

# Metode procjene VaR-a

---

Tokić, Vedrana

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Department of Mathematics / Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel za matematiku**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:126:040157>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-11-20**



**mathos**

Repository / Repozitorij:

[Repository of School of Applied Mathematics and Informatics](#)



Sveučilište J. J. Strossmayera u Osijeku  
Odjel za matematiku

**Vedrana Tokić**

# **Metode procjene VaR-a**

Diplomski rad

Osijek, 2021.

Sveučilište J. J. Strossmayera u Osijeku  
Odjel za matematiku

**Vedrana Tokić**

**Metode procjene VaR-a**

**Diplomski rad**

Voditelj: izv. prof. dr. sc. Nenad Šuvak

Osijek, 2021.

# Sadržaj

Uvod	4
<b>1. Osnovni pojmovi</b>	<b>5</b>
1.1. Podaci	6
<b>2. Value at Risk</b>	<b>10</b>
2.1. Definicija VaR-a	10
2.2. Neparametarska procjena	11
2.3. Parametarska procjena	13
2.3.1. EWMA model	13
2.3.2. GARCH model	15
2.4. Semiparametarska procjena	18
2.4.1. Teorija ekstremnih vrijednosti	18
<b>3. Problemi u primjeni VaR-a</b>	<b>29</b>
3.1. VaR kao kvantil	29
3.2. Koherentnost	30
<b>Literatura</b>	<b>32</b>

# Uvod

Financijsko tržište je skup odnosa ponude i potražnje za financijskim proizvodima i uslugama. Uz to, financijsko tržište omogućava protok viška sredstava pojedinaca do onih koji ta sredstva u određenom trenutku trebaju. Od trenutka nastanka takvih tržišta do danas, ljude zanima kakvom riziku se izlažu sudjelovanjem u takvim transakcijama i žele biti u mogućnosti uspoređivati takve rizike. Tu se javljaju problemi jer na kraju dana trgovanja, rizik i dalje ne znamo, dok primjerice cijene i povrate znamo. Razlog tomu leži u činjenici da je rizik latentna varijabla i kao takvog ga je teško kvantificirati.

Ovaj rad bavit će se načinima procjene jedne mjere rizika na financijskom tržištu - VaR-a, odnosno *value at risk*. Problematika procjene sistematizirana je u 3 glavne teme. Prva je, ujedno i najjednostavnija, neparametarska procjena koja ne traži dodatne pretpostavke o distribuciji podataka s kojima računamo VaR. Glavni problem ove metode je što svi podaci nose istu težinu, a iz iskustva znamo kako je tržište dinamično i podložno naglim promjenama pa se stoga ovakva procjena ne pokazuje kao adekvanta u ozbiljnijem upravljanju rizikom. Kako bismo takve promjene bolje opisali, prelazimo na parametarsku procjenu koja koristi modele uvjetne volatilnosti. Posljednja metoda procjene je semiparametarska koja se oslanja na teoriju ekstremnih vrijednosti (EVT) i proučavanje repnog ponašanja distribucije podataka. Svoju primjenu nalazi u granama privrede koju zanimaju nešto ekstremniji događaji kao što su mirovinski fondovi ili osiguranja. Posljednje poglavlje posvećeno je diskutiranju nedostataka VaR-a kao mjere rizika.

# 1. Osnovni pojmovi

Za početak upoznat ćemo se s osnovnim definicijama i oznakama koje će se koristiti u radu. Financijski instrument je ugovor koji svome vlasniku obećava budući tijek dohotka u vidu kamate, dividendi ili drugih dobitaka. Na financijskom tržištu razlikujemo dva osnovna financijska instrumenta: nerizični i rizični. Pod nerizičnim financijskim instrumentom smatrat ćemo novac (u domaćoj valuti) s obzirom da novac posjeduje perfektu likvidnost, dok ćemo pod rizičnim financijskim instrumentom smatrati dionice, obveznice, novac u stranoj valuti itd. Kako njihovu vrijednost ne možemo sasvim točno predviđati, modelirat ćemo ih slučajnim varijablama.

**Definicija 1.1.** *Portfelj u trenutku  $t$  je slučajni vektor  $\phi_t = (\phi_t^0, \dots, \phi_t^d)$  s vrijednostima u  $\mathbb{R}^{d+1}$ , pri čemu  $\phi_t^i$  modelira broj jedinica  $i$ -te rizične financijske imovine,  $i = 1, \dots, d$ , a  $\phi_t^0$  predstavlja broj jedinica nerizične financijske imovine, tj. novca.*

S  $V_t(\phi)$  označavat ćemo vrijednost (cijenu) portfelja  $\phi$  u trenutku  $t$ . Pojam kojim barataju sudionici tržišta, uz cijene financijske imovine koje označavamo sa  $S_t$ , su povrati koje izvodimo iz cijena.

