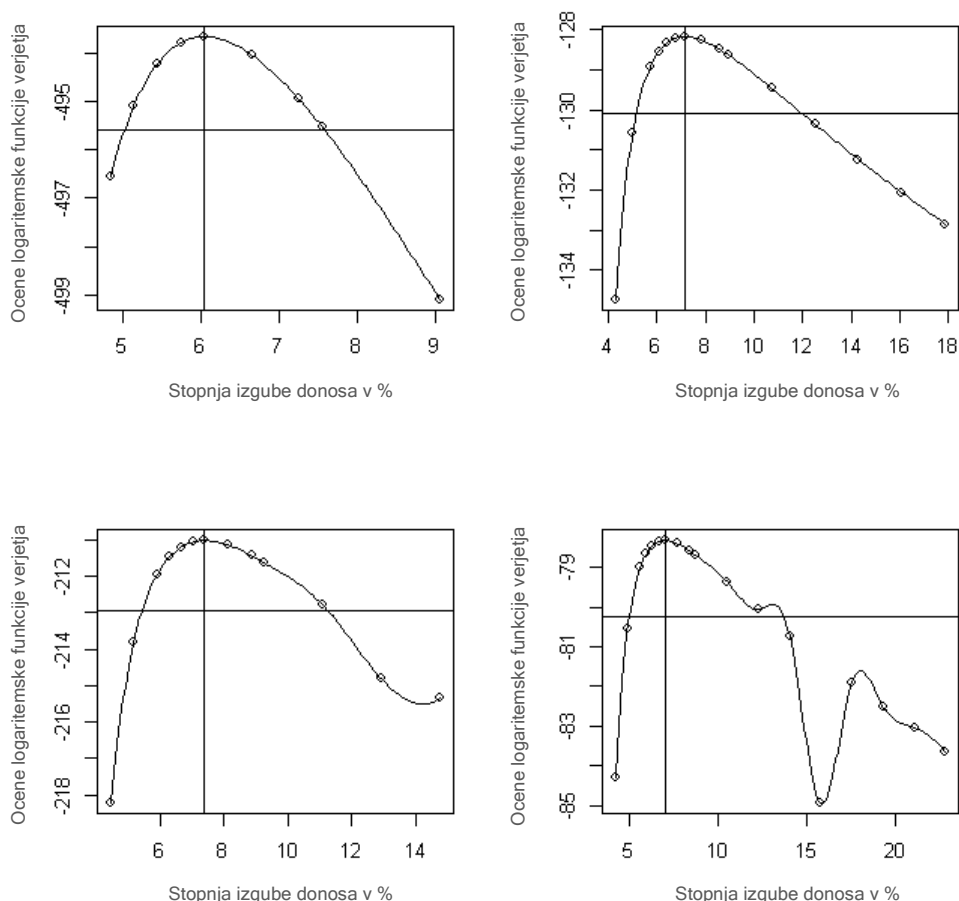
Še bolj kot enoletna stopnja poplačila nas zanimajo stopnje poplačila daljših obdobj. V spodnji tabeli si bomo kot primer ogledali 30-letno stopnjo poplačila. Tudi tu bomo razlikovali med različno dolgimi skupinami podatkov, za katere smo že v Tabeli 5 zapisali ocenjene parametre GEV-porazdelitve.

Tabela 7: 30-letna stopnja izgube donosa za negativne donose IBM-ove delnice

	Spodnji IZ	Ocenjena stopnja poplačila	Zgornji IZ
Mesečni maksimumi	5,029	6,047	7,560
Četrtni maksimumi	5,474	7,391	11,246
Polletni maksimumi	5,173	7,144	11,927
Letni maksimumi	5,064	7,034	13,681

Opazimo, da daljše, kot je obdobje opazovanja maksimumov, višja je ocenjena stopnja poplačila in bolj širok je 95 % interval zaupanja. Te podatke smo na Sliki 18 ponazorili tudi grafično.

Slika 18: 30-letna stopnja poplačila za mesečne (zgoraj levo), četrtnne (spodaj levo), polletne (zgoraj desno) in letne (spodaj desno) maksimume



Opazimo, da sta ocena stopnje poplačila in njen 95 % interval zaupanja vedno bolj asimetrično oblikovana. Če pogledamo 30-letno stopnjo poplačila za letne maksimume, vidimo, da ocenjena stopnja poplačila znaša 7 %, vendar njen zgornji 95 % interval zaupanja dosega sko-

raj še enkrat višje vrednosti - 13,7 %. Ta številka zglada relativno visoka, vendar če pogledamo oktobrski borzni zlom leta 1987, je ta dosežen skoraj pri isti vrednosti.

7.7 Drugi parametrični pristop

Namesto da se osredotočamo le na ekstreme v obdobjih, se bomo v drugem pristopu osredotočali na vse presežke čez neko vnaprej določeno mejo. To se imenuje "Peak Over Threshold Method" (POT-metoda). V našem primeru opazujemo negativne dnevne logaritemske donose delnic IBM, ki jih označimo z r_t . Tem negativnim donosom moramo vnaprej določiti višino mejne vrednosti u . Na primer, da se i -ti presežek zgodi ob času t_i (označeno z r_{t_i}), potem se POT-metoda osredotoča na podatke v območju $(t_i, r_{t_i} - u)$, kjer je $r_{t_i} \geq u$. Kot že prej omenjeno, POT-metoda ne zahteva izbire skupin podatkov dolžine n , ampak zahteva izbiro mejne vrednosti u . Različne izbire mejne vrednosti u vodijo k različnim ocenam parametra oblike ξ in posledično tudi repnega indeksa $\alpha = \frac{1}{\xi}$.

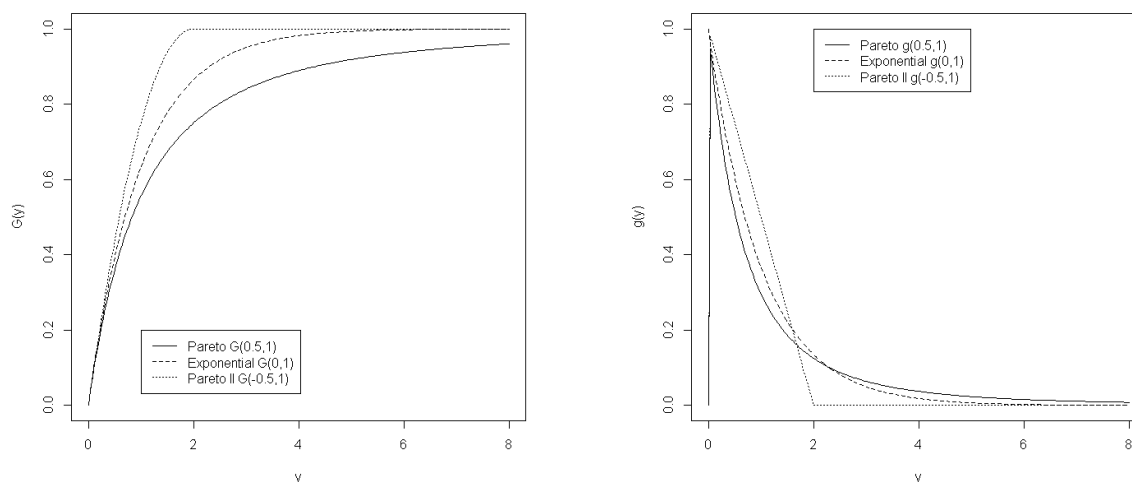
V prvem delu magistrske naloge smo zapisali ključni izrek POT-metode - Izrek Pickands-Balkema-de Haan, ki pravi, da za dovolj velik u velja aproksimacija $F_u(y) \approx G_{\xi, \sigma}(y)$ za širok razred porazdelitev F . Za implementacijo tega rezultata moramo najprej določiti višino mejne vrednosti u in potem oceniti neznan parametra ξ in σ posplošene Pareto porazdelitve. To bomo storili v naslednjih podpoglavjih.

7.7.1 Paretova porazdelitev

Najprej si bomo pogledali primer posplošene Pareto porazdelitve, saj drugi izrek EVT pravi, da lahko za dovolj visoko mejno vrednost u porazdelitveno funkcijo presežka aproksimiramo s posplošeno Pareto porazdelitvijo. Slednja je sestavljena iz dveh tipov Pareto porazdelitev in eksponentne porazdelitve, primeri vseh treh porazdelitev pa so prikazani na Sliki 19. Na levi sliki je uprizorjena kumulativna porazdelitvena funkcija vseh treh porazdelitev, na desni pa funkcija gostote verjetnosti. Izbrano je zaporedje y , ki predstavlja 200 naključno generiranih števil, ta pa zavzemajo vrednosti med 0 in 8. Posplošena Paretova porazdelitev (GPD-porazdelitev) je določena z dvema parametroma, ξ in σ . V našem primeru je za σ povsod izbrana vrednost 1, vrsta porazdelitve pa je odvisna od ξ . Paretova porazdelitev tipa I je v našem primeru uprizorjena s $\xi = 0,5$, eksponentna porazdelitev s $\xi = 0$, Paretova porazdelitev tipa II⁶ pa s $\xi = -0,5$.

⁶ Opomba: Paretovo porazdelitev tipa II smo v našem primeru definirali le za vrednosti $y < 2$.

Slika 19: Kumulativna porazdelitvena funkcija (levo) in funkcija gostote verjetnosti (desno) za Paretovo porazdelitev tipa I in tipa II ter eksponentno porazdelitev



7.7.2 Graf kvantilov na podlagi drugega parametričnega pristopa

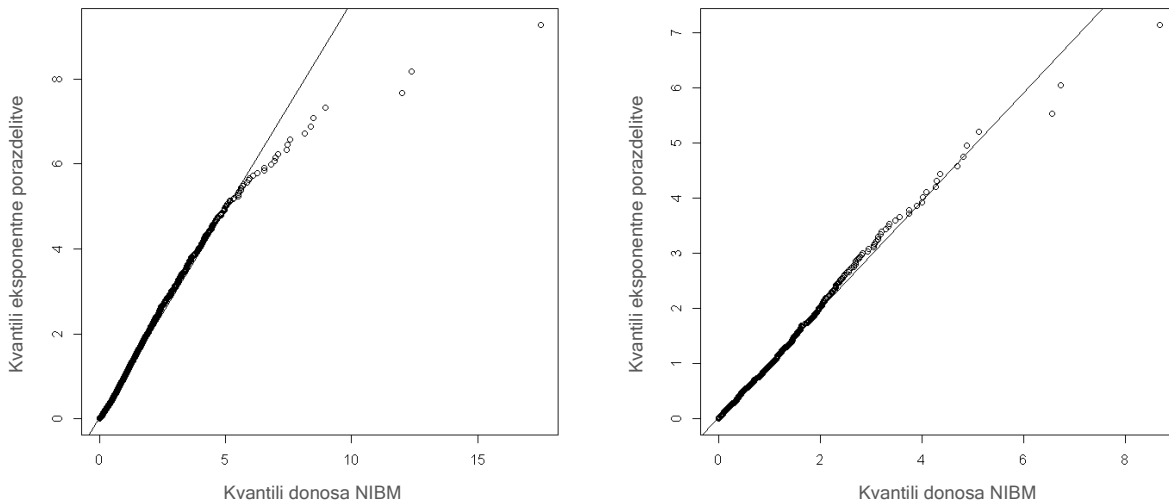
Že v Podpoglavju **Error! Reference source not found.** smo ugotovili, da je graf kvantilov enostavno grafično orodje, s katerim lahko sklepamo, kakšno je obnašanje repa opazovanih izgub vrednosti IBM-ove delnice. V tem primeru smo za teoretično porazdelitev izbrali eksponentno porazdelitev, torej nas bo zanimalo ali se opazovani podatki res porazdeljujejo po eksponentni porazdelitvi.