**Definicija 1.2.** *Povrat je relativna promjena vrijednosti financijskog instrumenta u danom vremenskom trenutku, a mjeri se u postocima.*

*Relativni povrat  $i$ -tog financijskog instrumenta je postotna promjena cijene u trenutku  $t$  u odnosu na prethodni trenutak  $t - 1$ ,*

$$R_t^i = \frac{S_t^i - S_{t-1}^i}{S_{t-1}^i}, \quad i \in \{0, \dots, d\}, \quad t \in \{1, \dots, T\}.$$

Relativni povrat cijelog portfelja  $\phi_t$  računamo kao skalarni produkt vektora  $\phi_t$  i  $R_t$ , tj.

$$R_{V_t}(\phi_t) = \sum_{k=0}^d \phi_t^k R_t^k = \sum_{k=0}^d \phi_t^k \frac{S_t^k - S_{t-1}^k}{S_{t-1}^k} = \sum_{k=0}^d \phi_t^k \left( \frac{S_t^k}{S_{t-1}^k} - 1 \right).$$

Osim relativnih povrata, često se koriste i log-povrati.

**Definicija 1.3.** *Prirodni logaritam bruto povrata  $(1 + R_t^i)$  nazivamo log-povrat  $i$ -tog financijskog instrumenta u trenutku  $t$ ,*

$$r_t^i = \ln(1 + R_t^i) = \ln\left(\frac{S_t^i}{S_{t-1}^i}\right) = \ln(S_t^i) - \ln(S_{t-1}^i).$$

Pokazuje se da su kod malih promjena cijena relativni i log-povrati približno jednaki, a log-povrati imaju bolja statistička svojstva od relativnih pa se u modeliranju puno češće koriste. No ono što svakako valja napomenuti je to da su relativni povrati točni, a log-povrati simetrični.

**Primjer 1.1.** Promotrimo jednostavan primjer investicije od 100 kn koja donosi relativni povrat od 20%, a zatim povrat od  $-20\%$ . Prvi povrat donosi dobit od 20 kn i investicija vrijedi 120 kn. Drugi povrat nosi gubitak od 24 kn stoga investicija na kraju vrijedi 96 kn. To nije slučaj s log-povratima. Pretpostavimo da imamo analognu situaciju, u ovom slučaju s log-povratima i pitamo se koliko će investicija vrijediti na kraju takvog razdoblja? Postavljamo jednadžbe:

$$\begin{aligned} 0.2 &= \ln(S_{t+1}) - \ln(100), \\ -0.2 &= \ln(S_{t+2}) - \ln(S_{t+1}). \end{aligned}$$

Nakon što ih zbrojimo, jednostavnim računom dolazimo do rezultata da je  $S_{t+2} = 100$  što bi značilo da na kraju razdoblja investicija opet vrijedi 100 kn iz čega jasno možemo vidjeti svojstvo simetričnosti ovih povrata.

Neka svojstva log-povrata koja često možemo susresti jesu da gotovo ne postoji korelacija, dok su kvadrati log-povrata korelirani; prijelazi visokih razina događaju se u klasterima i distribucija log-povrata je vrlo često lijevo asimetrična, te ima teške repove.

**Definicija 1.4.** Za funkciju distribucije  $F$  slučajne varijable  $X$  kažemo da ima teške repove s repnim indeksom  $\alpha > 0$  ako vrijedi:

$$F(x) = 1 - L(x)x^{-\alpha},$$

pri čemu je  $L$  funkcija sa svojstvom da za svaki  $x > 0$   $\frac{L(tx)}{L(t)} \rightarrow 1$  kad  $t \rightarrow \infty$ . Funkciju  $L$  nazivamo sporo varirajuća funkcija.

Što je parametar  $\alpha$  manji, teži su repovi distribucije, što znači da su realizacije ekstremnih vrijednosti vjerojatnije u odnosu na uobičajene distribucije (npr. normalna distribucija). Primjeri distribucija s težim repovima jesu Pareto distribucija i Studentova t-distribucija.

## 1.1. Podaci

Metodama procjene kojima ćemo se baviti u radu demonstrirat ćemo na primjeru cijena dionica dioničkog društva Saponia. Podaci su preuzeti s internetske stranice Zagrebačke burze <sup>1</sup> i odnose se na cijene u periodu od 22.8.2016. do 20.8.2021. Radi se o uzorku od ukupno 824 podatka, a nas će zanimati cijena dionica u trenutku zatvaranja burze.

Saponia je osječka kemijska industrija, bavi se proizvodnjom deterdženata, sapuna i ostalih sredstava za čistoću i osobnu higijenu. Osnovana je 1894., a 1998. preuzima ju tvrtka Mepas iz Širokog Brijega. Gotovo 53% prodaje ostvaruje na domaćem tržištu, a ostatak na regionalnom, no proizvodi se plasiraju i u ostale europske zemlje, a također i u SAD, Kanadu i Australiju. Neki od njihovih najpoznatijih proizvoda jesu: Faks helizim, Ornel, Likvi, Arf i Kalodont. Ime

---

<sup>1</sup><https://zse.hr/>

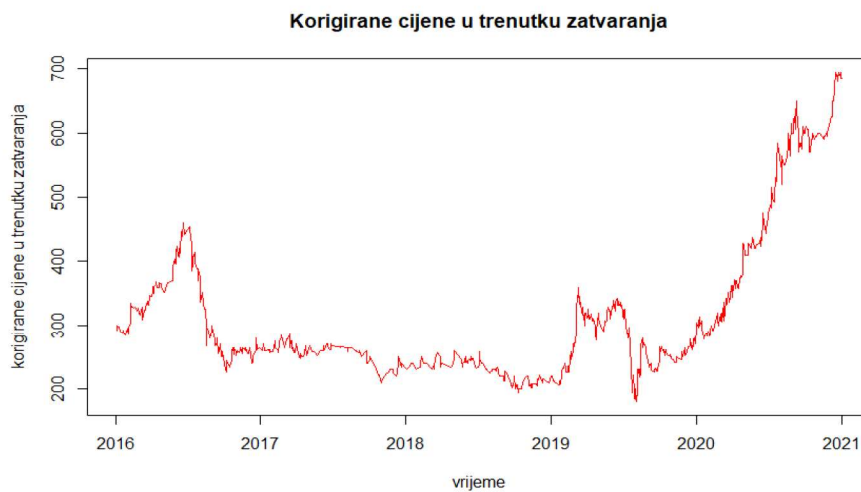
”Saponia” tvornica dobiva 1953. godine i nosi ga i danas. <sup>2</sup>

Proučimo osnovna svojstva cijena dionica ovog društva.

Minimum	Donji kvantil	Medijan	Aritmetička sredina	Gornji kvantil	Maksimum
180.00	242.00	270.00	320.98	344.25	695.00

Tablica 1: Deskriptivna statistika cijene dionica

Tablica 1 sugerira nam da distribucija cijena nije simetrična jer medijan i aritmetička sredina dosta odstupaju. Uz to možemo primijetiti veliku razliku između minimalne i maksimalne cijene dionica.



Slika 1: Grafički prikaz kretanja cijena dionica od 22.8.2016. do 20.8.2021.

Ono što se na slici 1 svakako ističe jest snažni rast u 2020. i 2021. godini. Ovo povežujemo s pojavom novog virusa koji je utjecao na drastičan porast prodaje dezinfekcijskih sredstava što je na Saponiju imalo vrlo pozitivan učinak.

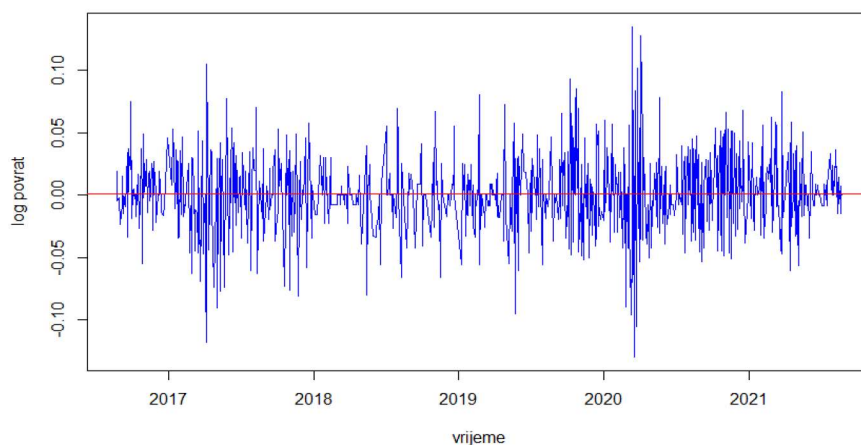
Nakon što cijene pretvorimo u log-povrate, dolazimo do sljedećih podataka:

Minimum	Donji kvantil	Medijan	Aritmetička sredina	Gornji kvantil	Maksimum
-0.129678	-0.016354	0.000000	0.001035	0.018898	0.134819

Tablica 2: Deskriptivna statistika log-povrata od 22.8.2016. do 20.8.2021.

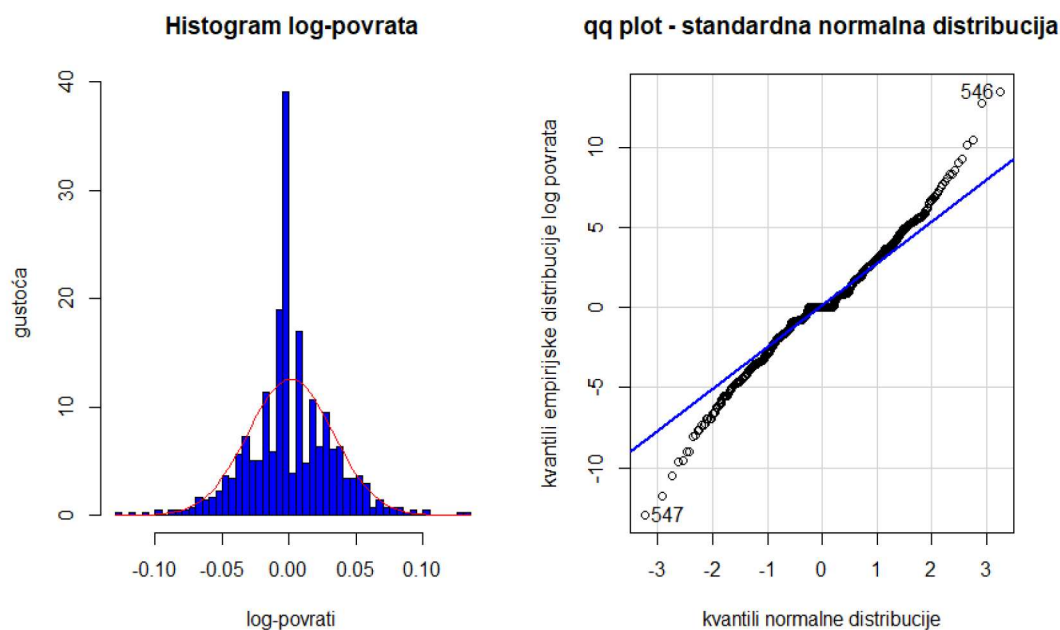
<sup>2</sup>Informacije preuzete s <https://www.saponia.hr/hr/>