Po mnenju Coxa in Snella (1968) lahko definiramo residue GPD-porazdelitve, katero smo prilegali opazovanim podatkom, kot:

$$W_i = \frac{1}{\xi} \ln \left(1 + \xi \frac{X_i - u}{\hat{\sigma} + \hat{\xi}(u - \hat{\mu})} \right)^{-\frac{1}{\xi}}. \quad (45)$$

Za tako definirane residue W_i velja, da so eksponentno n.e.p. slučajne spremenljivke z ničelno hipotezo, da se prilegajo GPD-porazdelitvi (McNeil & Saladin, 2000). Če bomo izbrali pravilno mejno vrednost, ki bo določila parametre GPD-porazdelitve, se bodo podatki prilegali naraščajoči premici. Za mejno vrednost najprej izberemo 0,1 %, da vidimo, kakšno obliko naši podatki povzemajo. Opazimo, da se začenjajo podatki proti repu odklanjati od premice in dobivajo konkavno obliko (leva graf Slike 20), kar nakazuje na porazdelitev debelega repa. Sedaj, kot primer, za mejno vrednost izberemo še 1 % (desni graf Slike 20). Vidimo, da se podatki že bolje prilegajo premici in predpostavka o GPD-porazdelitvi je tako upravičena. Z izbiro mejne vrednosti 1 % bomo iz opazovanja izključili vse tiste vrednosti, ki se nahajajo pod to mejno vrednostjo.

Slika 20: Graf kvantilov za izbrano mejno vrednost $u = 0,1$ (levo) in mejno vrednost $u = 1$ (desno)

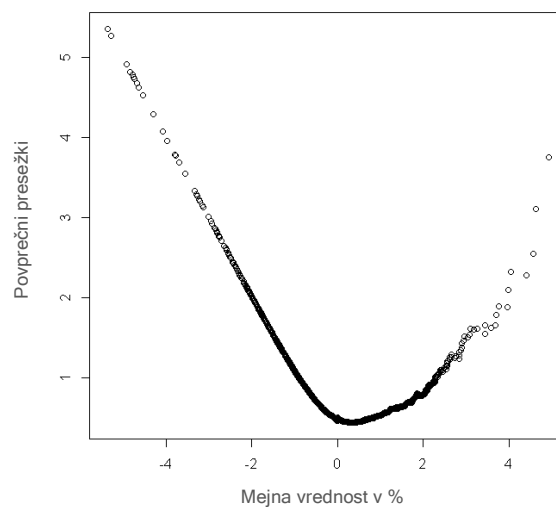


7.7.3 MEF na podlagi drugega parametričnega pristopa

Drugo grafično orodje, s katerim si lahko pomagamo pri sklepanju obnašanja porazdelitve v repu, je funkcija povprečnih odmikov (MEF) oziroma graf povprečnih odmikov (ME-Plot). Z njim bomo dodatno potrdili obstoj odebeljenega repa pri porazdelitvi opazovanih podatkov in z grafa ocenili mejno vrednost.

MEF ocenjuje funkcijo presežka, ki opisuje pričakovan presežek, ko se presežek že pojavi. Če so podatki iz GPD-porazdelitve s pozitivnim ξ (porazdelitev debelega repa), potem je MEF uprizorjen kot premica s pozitivnim naklonom nad neko določeno mejno vrednostjo u . Če podatki prihajajo iz porazdelitve, ki je iz Gumbelovega maksimalnega območja privlačnosti (porazdelitev suhega repa), dobimo premico z negativnim naklonom.

Slika 21: Graf povprečnih odmikov za negativne donose delnice IBM

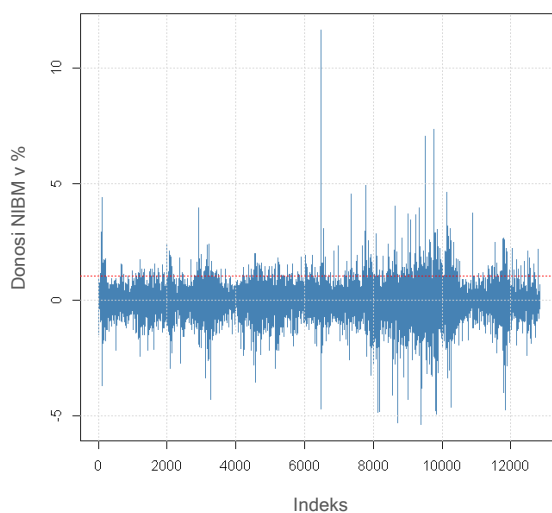


S Slike 21 lahko ocenimo, da podatki prikazujejo naraščajočo premico nekje pri vrednosti 1 %. Torej lahko ponovno sklepamo, da bi bila mejna vrednost 1 % pravilna izbira in da se podatki od tu naprej porazdeljujejo po GPD-porazdelitvi s pozitivnim oblikovnim parametrom ξ .

7.7.4 Računski pristop določanja mejne vrednosti

Pri iskanju mejne vrednosti, ki bi ustrezala GPD-aproksimaciji, lahko uporabimo tudi funkcijo *findThreshold*. Ta izračuna višino mejne vrednosti, da določeno število leži nad to mejno vrednostjo. V našem primeru bomo vzeli, da nad mejno vrednostjo leži 5 % vseh opazovanj, in tako dobili sledeči rezultat: 5 % najvišjih opazovanj predstavlja 641 vrednosti ter določa mejno vrednost $u = 1,03$ %. Na Sliki 22 lahko vidimo grafični prikaz te mejne vrednosti za opazovane podatke negativnih donosov delnice IBM.

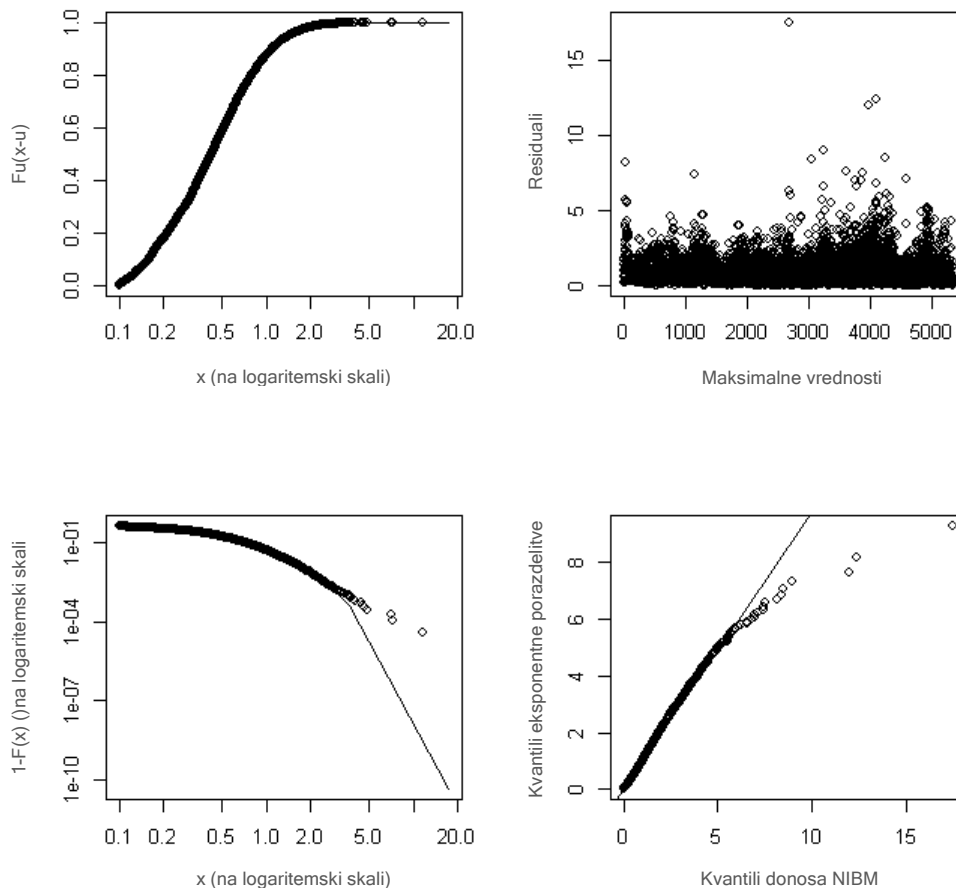
Slika 22: Prikaz izbire mejne vrednosti u , nad katero leži 5 % vseh negativnih donosov



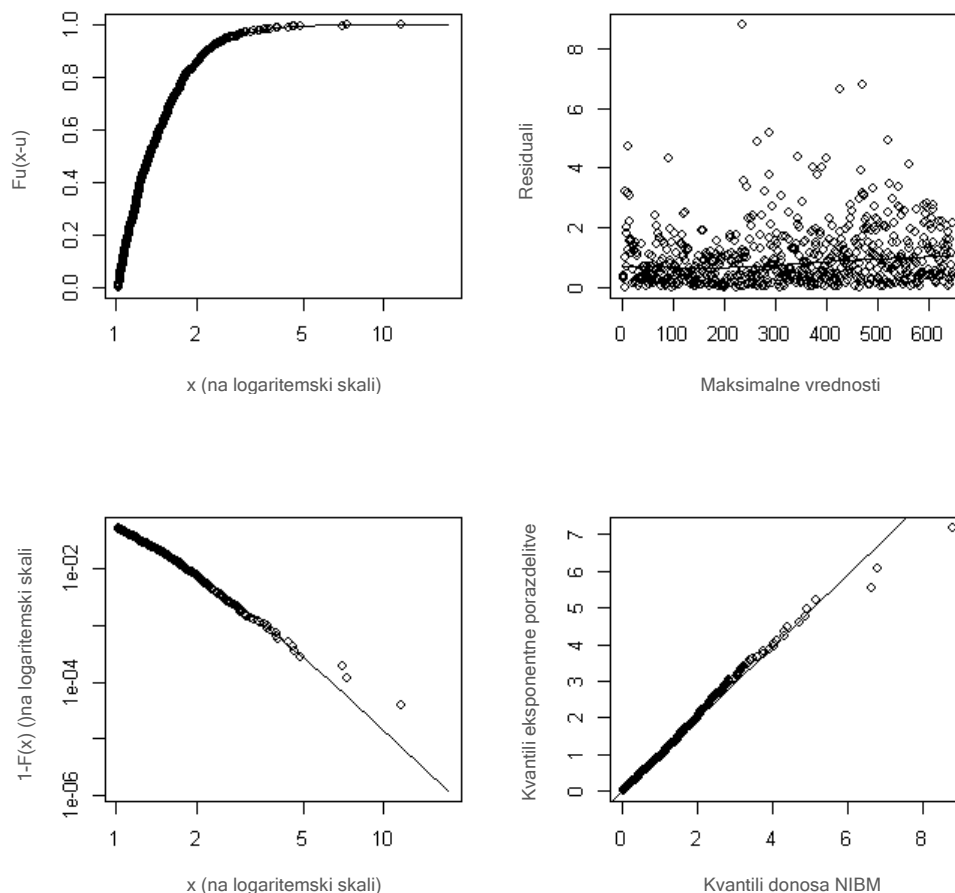
7.7.5 Izbira mejne vrednosti

Vemo, da je za izračun ocen parametrov ξ in σ po metodi največjega verjetja potrebno vnaprej določiti mejno vrednost u . Ta mejna vrednost mora biti dovolj visoka, da bo GPD-aproksimacija ustrezna in hkrati dovolj nizka, da bo na voljo zadostna količina opazovanj za natančno prileganje porazdelitve opazovanim podatkom. Da bi se povsem prepričali o pravilni izbiri mejne vrednosti, si lahko pogledamo, kako se podatki prilegajo GPD-porazdelitvi pri različnih mejnih vrednostih.