Slika 2: Grafički prikaz kretanja log-povrata kroz vrijeme

Na slici 2 i iz tablice 2 možemo primijetiti da nema prevelikog odstupanja log-povrata od prosjeka za period od 22.8.2016. do 20.8.2021. koji je na slici prikazan crvenim pravcem i koji je približno jednak 0. Također primjećujemo klasteriranje o kojem je bilo riječi ranije u radu. Postavlja se pitanje možemo li nešto više reći o njihovoj distribuciji.



Slika 3: Grafički prikazi log-povrata i normalne distribucije

Na slici 3 lijevo je prikazan histogram log-povrata dok je crvenom bojom označena gustoća distribucije normalne slučajne varijable s očekivanjem i varijancom jednakim očekivanju i vari-

janci uzorka. Vidimo da možemo opravdano sumnjati u normalnu distribuiranost log-povrata, maksimum distribucije log-povrata puno je veći od maksimuma normalne distribucije. Desna slika uspoređuje kvantile empirijske distribucije log-povrata i kvantile spomenute normalne distribucije. Otkriva nam da distribucija log-povrata ima teže repove nego normalna distribucija, što znači da su ekstremni događaji vjerojatniji. Ovakva distribucija, s većim maksimumom i težim repovima, naziva se još i leptokurtična distribucija.

## 2. Value at Risk

### 2.1. Definicija VaR-a

Kako su financijska tržišta postajala kompleksnija, pojavila se i potreba za točnijim i efikasnijim mjerenjem rizika. Jedna od mjera rizika koja se nametnula u počecima kvantificiranja rizika jest svakako volatilitet, koju definiramo kao standardnu devijaciju povrata. Glavni problem ove mjere rizika je to što je ona dovoljno informativna samo u slučaju ako distribucija povrata ovisi samo o očekivanju i standardnoj devijaciji, kao što je primjerice normalna distribucija. Odmah možemo i na primjeru podataka predstavljenim u ovom radu primijetiti da takva pretpostavka često nije održiva.

Stoga se kao jedna od glavnih mjera rizika na današnjem tržištu nametnuo *VaR - Value at Risk* ili *vrijednost pod rizikom*. VaR je prvu široku primjenu stekao u bankarskom sektoru i rezultat je rada analitičara. Čelnici velikih američkih banaka tražili su od svojih zaposlenika dnevni izvještaj o izloženosti riziku i to sumiranom u jednoj brojci. Danas ga primjenjuju gotovo sve financijske institucije zahvaljujući Baselskim sporazumima. Glavne prednosti ove mjere rizika nad ostalim su činjenice da je neovisna o tipu distribucije iz koje podaci čija nas rizičnost zanima dolaze te u samo jednom broju daje procjenu ukupnoj izloženosti riziku.

**Definicija 2.1.** *VaR je gubitak generiran na portfelju u vremenskom intervalu  $(t, t + l]$  takav da se s vjerojatnošću  $p \in (0, 1)$  realiziraju gubici veći ili jednaki VaR-u i s vjerojatnošću  $1-p$  realiziraju gubici manji od VaR-a. Oznaka:  $VaR(p)$ .*

**Primjer 2.1.** *Neka je  $p=0.1$ , vremenski interval 1 dan jer baratamo s dnevnim cijenama i  $VaR(0.1)=10000$  kn. To znači da se s vjerojatnošću 0.1 na dnevnoj bazi realizira gubitak veći ili jednak 10000 kn.*

$VaR(p)$  je dakle minimalni gubitak tijekom određenog vremenskog intervala uz određenu vjerojatnost  $p$ . Kako bi došli do formalnije matematičke definicije prvo ćemo s  $\Delta V_l(\phi)$  označiti promjenu vrijednosti portfelja u vremenskom intervalu duljine  $l$ , a funkciju distribucije te slučajne varijable s  $F_l$ . U tom slučaju  $VaR(p)$  definiramo kao:

$$p = P(\Delta V_l(\phi) \leq VaR(p)) = F_l(VaR(p)).$$

Pretpostavljamo da je portfelj u posjedu investitora te da trpi gubitak ukoliko je  $\Delta V_l(\phi) < 0$ . Vjerojatnost da investitor pretrpi gubitak veći (negativniji) od  $VaR(p)$  jednaka je  $p$ . U slučaju da investitor nije u posjedu portfelja čija ga rizičnost zanima, trpi gubitak kada je  $\Delta V_l(\phi) > 0$ . Definiciju VaR-a tada će prilagoditi na način da promatra  $-\Delta V_l(\phi)$ . Mi ćemo se zadržati na prvoj navedenoj interpretaciji.

Ono što se može primijetiti u oba slučaja jest da nas zanimaju repovi funkcije distribucije  $F_l$  (u prvoj interpretaciji lijevi, a u drugoj desni rep), odnosno da je VaR zapravo *kvantil* distribucije  $F_l$  i to  $p$ -tog reda. U praksi distribucija  $F_l$  zapravo nije poznata pa se procjena VaR-a

svodi na procjenu te funkcije distribucije a potom i na procjenu njezinog  $p$ -tog kvantila. Kako je promjene apsolutne vrijednosti dobitaka ili gubitaka nemoguće usporediti dođe li do promjene u primjerice sastavu portfelja, VaR ima smisla računati iz distribucija relativno izraženih varijabli, tj. u ovom slučaju log-povrata.

Postoje 3 osnovna koraka u računanju VaR-a.

1. korak: Odabir vjerojatnosti  $p$ . Sam izbor vjerojatnosti nema nekih ograničenja ili pravila, varira ovisno o korisniku i njegovim afinitetima. Ovisit će o tome što njemu znači velik gubitak, onaj koji se realizira s vjerojatnošću 1%, 5% ili neka druga razina? U praksi se ipak najčešće koristi raspon od 1% do 5%, no i druge razine nalaze svoju primjenu na financijskom tržištu, npr. za mirovinske fondove imalo bi smisla koristiti i vjerojatnosti manje od 1% dok bi za burze imalo smisla koristiti i neke veće vjerojatnosti.

2. korak: Odabir vremenskog razdoblja za koje se računa VaR, kojeg ćemo nazivati *period držanja*. Izbor perioda držanja ovisi o vremenskom razdoblju unutar kojeg se očekuje da će portfelj ili neki njegov dio biti izložen mogućem gubitku. Brokери koji trguju dionicama računat će dnevni VaR jer se njihove pozicije svakodnevno mijenjaju, dok će velika poduzeća računati kvartalni VaR.

3. korak: Identificiranje vjerojatnosne distribucije log-povrata portfelja. Često nismo u mogućnosti točno odrediti distribuciju iz koje log-povrati dolaze stoga se koristimo raznim metodama procjena od kojih će neke biti predstavljene u radu. Standardno ju procjenjujemo koristeći statističke modele i povijesne podatke.

U interpretaciji i usporedbi period držanja i vjerojatnost  $p$  su najvažniji. Za usporedbu dvaju različitih portfelja preko VaR-a moraju se koristiti jednake vjerojatnosti  $p$  i jednako dugi periodi držanja, čak i potpuno isti vremenski intervali.

Postoje 3 glavne metode za procjenu VaR-a: parametarske, neparametarske i semiparametarske.

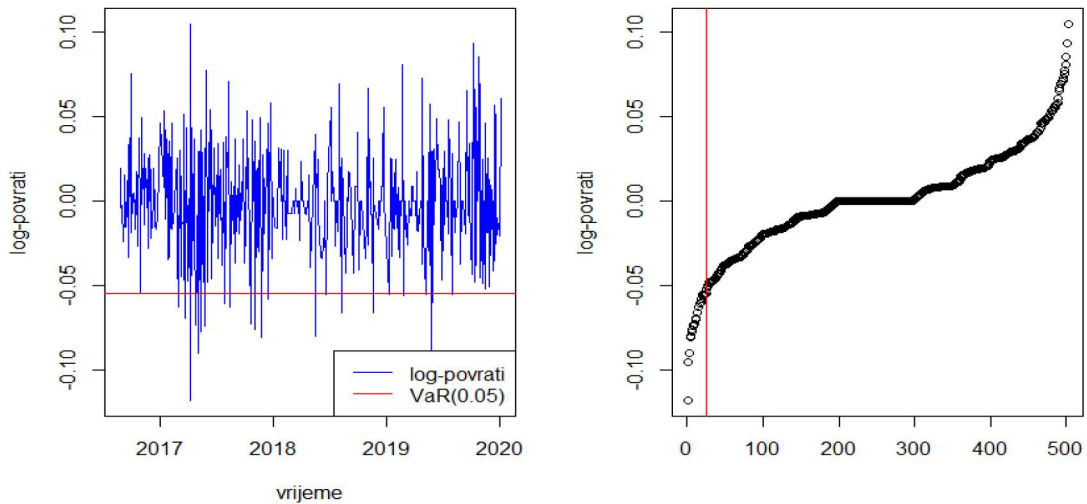
## 2.2. Neparametarska procjena

Ovakva metoda procjene ne pretpostavlja unaprijed poznatu funkciju distribucije slučajne varijable  $\Delta V_t(\phi)$ , tj.  $F_t$ , što poprilično pojednostavljuje proces procjene VaR-a. Najčešća neparametarska procjena je tzv. *povijesna simulacija* koja se zasniva na pretpostavci da se povijest ponavlja, odnosno na očekivanju da će se jedan od prethodno zapaženih povrata pojaviti i kao povrat nekog narednog perioda.