Slika 23: Diagnostični grafi za GPD-porazdelitev, ki smo jo prilegali podatkom negativnih donosov IBM-ove delnice pri izbrani mejni vrednosti $u = 0,1$: porazdelitvena funkcija presežkov (levo zgoraj), rep porazdelitvene funkcije (levo spodaj), razsevni diagram residualov (desno zgoraj) in graf kvantilov (desno spodaj)



Slika 24: Diagnostični grafi za GPD-porazdelitev, ki smo jo prilegali podatkom negativnih donosov IBM-ove delnice pri izbrani mejni vrednosti $u = 1,03$: porazdelitvena funkcija presežkov (levo zgoraj), rep porazdelitvene funkcije (levo spodaj), razsevni diagram residualov (desno zgoraj) in graf kvantilov (desno spodaj)



Na Sliki 23 in Sliki 24 so uprizorjeni grafi, ki predstavljajo prileganje GPD-porazdelitve podatkom negativnih donosov IBM-ove delnice. Sliki se razlikujeta v izbrani mejni vrednosti. Za Sliko 23 je izbrana mejna vrednost enaka $u = 0,1$, za Sliko 24 pa $u = 1,03$. Na obeh slikah zgornji levi graf prikazuje porazdelitveno funkcijo presežka čez izbrani mejni vrednosti. Opazimo, da ima na Sliki 23 porazdelitvena funkcija presežka obliko črke S, kar nakazuje na prenizko izbrano mejno vrednost. Na Sliki 24 te oblike ne vidimo več, zato je izbira mejne vrednosti $u = 1,03$ ponovno upravičena. Podobno izrojenost vidimo na levem spodnjem grafu, ki prikazuje porazdelitev repa pri izbrani mejni vrednosti. Za mejno vrednost $u = 0,1$ je opažena izrojenost od predvidenega gibanja, medtem ko se za izbiro $u = 1,03$ opazovani podatki prilegajo premici, ki predstavlja ujemanje teoretične porazdelitve in opazovanih podatkov. Na grafu desno zgoraj so uprizorjeni residuali GPD-porazdelitve pri izbranih mejnih vrednostih, ki ne kažejo avtokorelacije ali heteroskedastičnosti v podatkih, na grafu desno spodaj pa je prikazan graf kvantilov, ki smo ga prikazali in analizirali v Podpoglavju 7.7.2.

Zgornje analize (v Podpoglavjih 7.7.2, 7.7.3 in 7.7.4), za izbor mejne vrednosti, ki bi bil ustrezen za GPD-aproksimacijo, predlagajo vrednost blizu $u = 1\%$. V nadaljevanju magistr-

ske naloge bomo uporabljali $u = 1,03 \%$, ki smo jo izračunali s pomočjo funkcije *findThreshold*, opisane v Podpoglavju 7.7.4.

7.7.6 Ocenjevanje GPD-porazdelitve po metodi največjega verjetja

Sedaj, ko smo določili višino mejne vrednosti, lahko parametre GPD-porazdelitve tudi ocenimo. V Tabeli 8 so uprizorjeni podatki in ocene parametrov ξ in σ GPD-porazdelitve, izračunane po metodi največjega verjetja ter pri izbrani mejni vrednosti $u = 1,03$.

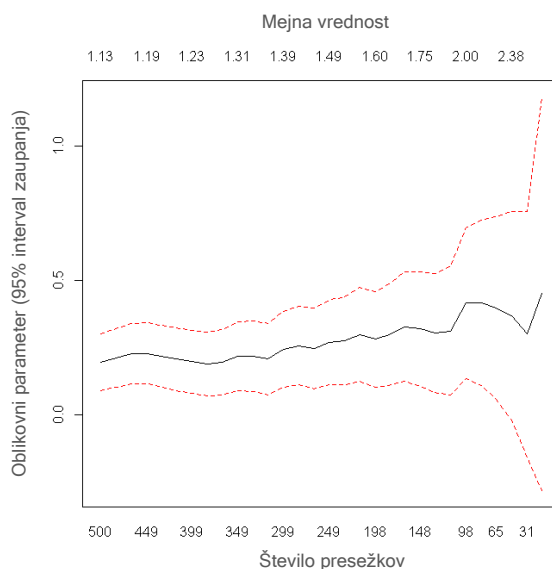
Tabela 8: Ocenjeni parametri GPD-porazdelitve pri mejni vrednosti $u=1,03$

Število vseh opazovanj	12.836	
Mejna vrednost	1,03	
Število opazovanj, ki presegajo mejno vrednost	639	
Ocene parametrov	$\hat{\xi} = 0,216$	$\hat{\sigma} = 0,418$
Asimptotična standardna napaka ocen parametrov	$SE(\hat{\xi}) = 0,046$	$SE(\hat{\sigma}) = 0,025$

Razberemo lahko, da so ocene po metodi največjega verjetja za $\hat{\xi}$ in $\hat{\sigma}$ enake 0,216 in 0,418, njuni standardni napaki pa znašata 0,046 in 0,025. Ker ξ ni blizu 0, ampak zavzema vrednost 0,216, kar je večje od 0, to nakazuje na porazdelitev debelega repa. Prav tako je GPD-ocena za $\hat{\xi}$ manjša kot GEV-ocena $\hat{\xi}$ za podatke, ki temeljijo na letnih, polletnih in četrletnih podatkih, vendar je precej blizu GEV-oceni $\hat{\xi} = 0,239$, ki temelji na mesečnih podatkih.

Na spodnji Sliki 25 si lahko ogledamo, za katere vrednosti bi bila drugačna izbira mejne vrednosti še stabilna. Na y-osi so uprizorjene ocene oblikovnega parametra ξ , izračunane po metodi največjega verjetja, ter njihov 95 % interval zaupanja. Na zgornji x-osi je uprizorjena naraščajoča mejna vrednost u , na spodnji x-osi pa padajoče vrednosti, ki presežejo te mejne vrednosti. Opazimo, da so ocene ξ precej stabilne za mejne vrednosti pod 1,8. Bolj ko se povečuje mejna vrednost, manj opazovanj imamo na voljo in širši je interval zaupanja za ξ .

Slika 25: Ocenjen oblikovni parameter ξ , ko se u in pripadajoče število presežkov spreminjata



7.7.7 Merjenje tveganja na podlagi drugega parametričnega pristopa

Potem ko smo izbrali ustrezno mejno vrednosti in ocenili parametre GPD-porazdelitve, lahko pogledamo kaj je prava dodana vrednost teorije ekstremnih vrednosti. To je bolj natančna ocena ekstremnih kvantilov oziroma mer tveganja. Ti nam bodo povedali, kolikšnemu tveganju smo izpostavljeni oziroma v našem primeru, koliko premoženja lahko izgubimo. Kot smo že omenili v Poglavju 4, sta VaR in ES najbolj pogosti meri tveganja. Zanimala nas bosta oceni VaR in ES na podlagi GPD-aproksimacije pri izbrani mejni vrednostji $u = 1,03$. Pogledali bomo oceno negativnih donosov IBM-ove delnice za kvantil reda 0,95 0,99 in 0,999 GPD-porazdelitve.

Tabela 9: Oceni VaR in ES, ki temeljita na GPD-aproksimaciji

Red kvantila – q	VaR _q	ES _q
0,95	1,028	1,560
0,99	1,831	2,583
0,999	3,590	4,826

Za primer IBM-ove delnice lahko s 95 % verjetnostjo trdimo, da bo dnevni donos presegel vrednost $-1,03$ %, povprečni donos⁷ pa $-1,56$ % (pod pogojem, da je donos nižji od $-1,03$ %). Podobno lahko z 99 % verjetnostjo trdimo, da bo dnevni donos presegel $-1,83$ %, povprečni donos (pod pogojem, da je donos nižji od $-1,83$ %) pa $-2,58$ %. Če pogledamo še bolj po repu porazdelitve, lahko z 99,9 % verjetnostjo trdimo, da bo dnevni donos presegel $-3,59$ %, povprečni donos (pod pogojem, da je donos nižji od $-3,59$ %) pa $-4,83$ %.

V Poglavju 4 smo zapisali, da se VaR osredotoča na rezultate v normalnih tržnih pogojih. Za primerjavo si pogledjmo ocene VaR in ES, ki sta izračunana na podlagi normalne porazdelitve, ki smo jo prilegali našim podatkom. Rezultati so zapisani v Tabeli 10.

Tabela 10: Oceni VaR in ES, ki temeljita na normalni porazdelitvi

Red kvantila – q	VaR _q	ES _q
0,95	1,139062	1,432142
0,99	1,617052	1,854727
0,999	2,152829	2,347012

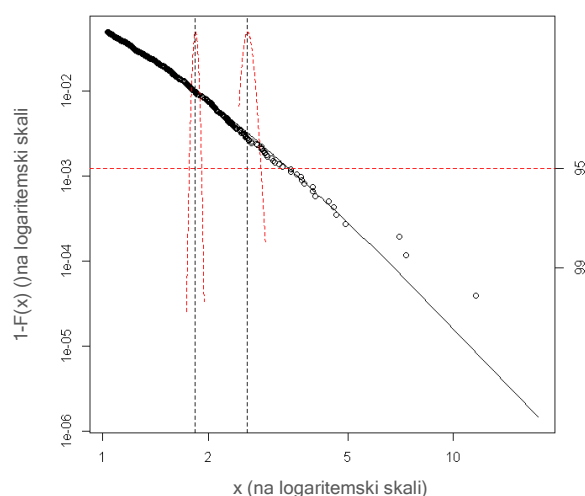
Če primerjamo Tabeli 9 in 10, vidimo, da so ocene VaR in ES podcenjene. Bolj kot se pomikamo proti repu porazdelitve, bolj je opazna podcenjenost. Za lažjo predstavbo si pogledjmo naslednji zgled za oceno VaR. Predpostavimo, da smo kupili za 10 mio EUR vrednih delnic IBM. Z 99 % verjetnostjo lahko trdimo, da bo dnevni donos presegel vrednost $-1,6$ % v primeru predpostavke o normalni porazdelitvi in $-1,8$ % v primeru porazdelitve debelega repa (GPD-porazdelitve). Izračunano na podlagi našega vložka, smo se zmotili v oceni za VaR za 20.000 EUR v primeru, če bi predpostavili normalno porazdelitev. Še večja podcenjenost se zgodi, če pogledamo ES, kjer znaša donos za 99 % kvantil $-1,9$ % v primeru normalne porazdelitve in $-2,6$ % v primeru GPD-porazdelitve. Razlika v oceni je kar za 70.000 EUR. To je 3,5-kratnik napake, ki smo jo storili pri oceni VaR. Zgornji zgled je temeljil na 99 % kvantilu. V primeru da se pomaknemo še bolj proti repu porazdelitve, se te razlike močno povečajo. V Poglavju 4 smo prav tako zapisali, da VaR ni najboljša mera tveganja, saj ne izpolnjuje pogo-

⁷ ES večkrat imenujemo tudi povprečni VaR

ja subaditivnosti, in da namesto nje raje uporabimo ES, ki pogoj izpolnjuje. Na podlagi tega manjšega zglada vidimo, kako visoke so lahko napake oziroma izgube v primeru napačno izbranih porazdelitev in izbire VaR namesto ES.