Za početak moramo odabrati vremenski period povrata koji ćemo promatrati. Zatim sortiramo povrate uzlazno i pronađemo  $p$ -ti kvantil, vrijednost od koje je  $p\%$  povrata manje ili jednako, a  $(1 - p)\%$  veće ili jednako. Ako želimo izračunati VaR dan nakon odabranog, pomičemo povrate za jedan dan unaprijed i ponavljamo postupak.

**Primjer 2.2.** *Pretpostavimo da dana 2.1.2021. investitor u svom posjedu ima 500 dionica Saponije. Tog dana tržišna cijena jedne dionice iznosila je 308 kn, što znači da je ukupna*

vrijednost portfelja jednaka 154000 kn. Radimo procjenu dnevnog VaR-a na osnovu 502 podataka i vjerojatnosti 0.05. Slijedi da je  $VaR(0.05)=8456.99$ , odnosno vjerojatnost da investitor u jednom danu izgubi 8456.99 kn ili više jednaka je 0.05.



Slika 4: Grafički prikaz log-povrata i VaR-a

Na slici 4 lijevo možemo vidjeti grafički prikaz log-povrata pri čemu crveni pravac predstavlja  $VaR(0.05)$  dok su na desnoj slici prikazani sortirani povrati i VaR je analogno označen crvenim pravcem.

Iako dosta intuitivna i jednostavna metoda koja na prvi pogled ne zahtjeva nikakve pretpostavke zapravo nosi jednu skrivenu pretpostavku, a to je da se distribucija povrata ne mijenja u vremenskom periodu u kojem računamo VaR, odnosno da su povrati nezavisni i jednako distribuirani (engl. kratica IID). Ta pretpostavka sa sobom nosi nekoliko problema. Prvi je taj da ako za uzorak pretpostavimo da je IID uzorak, svi poduzorci jednake duljine će također biti IID, a već smo vidjeli tendenciju klasteriranja volatilnosti log-povrata što implicira da će procjena VaR-a biti neosjetljiva na promjene u volatilnosti. Također znamo da je za točnost procjene bitno da  $n$ , odnosno veličina uzorka bude što veća pa dolazimo do kontradikcije u definiranju vremenskog perioda na kojem ćemo raditi procjenu: ne smije biti prevelik jer želimo podatke iz jednog klastera volatilnosti, a s druge strane želimo što veći period kako bi statističko zaključivanje bilo što točnije.

Nadalje, svaka od opservacija u ovoj metodi nosi istu težinu. To je problem zbog naglih promjena u volatilnosti, no u odsutnosti toga povijesna simulacija ponaša se bolje od ostalih metoda jer ima manju osjetljivost na stršće vrijednosti.

Pogledajmo što će se dogoditi s VaR-om ako mijenjamo vremenski interval u kojem promatramo log-povrate.

**Primjer 2.3.** *Pretpostavimo sada da dana 20.8.2021. investitor posjeduje 500 dionica Saponije i napravimo procjenu dnevnog VaR-a na osnovu svih 824 podatka o cijenama pri čemu je  $p=0.05$ .  $VaR(0.05)$  tada iznosi 17236.34 kn. Uzmemo li procjenu na osnovu 502 podatka unatrag od dana 20.8.2021., tada iznosi 16312.16 kn.*

Ovaj primjer demonstrira važnost duljine vremenskog perioda na kojem ćemo promatrati VaR, kao i upotrebu što recentnijih podataka. Primjećujemo kako se VaR gotovo udvostručio kada smo promatrali zadnjih 502 podatka nego u Primjeru 2.2. Osim toga, što je dulji vremenski period promatranja, procjeni VaR-a će trebati više da se prilagodi promjeni rizika. Preporuka je da za povijesnu simulaciju veličina uzorka bude minimalno  $\frac{3}{p}$  (pogledati [3]).

## 2.3. Parametarska procjena

Glavna razlika parametarske i neparametarske procjene je u činjenici da se ova metoda bazira na pretpostavljenoj parametarskoj distribuciji  $\Delta V_t(\phi)$  pri čemu procjenjujemo parametre, ali ne i samu distribuciju. Najpoznatiji modeli unutar ovakve procjene jesu EWMA i GARCH modeli.

### 2.3.1. EWMA model

Poznat i pod nazivom RiskMetrics model, procjena VaR-a izvodi se pomoću EWMA modela, eksponencijalnog modela pomičnih prosjeka s težinama, kojim procjenjujemo uvjetnu volatilitet. Nastao je kao poboljšanje modela pomičnih prosjeka u kojem svi povijesni povrati imaju istu težinu. Neka su  $(r_{t-l}, \dots, r_{t-1})$  (povijesni) log-povrati na osnovu kojih radimo procjenu, što znači da radimo s uzorkom duljine  $l$ . Uvjetnu volatilitet u trenutku  $t$  označavamo sa  $\sigma_t$  i procjenjujemo ju na osnovu  $l$  log-povrata i zadane konstante  $\lambda \in (0, 1)$  koju nazivamo *faktor raspadanja*:

$$\hat{\sigma}_t^2 = \frac{1 - \lambda}{\lambda(1 - \lambda^l)} \sum_{k=1}^l \lambda^k r_{t-k}^2.$$

Konstanta  $\lambda$  koristi se kako bi težina log-povrata bila manja što dalje u prošlost idemo, odnosno što je log-povrat "stariji", odakle joj i potječe naziv. EWMA model možemo zapisati i kao sumu procjene prethodne volatiliteta i kvadriranih log-povrata:

$$\hat{\sigma}_t^2 = (1 - \lambda)r_{t-1}^2 + \lambda\hat{\sigma}_{t-1}^2,$$

iz čega se jasno može uočiti da je suma težina jednaka 1.

Naziv RiskMetrics potječe od jedne od vodećih i najvećih američkih banaka, J.P. Morgana. Aktivno se baveći procjenom rizika, 1994. godine pustili su u javnost ovaj sustav procjene rizika koji se vrlo brzo proširio i postao standardom. Oni su predložili da  $\lambda$  za dnevne povrate bude jednak 0.94 što je ostalo u uporabi i danas. To je ujedno i glavni nedostatak ovog modela. Nerealno je očekivati da je  $\lambda$  konstanta koja je jednaka za sve tipove financijskih instrumenata

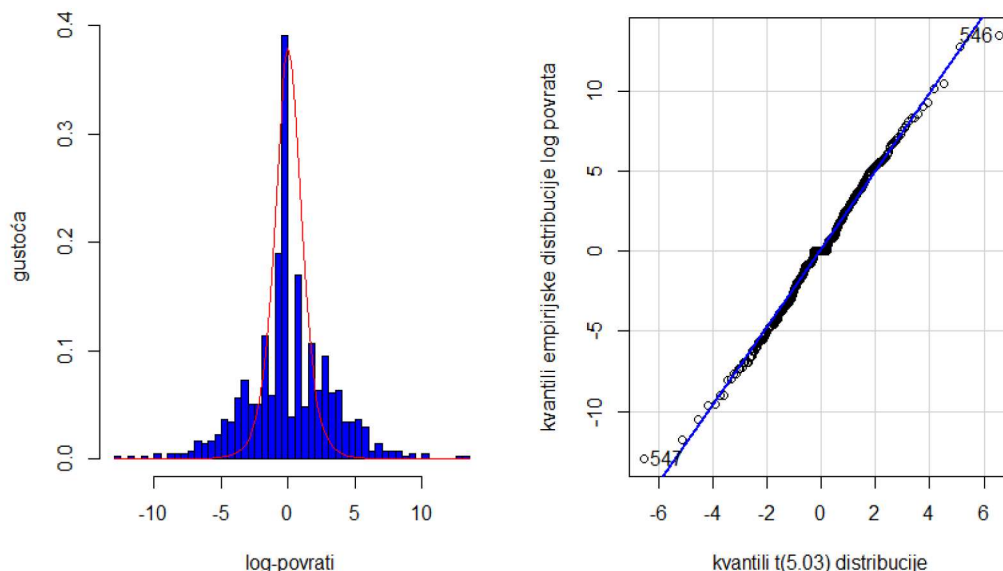
ili da se ne mijenja kroz vrijeme. U tom smislu GARCH modeli o kojima će biti riječi kasnije daju bolju procjenu, ali s obzirom da je EWMA model jednostavniji za uporabu još uvijek se često koristi.

Kako EWMA model daje procjenu uvjetne volatilnosti, VaR u trenutku  $t$  računamo kao:

$$VaR(p) = -\sigma_t \cdot F_r^{-1}(p) \cdot V_{t-1}(\phi), \quad (1)$$

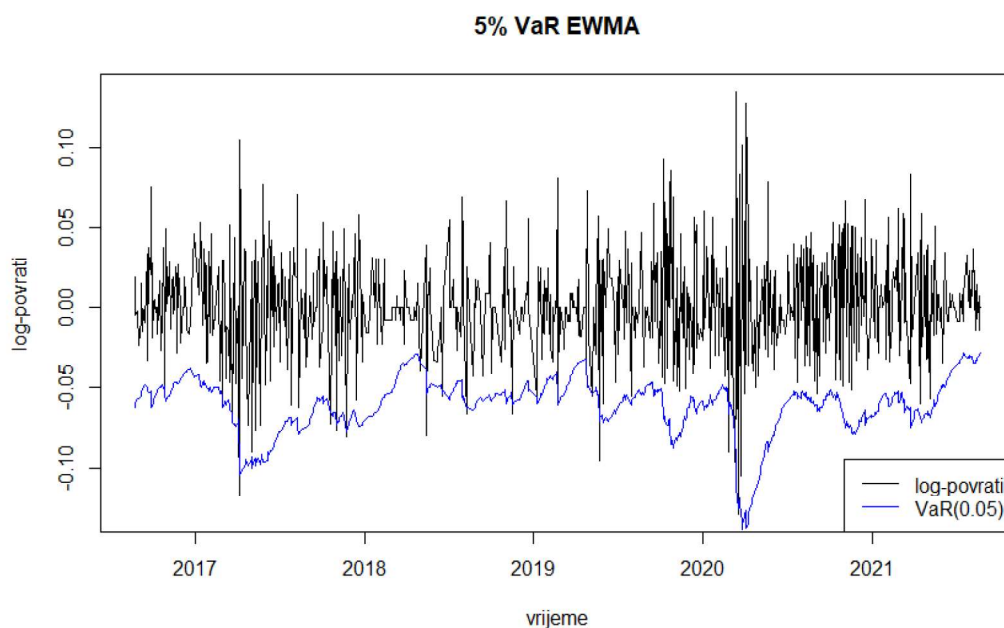
pri čemu je  $F_r$  funkcija distribucije log-povrata.