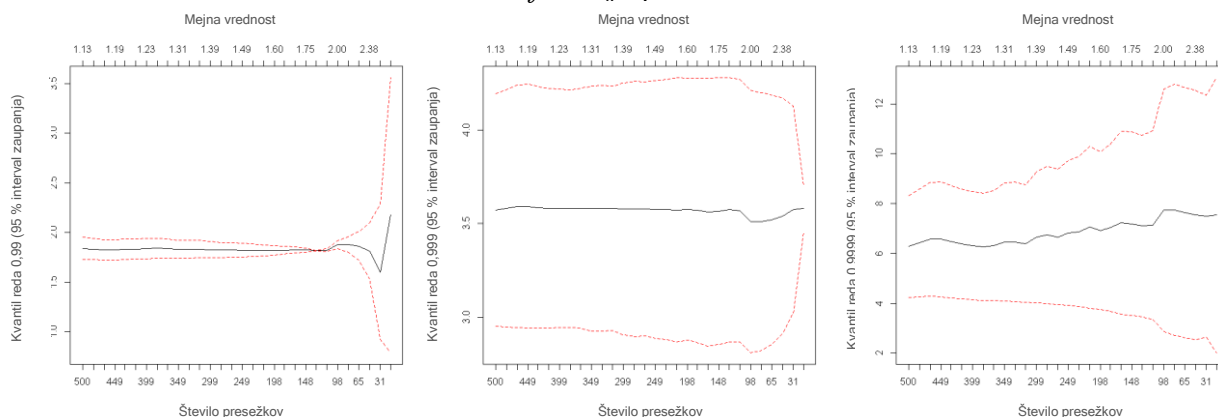
Za lažjo predstavo lahko tudi grafično uprizorimo ocene in asimptotične intervale zaupanja za zgoraj izračunane ocene VaR in ES. Ponovno se bomo osredotočali samo na ocene, ki temeljijo na GPD-aproksimacijah z mejno vrednostjo $u = 1,03$, saj smo že večkrat dokazali, da je ta izbira porazdelitve ustrezna. Na spodnji Sliki 26 sta uprizorjena 95 % intervala zaupanja za oceno VaR_{99} (levi interval zaupanja) in oceno ES_{99} (desni interval zaupanja). Širši kot je interval zaupanja, več močno volatilnih podatkov je bilo opazovanih v repu porazdelitve.

Slika 26: Asimptotični 95 % interval zaupanja za VaR_{99} in ES_{99} , ki temelji na GPD-porazdelitvi z mejno vrednostjo $u = 1,03$, katero smo prilegali negativnim donosom IBM-ove delnice



Preverimo še občutljivost VaR ocen oziroma ocen kvantilov na spremembo mejne vrednosti u . Na Sliki 27 je na levi strani uprizorjen kvantil reda 0,99, ki je stabilen za mejno vrednost pod 1,8 % in nakazuje, da bo enkrat v stotih dnevih negativni donos IBM-ove delnice znašal približno 1,6 %. Seveda, se lahko zgodi tudi, da bo enkrat v stotih dnevih negativni donos presegel 3,5 %, kar lahko vidimo iz zgornjega 95 % intervala zaupanja te slike. Podobno lahko pogledamo, kaj se zgodi pri kvantilu reda 0,999. Uprizorjen je na sredini Slike 27 in je stabilen za vse mejne vrednosti. Rečemo lahko, da se bo enkrat v štirih letih zgodil negativni donos višji od vrednosti 3,6 %. Če se pomaknemo še bolj po repu porazdelitve in pogledamo kvantil reda 0,9999, nam pove, da bo enkrat v štiridesetih letih negativni donos IBM-ove delnice presegel dobrih 6 %, lahko pa preseže tudi donos, višji od -12 %.

Slika 27: Ocene kvantilov reda 0,99 in 0,999 ter 0,9999 kot funkcija mejnih vrednosti ali kot funkcija presežkov



7.8 Neparometrični pristop

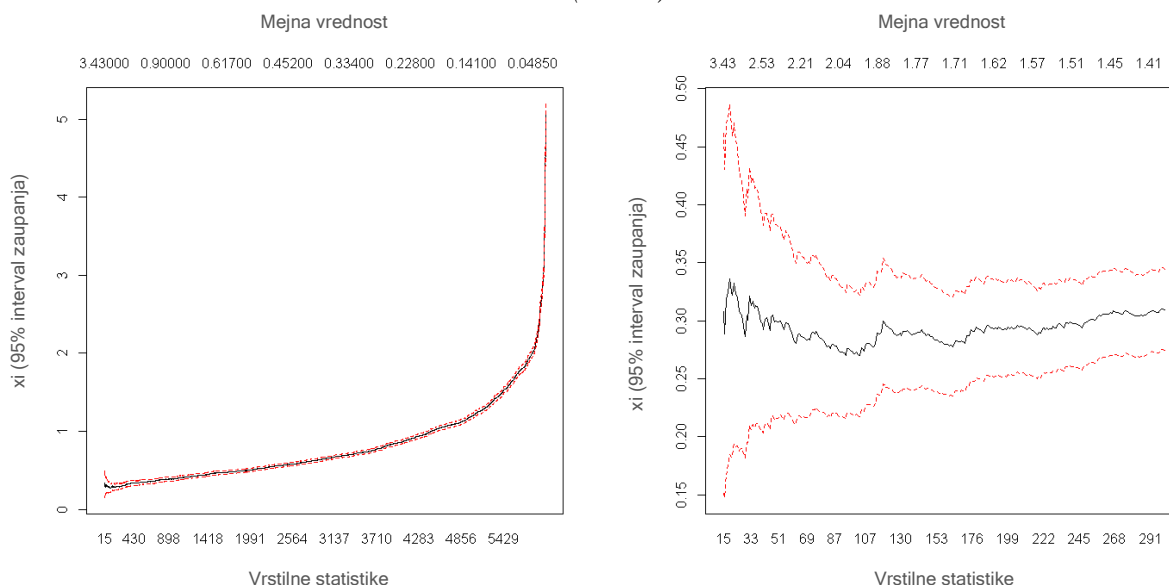
Parameter oblike ξ ali, ekvivalentno, repni indeks $\alpha = \frac{1}{\xi}$ GEV-porazdelitve ali GPD-porazdelitve je mogoče oceniti tudi neparometrično, in sicer na več različnih načinov. Zelo popularno metodo neparometričnega ocenjevanja je razvil Hill (1975). Ta se osredotoča na primer, ko je $\xi > 0$ ($\alpha > 0$), torej se podatki porazdeljujejo po eni izmed porazdelitev debelega repa, ki spada v maksimalno območje privlačnosti Fréchetove porazdelitve (glej Prilogo 2).

7.8.1 Hillova neparometrična ocena repnega indeksa

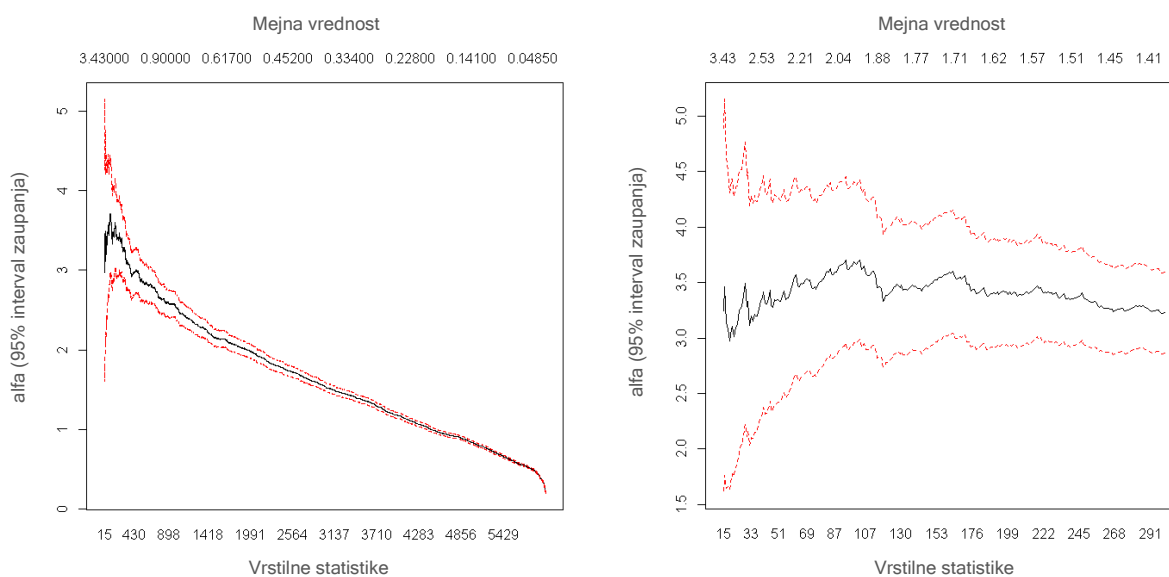
Hillovi cenilki $\hat{\xi}^{Hill}$ in $\hat{\alpha}^{Hill}$ sta odvisni od števila k . Spomnimo se, da k igra enako vlogo v analizi GEV-porazdelitve, kot ga igra N_u v analizi GPD-porazdelitve. Tako k kot N_u predstavljata število opazovanj v repu porazdelitve. Hillovi cenilki, $\hat{\xi}^{Hill}$ in $\hat{\alpha}^{Hill}$, sta zato uprizorjeni kot funkciji mejnih vrednosti. Tako lahko vidimo, za katere mejne vrednosti sta Hillovi cenilki stabilni.

Na Sliki 28 in Sliki 29 sta uprizorjeni Hillovi cenilki $\hat{\xi}^{Hill}$ in $\hat{\alpha}^{Hill}$. Na levem grafu sta cenilki uprizorjeni za vsa opazovanja, na desnem pa za prvih 300 vrstilnih statistik, kjer je cenilka najmanj stabilna. Na y-osi je uprizorjena Hillova cenilka in njen 95 % interval zaupanja, na zgornji x-osi njene pripadajoče mejne vrednosti ter na spodnji x-osi pripadajoče vrstilne statistike. Ker vemo, da velja $\hat{\alpha}^{Hill} = \frac{1}{\hat{\xi}^{Hill}}$, so temu razmerju primerno uprizorjeni tudi spodnji grafi. S Slike 28 lahko razberemo, da se $\hat{\xi}^{Hill}$ stabilizira pri vrednosti 0,30, pripadajoča vrednost $\hat{\alpha}^{Hill}$ pa se stabilizira pri vrednosti 3,3, kar lahko razberemo tako s Slike 29 kot tudi iz razmerja $\hat{\alpha}^{Hill} = \frac{1}{\hat{\xi}^{Hill}}$.

Slika 28: Hillova cenilka $\hat{\xi}^{Hill}$ za vsa opazovanja (levo) in za prvih 300 vrstilnih statistik (desno)



Slika 29: Hillova cenilka $\hat{\alpha}^{Hill}$ za vsa opazovanja (levo) in za prvih 300 vrstilnih statistik (desno)



Podatke z zgornjih slik bomo nazorneje prikazali v Tabeli 11. Opazimo, da je Hillova cenilka za negativne dnevne donose IBM-ove delnice stabilna, razen v primeru, ko je število vrstilnih statistik majhno (t. j. kjer se nahaja največ ekstremnih vrednosti). Za majhno število vrstilnih statistik je volatilitnost visoka in stabilizacija ni tako očitna. Ocenjen parameter oblike $\hat{\xi}^{Hill}$ pa se za višje vrstilne statistike stabilizira okoli vrednosti 0,30, kar pa smo ugotovili tudi že s Slike 28.

Tabela 11: Rezultati Hillove cenilke s standardnimi napakami za donose IBM-a

Število vrstilnih statistik	10	50	130	170	250	430
Hilova ocena za ξ	0,344	0,305	0,293	0,282	0,302	0,336
Standardna napaka	0,109	0,043	0,026	0,022	0,019	0,016

SKLEP

Teorija ekstremnih vrednosti se je do danes uveljavila na raznovrstnih področjih, predvsem v hidrologiji, inženirstvu, aktuarstvu, financah ter upravljanju s tveganji. Skozi čas se je pojavila potreba po razvoju omenjene teorije, saj so se statistiki pogosteje srečevali z ekstremnimi naravnimi katastrofami, močnimi neurji, tsunamiji, obilnimi padavinami, dolgoletnimi poplavami ter tudi močnimi potresi. Ker so vsi ti dogodki ogrožali človeška življenja in za sabo puščali ogromno materialno škodo, se je kmalu začelo razvijati tudi zavarovalništvo, kjer so se ljudje želeli zavarovati pred takšnimi katastrofalnimi posledicami. Napovedovanje ekstremnih dogodkov oziroma katastrof je tako postal osnovni cilj hidrologov, statistikov, matematikov in aktuarjev.

S takšnimi problemi pa se niso soočali samo navedeni, temveč tudi finančni analitiki, katerim so svetovne finančne in gospodarske krize povzročale visoke izgube donosov finančnih instrumentov. Ekstremni dogodki, kot so naravne katastrofe in finančne krize, imajo močan vpliv na naša življenja. Zato je zelo pomembno, da znamo modelirati in analizirati tudi take vrste dogodkov. Tradicionalne parametrične statistične in ekonometrične metode temeljijo na ocenjevanju celotne porazdelitve, kjer se po navadi osredotočamo na osrednja območja porazdelitve. Te metode niso zasnovane za modeliranje ekstremnih kvantilov in dogodke na repu porazdelitvene funkcije. Zato so strokovnjaki začeli razvijati teorijo ekstremnih vrednosti, ki je upoštevala obe navedeni zahtevi.

Kot že omenjeno, so se z ekstremnimi dogodki prvi začeli ukvarjati statistiki. Ti so se v 18. stoletju preživljali tudi s svetovanjem igralcem na srečo in posledično odkrili t. i. normalno porazdelitev. Po skoraj stotih letih so ugotovili, da normalna porazdelitev ni vedno ustrezna in da je potrebno opisati tudi druge vrste porazdelitev. Zanimala jih je predvsem verjetnost najslabšega možnega izida, kako visoko vrednost lahko doseže in kako pogosto se lahko zgodi. Zanimala jih je torej verjetnost in vrednost ekstremnih dogodkov. Ekstremni dogodki se po mnenju Rocca (2011) zgodijo redko in z nizko verjetnostjo, torej ne pripadajo osrednjemu delu normalne porazdelitve, ampak repu porazdelitve. Več ekstremnih dogodkov kot opazujemo, več imamo vrednosti v repu porazdelitve in debelejši je ta rep. Te porazdelitve imenujemo porazdelitve z debelim repom.

Tako se je v začetku 19. stoletja pričel razvoj teorije ekstremnih vrednosti, ki se je osredotočal ravno na rep porazdelitve. Do danes sta se razvila dva temeljna pristopa. Prvi pristop se imenuje metoda maksimumov skupin podatkov. Ta v isto skupino združi dogodke, ki so se zgodili v enako dolgem časovnem obdobju, nato pa poišče najvišjo vrednost v vsakem takem obdobju in te vrednosti analizira. Pri tem pristopu je uporabljen Fisher-Tippetov izrek, ki pravi, da je ena izmed treh porazdelitev ekstremnih vrednosti v limiti prilegana opazovanim vrednostim. Drugi pristop k EVT se imenuje model preseganja mejne vrednosti in analizira le opazovanja, ki presegajo vnaprej izbrano mejno vrednost. Pri tem pristopu pa je uporabljen Izrek Pickands-Balkema-de Haan, ki je za visoke mejne vrednosti povezal limitno obliko funkcije presežka s posplošeno Pareto porazdelitvijo. Oba zgoraj opisana pristopa sta parametrična pristopa k teoriji ekstremnih vrednosti. Tekom magistrske naloge smo spoznali tudi neparametrični pristop, natančneje Hillovo cenilko, ki ocenjuje repni indeks porazdelitve z debelim repom.

Običajne parametrične in neparametrične metode za ocenjevanje porazdelitev in gostot porazdelitev, ki ne temeljijo na EVT, delujejo dobro v empiričnih porazdelitvah, kjer imamo na

voljo veliko število opazovanj, vendar se slabo prilegajo v repih porazdelitev, kjer imamo najvišje ekstremne vrednosti. Reševanje tega problema je postal objektiven cilj analize ekstremnih vrednosti, ki ocenjuje verjetnostno obnašanje nenavadno velikih vrednosti oziroma izgub in razvija orodje, ki se ukvarja ravno z upravljanjem ekstremnih tveganj. Upravljanje ekstremnih tveganj največkrat zahteva ocene kvantilov in repnih verjetnosti, ki presegajo večino naših opazovanih vrednosti, zato mora biti prileganje v repih porazdelitve čim bolj natančno. Metode teorije ekstremnih vrednosti se osredotočajo na modeliranje obnašanja repov porazdelitve in zato analizirajo le ekstremne vrednosti namesto celotnega obsega podatkov.

Kot že omenjeno, imamo v teoriji ekstremnih vrednosti na voljo dva parametrična pristopa. Katerega izmed parametričnih pristopov EVT uporabimo, je odvisno od razpoložljivosti podatkov, želenega časovnega obdobja in izbire vrste tveganja. Za večje obsege podatkov z enako dolgimi časovnimi obdobji je najbolj ustrezen pristop metoda maksimumov skupin podatkov. Ta metoda je enostavna za uporabo in omogoča izračun stopnje poplačila ter napoved dolžine obdobja, kjer bo ta stopnja poplačila dosežena. Ta ocena se uporablja predvsem pri primerih stres testov. Metoda preseganja mejne vrednosti modelira podatke bolj učinkovito, saj uporablja le presežke čez vnaprej izbrano mejno vrednost. V primeru omejenosti s podatki je ta metoda ustrežnejša. S pomočjo metode preseganja mejne vrednosti, s katero je tesno povezana tudi GPD-porazdelitev, lahko natančneje izračunamo meri tveganja VaR in ES.

V prvem delu magistrske naloge smo spoznali teoretično podlago teorije ekstremnih vrednosti, v empiričnem delu magistrske naloge pa smo se osredotočili predvsem na uporabo EVT v finančah, natančneje na donosih delnice podjetja IBM. Tako delniški donosi, kot tudi druge analize EVT v finančah se nanašajo na merjenje in upravljanje finančnih tveganj ali negotovosti v prihodnjih donosih. S pomočjo poznanih preteklih donosov tako lahko ocenimo verjetnost, velikost in dolžino naslednje prihajajoče večje izgube donosa.

Povedano drugače, z EVT lahko odgovorimo na vprašanja, kot so:

- Kolikšna je verjetnost, da se v naslednjih 10 letih zgodi novo krizno obdobje?
- Kolikšna je verjetnost, da bo maksimalna vrednost v naslednjem letu preseгла vse prejšnje maksimalne vrednosti? Z drugimi besedami, kolikšna je verjetnost, da bo v naslednjem letu dosežen nov rekord?
- Kolikšna je pričakovana dolžina obdobja (npr. v letih), preden se bo zgodila določena maksimalna vrednost?
- Kolikšna je verjetnost, da bo maksimalna vrednost v naslednjem letu preseгла mejno vrednost u ?
- Kolikšna je 30-letna stopnja poplačila (angl. *return level*) maksimalnih vrednosti? Torej, kako visoka je maksimalna vrednost, ki je v povprečju presežena enkrat letno v 30 letih?
- Kolikšna je maksimalna izguba, ki jo lahko utrpimo, če se nam zgodi dogodek, ki ustreza kvantilu reda 0,999?

Na večino teh vprašanj smo tekom empiričnega dela magistrske naloge tudi odgovorili.

Po mnenju Embrechts et al. (1997) vemo, da je za finančne podatke značilno, da opisujejo porazdelitev debelega repa. Za EVT pa je značilno in dokazano, da je zelo uporabno orodje za ocenjevanje in napovedovanje obnašanja ekstremov v finančnih produktih (Embrechts et al., 1997). Donosi delnice podjetja IBM so odličen primer porazdelitve debelega repa, kateri je tudi izbran za opazovanje. Opazovano je obdobje od začetka leta 1962 do konca leta 2012.

Med tem časom je bil finančni trg dovzeten večjim finančnim krizam. Z vsako krizo se na finančnih trgih pojavi tudi negotovost pri investitorjih. Tako se je zaradi finančne nestabilnosti, nepredvidljivih gospodarskih razmer in posledičnega nezaupanja investitorjev povečala volatilitnost na finančnih trgih. Visoka volatilitnost pa je seveda povzročila tudi visoka nihanja v donosih delnice IBM, kar je ključno za nastanek porazdelitve z debelim repom. Ker nas zanima najslabši možen izid (angl. *worst case scenario*), torej največji možen padec donosa, smo se osredotočili na negativne donose delnice IBM. Pogledali smo osnovne statistike teh podatkov in z asimetričnostjo večjo od 0 ter sploščenostjo višjo od 3 nakazali, da podatki res opisujejo porazdelitev z debelim repom. Najvišji dnevni padec donosa je znašal 11,6 % in predstavlja Črni ponedeljek oziroma zlom finančnega trga oktobra 1987. S pomočjo grafa kvantilov za celotno opazovanje smo dodatno potrdili obstoj porazdelitve debelega repa. Tako je uporaba EVT na donosih delnice IBM upravičena, kar smo potrdili tudi v nadaljevanju empiričnega dela magistrske naloge.

Najprej smo pogledali prvi parametrični pristop k EVT, kjer smo pogledali maksimalne vrednosti v različnih neprekrivajočih se obdobjih: mesečna obdobja, četrtna, polletna in letna obdobja. V vsakem izmed teh obdobji smo izbrali maksimalne vrednosti, katerim smo nato prilegali GEV-porazdelitev. Pri tem smo uporabili metodo največjega verjetja. Ugotovili smo, da je GEV-porazdelitev ustrezna ne glede na to kako dolgo obdobje izberemo. Ocena repnega indeksa ξ je bila pozitivna za vsa izbrana obdobja, kar nam pove, da se podatki porazdeljujejo po Fréchejevi porazdelitvi, torej porazdelitvi debelega repa, ne glede na dolžino izbranega obdobja. Ocene tega parametra so bile stabilne tako za četrtna, polletna kot tudi za letna obdobja, vrednosti pa so znašale okoli $\xi = 0,37$. Z uporabo ocenjenih parametrov lahko napovemo verjetnost naslednjega ekstremnega dogodka v določenem časovnem obdobju. Daljše kot je napovedano obdobje, višja je verjetnost, da se bo zgodil ekstremni dogodek. Na podlagi ocenjenih parametrov za naše opazovane podatke lahko napovemo, da se bo z 1 % verjetnostjo naslednji ekstremni dogodek zgodil v prihajajočem letu. Prav tako lahko sedaj ocenimo stopnjo poplačila. Ta predstavlja višino negativnega donosa, ki bo v povprečju presežen enkrat v izbranem obdobju. Primer 30-letne stopnje poplačila za podatke delnice IBM predstavlja višino negativnega donosa, ki bo presežena enkrat letno v naslednjih 30 letih. Taka ocenjena stopnja poplačila znaša okoli 7 %, njen zgornji 95 % interval zaupanja pa lahko doseže vrednosti tudi višje od 13 %. Spomnimo se, da je najvišji padec delnice IBM znašal 11,6 %, kar je skoraj toliko kot ocenjen zgornji interval 30-letne stopnje poplačila.

Nato smo si pogledali še drugi parametrični pristop k EVT, kjer smo se osredotočali le na presežke čez določeno mejno vrednost. Najprej smo morali določiti ustrezno višino mejne vrednosti. To smo storili s pomočjo grafa kvantilov, grafa povprečnih odmikov (ME-Plot) in funkcije, ki je določila mejno vrednost tako, da je nad njo ležalo 5 % vseh opazovanj. Vsi trije indikatorji določajo izbiro mejne vrednosti blizu 1 %. Na podlagi te izbrane mejne vrednosti smo prilegali GPD-porazdelitev podatkom, ki presegajo to mejno vrednost. Povedano drugače, zanimalo nas je, ali je za vrednosti v repu porazdelitve GPD-aproksimacija ustrezna. Ugotovili smo, da slednje drži, ocenjen repni indeks ξ pa je znašal 0,216. Ta ocena je zelo blizu oceni GEV-porazdelitve za mesečne podatke, kjer pa je ξ znašal 0,239. Tudi tu potrjujemo, da se opazovani podatki porazdeljujejo po porazdelitvi z debelim repom, saj je $\xi > 0$.

V empiričnem delu magistrske naloge smo pogledali tudi neparametrični pristop k EVT. Pogledali smo Hillovo neparametrično oceno repnega indeksa. Ta se stabilizira nekje okoli 0,3 %, kar je nedaleč od ocen obeh parametričnih pristopov.

Ugotovili smo, da v obeh parametričnih pristopih, kot tudi v neparametričnem pristopu k EVT, potrjujemo obstoj porazdelitve z debelim repom za opazovane podatke IBM-ove delnice. S tem smo tudi potrdili temeljno tezo magistrske naloge, ki je, da je uporaba teorije ekstremnih vrednosti upravičena na donosih delnice izbranega podjetja, podjetja IBM. Zaradi potrjenega obstoja porazdelitve z debelim repom imamo vedno več podatkov v repu porazdelitve in zato nas zanima kolikšna je ocena kvantila v tako odebeljeni porazdelitvi. Višji kot je opazovan kvantil, bolj ekstremnim podatkom ustreza. Primer ocenjevanja visokega kvantila sta meri tveganja Value at Risk (VaR) in pričakovani izpad (ES), njuni izračuni na podlagi EVT pa predstavljajo pravo dodano vrednost v tej magistrski nalogi.

VaR pristop (Jorion, 2001) je do sedaj požel že veliko kritik. Bistvena kritika zadeva aproksimacijo z normalno porazdelitvijo, ki jo uporablja večina parametričnih metod. S to aproksimacijo zanemarimo tveganje visokih kvantilov, še posebej v primeru porazdelitve z debelim repom, ki je značilna za finančne podatke. Nekaj študij je poskušalo rešiti ta problem z uporabo primernejših porazdelitev (na primer s Studentovo t-porazdelitvijo), vendar se vse VaR metode osredotočajo na osrednja opazovanja porazdelitve. Z drugimi besedami, osredotočajo se na donose v normalnih tržnih pogojih. Investitorji in upravljavci tveganj so se kmalu začeli ukvarjati tudi z dogodki, ki se zgodijo v ekstremnih tržnih pogojih. S pomočjo uporabe EVT v VaR smo tako tudi v našem primeru dokazali, da je slednja bolj primerna, kot ocena VaR v normalnih tržnih pogojih. Za razliko od normalnega VaR, se v VaR, ki upošteva EVT, ne predpostavlja, kakšna naj bi bila opazovana porazdelitev. Tako lahko z uporabo EVT ocenimo vrednost visokih kvantilov, kar je zelo uporabno pri napovedovanju ekstremnih situacij, kjer lahko veliko izgubimo. V našem primeru smo ocenili, da pod predpostavko, da kupimo za 10 mio EUR vrednih delnic IBM, lahko s 99 % verjetnostjo (t. j. 99 % kvantil) trdimo, da bo dnevni donos delnice IBM presegel vrednost -1,6 % v normalnih tržnih pogojih in -1,8 % v primeru porazdelitve debelega repa. Torej smo se pod našo predpostavko o normalni porazdelitvi zmotili v VaR oceni za 20.000 EUR. Še večja podcenjenost se zgodi, če pogledamo oceno ES, kjer bi naša napaka pod enako predpostavko znašala kar 70.000 EUR. To je 3,5-kratnik napake, ki smo jo storili že pri oceni VaR. V primeru, da se pomaknemo še bolj proti repu, se te razlike močno povečajo. V Poglavju 4 smo zapisali, da VaR ni najboljše mera tveganja, saj ne izpolnjuje pogoja subaditivnosti in da namesto nje raje uporabimo ES, ki ta pogoj izpolnjuje. Na podlagi zgoraj opisanega zgleda vidimo, kako visoke so lahko napake oziroma izgube v primeru napačno izbranih porazdelitev in izbire VaR namesto ES.

Tako smo preko širšega teoretičnega in praktičnega primera predstavili in upravičili uporabo teorije ekstremnih vrednosti. Podali smo njene zaključke in ocenili mere tveganja, ki predstavljajo pravo dodano vrednost uporabe EVT. Ker se danes soočamo s kriznimi časi in vedno bolj ekstremnimi dogodki na vseh področjih, moramo izpopolniti znanje in izračunati natančnejše ocene ter napovedi možnih ekstremnih dogodkov v prihodnosti. Teorija ekstremnih vrednosti v Sloveniji je precej nova in nepoznana, zato upamo, da bomo s to magistrsko nalogo spodbudili še kakšnega raziskovalca k tej analizi, ki bo z EVT bolje ocenil ter predvidel stanje, ki se v najslabšem primeru lahko zgodi tudi na njegovem področju. Z večimi analizami na različnih, vendar med sabo povezanih področjih, bi tako lahko lažje predvideli prihajajoče krizno stanje ali morda celo obdobje.

LITERATURA IN VIRI

1. Alexander, C. (2001). *Market Models: A Guide to Financial Data Analysis*. Chichester: John Wiley & Sons Ltd.
2. Alexander, C. (2008). *Market Risk Analysis, Vol. 1: Quantitative Methods in Finance*. Chichester: John Wiley & Sons Ltd.
3. Artzner, P., Delbaen, F., Eber, J., & Heath, D. (1999). Coherent Measures of Risk, *Mathematical Finance*, 9(3), 203-228.
4. Balkema, A., & de Haan, L. (1974). Residual life time at great age. *Annals of Probability*, 2(5), 792-804.
5. Basel committee on Banking Supervision (2001). *Progress towards completion of the new Basel Capital Accord*. (2001). Basle: Basel committee on Banking Supervision.
6. Beirlant, J., Teugels, J., & Vynckier, P. (1996). *Practical analysis of extreme values*. Leuven: Leuven University Press.
7. Beirlant, J., Goegebeur, Y., Segers, J., & Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*. Chichester: John Wiley & Sons Ltd.
8. Bensalah, Y. (2000). Steps in Applying Extreme Value Theory to Finance: A Review. *Working paper*, 20. Ottawa: Bank of Canada.
9. Byström, H. N. E. (2005). Extreme value theory and extremely large electricity price changes. *International Review of Economics and Finance*, 14(1), 41-55.
10. Byström, H. N. E. (2006). Using extreme value theory to estimate the likelihood of banking sector failure. *The European Journal of Finance*, 12(4), 303-312.
11. *Cepitev delnic*. Najdeno 26. novembra 2012 na spletni strani <http://www.ljse.si/cgi-bin/jve.cgi?doc=791&crka=R&sid=S4pc6gWRRaElr1Ko>
12. Chan-Lau, J. A., Mitra, S., & Ong, L. L. (2007). Contagion risk in the international banking system and implication for London as a global financial center. *Working Paper*, WP/07/74. Washington: International Monetary Fund.
13. Chavez-Demoulin, V., & Embrechts, P. (2004). Advanced Extremal Models for Operational Risk. *Working paper*. Zürich: ETH-Zentrum.
14. Chavez-Demoulin, V., Embrechts, P., & Nešlehová, J. (2006). Quantitative Models for Operational Risk: Extremes, Dependence and Aggregation, *Journal of Banking and Finance*, 30(10), 2635-2659.
15. Chernobai, A., Rachev, S. T., & Fabozzi, F. J. (2007). *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis*. Hoboken: John Wiley & Sons Ltd.
16. Coles, S. (2001). *An introduction to Statistical Modeling of Extreme Values*. London: Springer-Verlag.
17. Cotter, J. (2006). Extreme value estimation of boom and crash statistics. *The European Journal of Finance*, 12(6-7), 553-566.
18. Cox, D., & Snell, E. (1968). A general definition of residuals (with discussion). *Journal of the Royal Statistical Society*, 2(30), 248-275.
19. Cox, D., & Hinkley, D. V. (1974). *Theoretical Statistics*. London: Chapman and Hall.
20. Danielson, J., & de Vries C. G. (1997). Tail Index and Quantile Estimation with Very High Frequency Data, *Journal of Empirical Finance*, 4, 241-257.
21. Danielson, J., & de Vries C. G. (2000). Value-at-Risk and Extreme Returns. V A. Trognon (ur.), *Annales d'Economie et de Statistique* (str. 239-270). Luxembourg: STATEC.
22. Davison, A. (1984). Modelling excesses over high thresholds with application. V J. De Oliveira (ur.), *Statistical Extremes and Applications*, 461-482.
23. Davidson, A., & Smith, R. (1990). Models for exceedances over high thresholds (with discussion), *Journal of the Royal Statistical Society*, 52, 393-442.

24. De Haan, L., & Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Boston: Springer.
25. Dekkers, A. L. M., & de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, 17, 1795-1832.
26. Diebold, F. X., Schuermann, T., & Strouhair, J. D. (1997). Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management, *Advances in Computational Management Science*, 2, 3-12.
27. *Donosi delnice podjetja IBM*. (b.l.) Na Yahoo Finance. Najdeno 13. oktobra 2012 na spletni strani <http://finance.yahoo.com/q/hp?s=IBM+Historical+Prices>.
28. Dowd, K. (2002). *Measuring market risk*. Chichester: John Wiley & Sons Ltd.
29. Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Berlin: Springer-Verlag.
30. Embrechts, P., McNeil, A., & Straumann, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. V M.A.H. Dempster (ur), *Risk management: value at risk and beyond* (str. 176-223). Cambridge: Cambridge University Press.
31. Falk, M., Hüsler, J., & Reiss, R. (1994). *Laws of Small numbers: extremes and rare events*. Basel: Birkhäuser.
32. Fernandez, V. (2003a). Extreme Value Theory: Value at Risk and Returns Dependence Around the World. *Working paper*. Santiago: Centro de Economiá Aplicada, Universidad de Chile.
33. Fernandez, V. (2003b). Extreme Value Theory and Value at Risk. *Revista de Análisis Económico*, 18(1), 57-85.
34. Fisher, R. A., & Tippett, H. C. (1928). Limiting Forms of the Frequency Distributions of the Largest or Smallest Member of a Sample, *Proceedings of the Cambridge Philosophy Society*, 24, 180-190.
35. Fréchet, M. (1927). Sur la Loi de Probabilité de l'écart Maximum, *Annales de la Société Polonaise de Mathématique (Cracow)*, 6, 93-117.
36. Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*. Malabar: Robert E. Krieger Publishing Company.
37. Garita, G., & Zhou, C. (2009). Can open capital markets help avoid currency crises?. *Working Paper*, 205. Amsterdam: Netherlands Central Bank, Research Department.
38. Gencay, R., & Selcuk, F., (2006). Overnight borrowing, interest rates and extreme value theory. *European Economic Review*, 50, 547-563.
39. Gilli, M., & Kellezi, E. (2006). An Application of Extreme Value Theory for Measuring Financial Risk. *Computational Economics*, 27(2-3), 207-228.
40. Greene, W. (2000). *Econometric Analysis*, (4th ed.). New York: Prentice Hall, Upper Saddle River.
41. Gnedenko, B. V. (1943). Sur la Distribution Limite de Terme Maximum d'un Séries Aléatoire, *The Annals of Mathematics*, 44, 423-453.
42. Gumbel, E. J. (1958). *Statistics of Extremes*. New York: Columbia University Press.
43. Haan, C. T. (2002). *Statistical methods in hydrology*. Iowa: Iowa State University Press.
44. Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution, *Annals of Statistics*, 13, 331-341.
45. Hogg, R., & Klugman, S. (1984). *Loss Distributions*. New York: Wiley.
46. Holton, G. (2003). *Value-at-Risk: Theory and Practice*. San Diego: Academic Press.
47. Hosking, J. R. M., & Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29, 339-349.
48. Jansen, D. W., & de Vries, C. G. (1991). On the frequency of large stock returns: Putting booms and busts into perspective. *The Review of Economics and Statistics*, 73(1), 18-24.

49. Jenkinson, A. F. (1955). The Frequency Distribution of Annual Maximum (or Minimum) Values of Meteorological Elements, *Quarterly Journal of The Royal Meteorological Society*, 81, 158-171.
50. Jorion, P. (2001). *Value at Risk*, (2nd ed.). New York: McGraw Hill.
51. Kellezi, E., & Gilli, M. (2000). *Extreme Value Theory for Tail-Related Risk Measures. Preprint*. Geneva: University of Geneva.
52. Kotz, S., & Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*. London: Imperial College Press.
53. Krehbiel, T., & Adkins, L. C. (2008). Extreme daily changes in U.S. Dollar London inter-bank offer rates. *International Review of Economics and Finance*, 17, 397-411.
54. Liang, B. (1999). On the performance of Hedge Funds. *Journal of Finance and Quantitative Analysis*, 35, 383-408.
55. Longin, F. M. (2000). From Value-at-Risk to Stress Testing: The Extreme Value Approach. *Journal of Banking and Finance*, 24, 1097-1130.
56. Longin, F. M. (2005). The choice of the distribution of asset returns: How extreme value theory can help?. *Journal of Banking and Finance*, 29, 1017-1035.
57. McNeil, A. J. (1998). Calculating Quantile Risk Measures for Financial Returns using Extreme Value Theory. *Working paper*. Zürich: ETH Zentrum, Department Mathematik.
58. McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7, 271-300.
59. McNeil, A. J., & Saladin, T. (2000). Developing scenarios for future extreme losses using the POT model, *Extremes and Integrated Risk Management*. London: RISK Books.
60. Mina, J., & Xiao, J. Y. (2001). *Return to RiskMetrics: The Evolution of a Standard*. New York: RiskMetrics.
61. Novak, S. Y., & Beirlant, J. (2006). The magnitude of a market crash can be predicted. *Journal of Banking and Finance*, 30, 453-462.
62. Pais, A., & Stork, P. A. (2011). Contagion risk in the Australian banking and property sectors. *Journal of Banking and Finance*, 35(3), 681-697.
63. Pickands, J. (1975). Statistical inference Using Extreme Order Statistics. *The Annals of Statistics*, 3(1), 119-131.
64. Reiss, R. D., & Thomas, M. (2001, 2007). *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields* (2nd and 3rd ed.). Basel: Birkhauser.
65. Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. New York: Springer-Verlag.
66. Resnick, S. I. (2007). *Heavy-Tail Phenomena*. New York: Springer-Verlag.
67. Rocco, M. (2011). Extreme Value Theory: A Survey. *Questioni di Economia e Finanza. Working paper*, 99. Rim: Bank of Italy
68. Rockafellar, T. R., & Uryasev, S. (2002). Conditional Value-at-Risk for general loss distributions. *Journal of Banking & Finance*, 26, 1443-1471.
69. *R-project*. Najdeno 10. septembra 2012 na spletnem naslovu <http://www.r-project.org>.
70. Smith, R. L. (1987). Estimating tails of probability distributions. *Annals of Statistics*, 15, 1174-1207.
71. Smith, R. L. (1999). Measuring risk with extreme value theory. *Working paper*. Chapel Hill: University of North Carolina at Chapel Hill, Department of Statistics.
72. Stanton, R. G. (1987). The Work of L. H. C. Tippett. *Ars Textrina* 7, 179-185.
73. Szegö, G. (2008). Measures of risk. *Journal of Banking & Finance*, 26, 1253-1272.
74. Tsay, R. S. (1999). Extreme value analysis of financial data. *Working paper*. Chapel Hill: University of North Carolina at Chapel Hill, Department of Statistics.

75. Turk, G. (2011). Verjetnosti račun in statistika, 1-246. Najdeno 16. januarja 2012 na spletnem naslovu <http://km.fgg.uni-lj.si/PREDMETI/sei/vrs.pdf>.
76. Von Mises, R. (1936). La Distribution de la plus Grande de n Valeurs. *Rev. of Math. Union Interbalcanique, 1*, 141-160.
77. Weibull, W. (1939). A Statistical Theory of the Strength of Materials. *The Royal-Swedish Institute for Engineering Research, 151*(1).
78. Wilson, T. (1998). *Value at Risk*. V C. Alexander (ur.), *Risk Management and Analysis Measuring and Modelling Financial Risk*. Chichester: John Wiley & Sons Ltd.
79. Yamai, Y., & Yoshiba, T. (2002). Comparative Analyses of Expected Shortfall and Value-At-Risk: Their Validity under Marker Stress. *Monetary and Economic Studies, 20*(3), 181-237.
80. Yamai, Y., & Yoshiba, T. (2005). Value-at-Risk versus Expected Shortfall: A practical perspective. *Journal of Banking and Finance, 29*, 997-1015.
81. Zhao, X. (2010). *Extreme Value Modelling with Application in Finance and Neonatal Research*. (Doktorska disertacija (še nepredložena)). Christchurch: University of Canterbury.
82. Zhou, C. (2009). Existence and consistency of maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis, 101*, 971-983.

PRILOGE

KAZALO PRILOG

Priloga 1:	Seznam kratic.....	1
Priloga 2:	Osnovna matematična teorija za lažje razumevanje magistrske naloge....	2
Priloga 3:	Programska oprema za teorijo ekstremnih vrednosti.....	6

Priloga 1: Seznam kratic

EVT (angl. *Extreme Value Theory*) – teorija ekstremnih vrednosti
GEV(angl. *Generalized Extreme Value Distribution*) – posplošena porazdelitev ekstremnih vrednosti
GPD (angl. *Generalized Pareto Distribution*) – posplošena Paretova porazdelitev
VaR (angl. *Value-at-Risk*) – tvegana vrednost
POT-metoda (angl. *Peak Over Threshold Method*) – metoda presežanja mejne vrednosti
MDA (angl. *Maximum Domain of Attraction*) – maksimalno območje privlačnosti
CLP (angl. *Central Limit Problem*) – centralni limitni problem
CLI – centralni limitni izrek
ES (angl. *Expected Shortfall*) – pričakovani izpad
MEF (angl. *Mean Excess Function*) – funkcija povprečnega presežka
CI (angl. *Confidence Interval*) – interval zaupanja
MLE (angl. *Maximum Likelihood Estimation*) – ocenjevanje po metodi največjega verjetja

Priloga 2: Osnovna matematična teorija za lažje razumevanje magistrske naloge

V magistrskem delu uporabljamo pojme s področja verjetnosti, statistike in matematike, mnogi od njih so bralcu najverjetneje že znani. Zaradi lažjega razumevanja bomo nekatere vseeno predstavili v nadaljevanju. Dokaze vseh definicij in izrekov, ki so zapisani v magistrski nalogi, lahko najdemo v Embrechts, Klüppelberg in Mikosch (1997).

• VERJETNOSTNA PORAZDELITEV

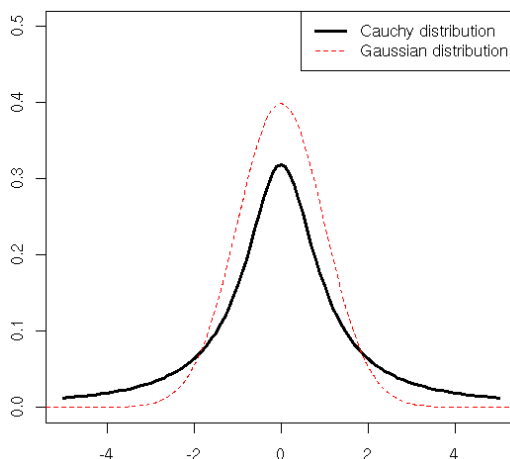
Porazdelitev podatkov matematično prikazujemo s porazdelitveno funkcijo in gostoto verjetnosti.

Porazdelitvena funkcija $F(x)$ slučajne spremenljivke X je funkcija, ki ima pri vsakem realnem x vrednost enako verjetnosti dogodka ($X < x$): $F(x) = P(X \leq x)$.

Funkcija gostote verjetnosti $f(x)$ je funkcija, ki jo dobimo z odvajanjem porazdelitvene funkcije: $f(x) = F'(x) \geq 0$. Gostota verjetnosti označuje verjetnost, da bo neka funkcija zavzela točno določeno vrednost.

Najbolj znana je zvonasta oblika Gaussove krivulje, ki ji drugače rečemo tudi normalna porazdelitev. Na sredini, kjer je vrednost verjetnostne gostote najvišja, ima tudi krivulja svoj vrh, proti repom, ko se verjetnost, da funkcija zavzame vrednosti v tistem območju zmanjšuje, pa se tudi normalna krivulja spušča, postaja vedno bolj ploščata in se vedno bolj bliža abscisi, torej ničelni osi. V nadaljevanju se bomo posvetili predvsem porazdelitvam, ki imajo debel oziroma težek rep. Taka porazdelitev je porazdelitev, ki se na levi ali na desni strani manj približa abscisni osi kot normalna krivulja. Za lažjo predstavbo je na Sliki 1 prikazan primer normalne porazdelitve ter porazdelitve z debelim repom (npr. Cauchyjeve porazdelitve). Matematično pa porazdelitev z debelim repom pomeni, da ima visoko stopnjo sploščenosti in asimetričnosti (Liang, 1999).

Slika 1: Primer normalne porazdelitve in porazdelitve z debelim repom



Za namene teorije ekstremnih vrednosti bomo potrebovali predvsem maksimume podatkov, ki jih v porazdelitvenih funkcijah vidimo v repih porazdelitve. Torej, naj bodo X_1, \dots, X_n neodvisne enako porazdeljene slučajne spremenljivke s porazdelitveno funkcijo F (tj. $F(x) = P(X \leq x)$) in naj bo $M_n = \max(X_1, \dots, X_n)$ n -ti maksimum tega procesa.

Porazdelitvena funkcija maksimuma M_n je:

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = P(X_1 \leq x) * \dots * P(X_n \leq x) = F(x)^n \quad (1)$$

Vendar po navadi porazdelitvene funkcije F ne poznamo, zato pogledamo, kam limitira F^n , ko gre $n \rightarrow \infty$, in katere porazdelitve se nam lahko pojavijo. Vemo, da v primeru, ko je $F(x) < 1$, gre v limiti $F^n \rightarrow 0$, ko gre $n \rightarrow \infty$. Da se izognemo taki možnosti, uporabimo centralni limitni izrek, kjer z normiranjem zagotovimo odsotnost neizrojene limite.

• FUNKCIJA NAJVEČJEGA VERJETJA

Predpostavimo, da imamo vzorec slučajnih spremenljivk X_1, \dots, X_n . Verjetnost, da se zgodi nek dogodek, podan z vektorjem naključnih spremenljivk, je odvisen od izbrane porazdelitvene funkcije in njenih parametrov. Cilj metode največjega verjetja je določiti take ocene parametrov, da bo verjetnost, da se je zgodil ta vzorec, največja (Turk, 2011).

Kot že predpostavljeno, imamo vzorec slučajnih spremenljivk X_1, \dots, X_n , kjer je za vsak X njegova funkcija gostote verjetnosti enaka $f_X(x)$, potem je funkcija verjetja (Turk, 2011):

$$L(\theta) = \prod_{i=1}^n f_X(x_i|\theta), \quad (2)$$

kjer je θ vektor parametrov porazdelitvene funkcije.

Poiskati moramo take vrednosti θ , da ima funkcija $L(\theta)$ maksimum.

Veliko verjetnostnih porazdelitev vsebuje eksponentno funkcijo, zato je najbolje, da funkcijo verjetja najprej logaritmiramo in šele nato odvajamo po vseh možnih parametrih. Logaritemska funkcija je monotona, zato so vrednosti parametrov θ , ki rešijo logaritemsko funkcijo verjetja $\ln L(\theta)$ (angl. *log-likelihood function*), tudi rešitve funkcije verjetja $L(\theta)$ (Haan, 2002).

• CENTRALNI LIMITNI PROBLEM

Veliko izračunov v statistiki in med njimi tudi EVT temelji na verzijah centralnega limitnega problema (CLP). Embrechts et al. (1997) povezujejo uporabo EVT z osnovno logiko, ki leži za centralnim limitnim izrekom. Ta pravi, da vsako zaporedje vsot enako porazdeljenih neodvisnih slučajnih spremenljivk z istim matematičnim upanjem in disperzijo konvergira k normalni porazdelitvi. Tako velja, da tudi sam rep porazdelitve predstavlja neko porazdelitveno funkcijo. Spodaj je zapisana matematična definicija centralnega limitnega izreka, povzeta iz knjige, ki so jo napisali Embrechts et al. Tudi vse nadaljnje definicije in izreki so povzeti iz zgornjega vira.

Definicija 9 (Centralni limitni izrek): Naj bodo X_1, \dots, X_n neodvisne enako porazdeljene slučajne spremenljivke s porazdelitveno funkcijo F , matematičnim upanjem μ in končno varianco σ^2 . Po centralnem limitnem izreku (v nadaljevanju CLI) velja, da

$$\frac{(X_1 + \dots + X_n - n\mu)}{\sigma\sqrt{n}} \Rightarrow Z \quad (3)$$

za slučajno spremenljivko Z , ki ima standardno normalno porazdelitev.

Brez zgoraj definiranih momentov se lahko vprašamo po obstoju konstant $a_n > 0$ in $b_n \in \mathbb{R}$:

$$\frac{(X_1 + \dots + X_n - b_n)}{a_n} \Rightarrow Y \quad (4)$$

za neko neizrojeno slučajno spremenljivko Y (katere porazdelitvena funkcija je H). Pri klasičnem problemu območja privlačnosti (natančneje opisano v nadaljevanju Priloge 2) nas zanimajo vse možne porazdelitve funkcije H , pri čemer poznamo porazdelitveno funkcijo F , tako da enačba

$$\frac{(X_1 + \dots + X_n - b_n)}{a_n} \Rightarrow Y \quad (5)$$

velja za neke določene a_n in b_n . Če ta enakost velja za neka F in H , rečemo, da F pripada območju privlačnosti za H in zapišemo: $F \in MDA(H)$. Denimo, CLI velja za poljuben F s končno varianco, ki tako pripada maksimalnemu območju privlačnosti normalne porazdelitve: $F \in MDA(\Phi)$ (kjer je Φ porazdelitvena funkcija normalne $N(0,1)$ porazdelitve) z $a_n = \sigma\sqrt{n}$ in $b_n = n\mu$.

• MAKSIMALNO OBMOČJE PRIVLAČNOSTI

Zelo pomembno vprašanje zadeva problem območja privlačnosti za obravnavane tri tipe porazdelitve ekstremnih vrednosti Φ_α , Ψ_α in Λ . Za te tri porazdelitve velja, da vse pripadajo maksimalnemu območju privlačnosti. Maksimalno območje privlačnosti definira pogoje, ki zagotavljajo, da bo porazdelitev na repu v maksimalnem območju privlačnosti pripadajoče porazdelitve ekstremnih vrednosti, če in samo če Fisher-Tippetov izrek (1928) drži. To je prvi dokazal Gnedenko (1943).

Definicija 10 (Maksimalno območje privlačnosti)⁸: Slučajna spremenljivka X (oziroma njena porazdelitvena funkcija F) pripada maksimalnemu območju privlačnosti ekstremne porazdelitvene funkcije H , če obstajata konstanti $a_n > 0$ in $b_n \in \mathbb{R}$, tako da

$$\frac{(X_1 + \dots + X_n - b_n)}{a_n} \Rightarrow H \quad (6)$$

drži. Pišemo $X \in MDA(H)$ (oziroma $F \in MDA(H)$).

Spodaj je zapisana ustrezna parametrizacija za opisane tri vrste porazdelitev ekstremnih vrednosti. Definicije do sedaj še neznanih pojmov so razložene v nadaljevanju Priloge 2.

⁸ Opomba:

$$\frac{(X_1 + \dots + X_n - b_n)}{a_n} \Rightarrow H$$

je ekvivalentno

$$\lim_{n \rightarrow \infty} P(M_n \leq a_n x + b_n) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x),$$

kjer je $M_n = \max(X_1, \dots, X_n)$ za $x \in \mathbb{R}$.

- **Maksimalno območje privlačnosti Fréchetove porazdelitve $\Phi_\alpha(x)$**

Porazdelitvena funkcija F pripada maksimalnemu območju privlačnosti Φ_α , $\alpha > 0$, natanko tedaj, ko velja $\bar{F}(x) = x^{-\alpha}L(x)$ za počasi spreminjajočo se funkcijo L .

Če je $F \in MDA(\Phi_\alpha)$, potem je $\frac{M_n}{a_n} \rightarrow \Phi_\alpha$, kjer je normirna konstanta podana z

$$a_n = F^{-1}\left(1 - \frac{1}{n}\right) = \left(\frac{1}{\bar{F}}\right)^\leftarrow(n). \quad (7)$$

- **Maksimalno območje privlačnosti Weibullove porazdelitve $\Psi_\alpha(x)$**

Porazdelitvena funkcija F pripada maksimalnemu območju privlačnosti Ψ_α , $\alpha < 0$, natanko tedaj, ko velja $x_F < \infty$ in $\bar{F}\left(x_F - \frac{1}{x}\right) = x^{-\alpha}L(x)$ za počasi spreminjajočo se funkcijo L .

Če je $F \in MDA(\Psi_\alpha)$, potem je $\frac{M_n - x_F}{a_n} \rightarrow \Psi_\alpha$, kjer sta normirni konstanti enaki

$$a_n = x_F - F^{-1}\left(1 - \frac{1}{n}\right) \quad (8)$$

in

$$b_n = x_F. \quad (9)$$

- **Maksimalno območje privlačnosti Gumbelove porazdelitve $\Lambda(x)$**

Predpostavimo, da je F von Misesova porazdelitvena funkcija. Potem je $F \in MDA(\Lambda)$.

Možna izbira normirnih konstant je

$$b_n = F^{-1}\left(1 - \frac{1}{n}\right) \quad (10)$$

in

$$a_n = a(b_n), \quad (11)$$

kjer je a pomožna funkcija.

• **POSPLOŠENI INVERZ MONOTONE FUNKCIJE**

Definicija 11 (Posplošeni inverz monotone funkcije): Predpostavimo, da je h nepadajoča funkcija, definirana na \mathbb{R} (množica realnih števil). Posplošeni inverz h je definiran kot:

$$h^\leftarrow(t) = \inf\{x \in \mathbb{R}: h(x) \geq t\}. \quad (12)$$

• **KVANTILNA FUNKCIJA**

Definicija 12 (Kvantilna funkcija): Posplošeni inverz porazdelitvene funkcije F :

$$F^{\leftarrow}(t) = \inf\{x \in \mathbb{R}: F(x) \geq t\}, 0 \leq t \leq 1, \quad (13)$$

se imenuje kvantilna funkcija porazdelitvene funkcije F . Količina $x_t = F^{\leftarrow}(t)$ definira t -ti kvantil funkcije F .

- **POČASI SPREMINJAJOČA SE FUNKCIJA**

Definicija 13 (Počasi spreminjajoča se funkcija)⁹: L je počasi se spreminjajoča funkcija (v neskončnost), če je pozitivna, Lebesguovo merljiva in če za njo velja

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1, \quad t > 0. \quad (14)$$

⁹ Opomba: Lebesguova mera na evklidskem prostoru je mera, ki dodeli običajno dolžino, ploščino in prostornino evklidske geometrije ustreznim podmnožicam n -razsežnega evklidskega prostora \mathbb{R}^n . Lebesgueva mera μ_L enotskega intervala $[0,1]$ realnih števil je npr. njegova dolžina v vsakodnevnem smislu besede in je enaka 1, kar zapišemo kot: $\mu_L([0,1]) = 1$.

Priloga 3: Programska oprema za teorijo ekstremnih vrednosti

Empirični del magistrske naloge je bil izveden s pomočjo statističnega programa R. Danes je R množično uporabljeno orodje na področju statistike. Njegova prednost je zagotovo v tem, da je odprtokodni program, torej je dostopen vsem uporabnikom. Zaradi odprtokodnosti se R razvija izjemno hitro, saj lahko vsak uporabnik predlaga nove metode in morebitnim bodočim uporabnikom omogoča tudi neposredno uporabo le-teh.

Program je namenjen statistični analizi podatkov, statističnemu modeliranju ter pripravi podatkov za nadaljnjo obdelavo. Zato je učinkovito orodje za najrazličnejša področja kot so aktu-arstvo, kvantitativne finance, finančna matematika, statistika, bioinformatika, kemija, ekonomija ter povsod drugje, kjer potrebujemo kvantitativne izračune in statistiko ter modeliranje. Ponuja vrsto različnih funkcij, metod in tudi grafično podprtih analiz. Omogočeno je tudi ustvarjanje samostojnih knjižnic, lahko pa se uporabi tudi že obstoječe. Vse te knjižnice so pregledno dostopne v spletni zbirki CRAN, katerih število se nenehno povečuje in do sedaj obsega že skoraj 4.600 različnih knjižnic.

Skladnost programa R z ostalimi programskimi orodji in programi je odlična. Dovoljuje uvoz/izvoz podatkov v poljubnem formatu. Odlikuje pa se tudi v izvozu slik, ki ohranijo visoko kakovost in jih lahko izvozimo v formate pdf, ps, jpg, emf, png ipd.

V empiričnem delu magistrske naloge smo za analizo potrebovali predvsem statistične metode in modeliranje, zato je bila izbira programa R ustrezna. Poleg osnovnih knjižnic smo za analizo teorije ekstremnih vrednosti potrebovali še sledeče knjižnice: `evd`, `evdbayes`, `evir`, `extRemes`, `extremevalues`, `evd`, `fBasics`, `fExtemes`, `ggplot2`, `ismev`, `Lmoments`, `MASS`, `maxLik`, `miscTools`, `POT`, `qualityTools`, `Rcmdr`, `reliaR`, `timeDate`, `timeSeries`, `zoo`.