**Primjer 2.4.** *Pretpostavimo da dana 20.8.2021. investitor posjeduje 500 dionica Saponije. Napraviti ćemo procjenu VaR-a EWMA modelom na osnovu svih dostupnih log-povrata, s tim da ćemo za izračun volatilnosti "prvoga dana", tj.  $\hat{\sigma}_1$  koristiti (bezuovjetnu) varijancu čitavog vremenskog niza. Koristimo  $\lambda = 0.94$  po uzoru na J.P. Morgana. Ono što se nameće kao prvi problem u rješavanju ovog problema je činjenica da ne znamo funkciju distribucije log-povrata pa ćemo ju unutar ovog primjera i procijeniti te ju zatim i koristiti kao stvarnu distribuciju log-povrata. Kako smo vidjeli u prvom poglavlju rada na slici 3, log-povrati ne prate normalnu distribuciju. Tražit ćemo distribuciju log-povrata među distribucijama s težim repovima, točnije pokušat ćemo sa studentovom t-distribucijom. Ono što trebamo procijeniti je stupanj slobode koji bi mogao odgovarati podacima kojima raspolažemo. Procjenjujemo ga uz pomoć R funkcije `fit.st()` unutar paketa `QRM` koja u pozadini koristi metodu maksimalne vjerodostojnosti. Dolazimo do stupnja slobode  $\nu = 5.03$ . Pogledajmo možemo li podatke opisati Studentovom distribucijom.*



Slika 5: Grafički prikazi log-povrata i t-distribucije

Na slici 5 možemo primijetiti kako ova distribucija puno bolje opisuje podatke. Lijevo je prikazan histogram log-povrata dok je crvenom bojom označena teorijska gustoća  $t$ -distribucije s 5.03 stupnjeva slobode. Gotovo da i nema odstupanja kvantila empirijske distribucije log-povrata i  $t(5.03)$  distribucije stoga ćemo ovu distribuciju koristiti u daljnjem modeliranju VaR-a. Sada možemo prijeći na računanje VaR-a EWMA modelom. Na sljedećoj slici možemo vidjeti procjenu 5%-nog VaR-a kroz vrijeme.



Slika 6: VaR(0.05) kroz vrijeme

Primjećujemo kako VaR prilično dobro prati oscilacije log-povrata no ipak treba određeno vrijeme za prilagodbu.  $VaR(0.05)$  na kraju promatranog vremenskog perioda iznosi  $-0.01411536$ , što za investitora koji posjeduje 500 dionica Saponije znači da s vjerojatnošću 0.05 u jednom danu gubi barem 9728.58 kn.

### 2.3.2. GARCH model

S obzirom da je GARCH generalizirani ARCH model, prvo ćemo uvesti ARCH model. Jedan je od prvih modela dizajniranih za opisivanje klasteriranja volatilnosti.

**Definicija 2.2.** Proces  $(X_t, t \in \mathbb{Z})$  je autoregresivni uvjetno heteroskedastični proces reda  $p$  ako vrijedi:

$$X_t = \sigma_t \varepsilon_t,$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2,$$



pri čemu su  $\alpha_j \geq 0$ ,  $(\varepsilon_t, t \in \mathbb{Z})$  je niz nezavisnih jednako distribuiranih slučajnih varijabli s očekivanjem 0 i varijancom 1 i  $\varepsilon_t$  nezavisan je od  $X_{t-j}$ ,  $j \geq 1, \forall t \in \mathbb{Z}$ . Oznaka  $X_t \sim ARCH(p)$ . Niz  $(\varepsilon_t, t \in \mathbb{Z})$  nazivamo niz inovacija.

Ovakav proces ima svojstvo uvjetne heteroskedastičnosti, odnosno svojstvo da se uvjetna varijanca uvjetno na prošlost mijenja u vremenu. Takvi modeli su posebno važni u financijskom modeliranju zbog same strukture log-povrata. Radi jednostavnosti zapisa, pokazat ćemo svojstvo na ARCH(1) procesu.

$$\begin{aligned} Var(X_t|X_{t-1}, X_{t-2}, \dots) &= E[X_t^2|X_{t-1}, X_{t-2}, \dots] - (E[X_t|X_{t-1}, X_{t-2}, \dots])^2, \\ E[X_t|X_{t-1}, X_{t-2}, \dots] &= E[\varepsilon_t \sqrt{\alpha_0 + \alpha_1 X_{t-1}}|X_{t-1}, X_{t-2}, \dots] \\ &= \sqrt{\alpha_0 + \alpha_1 X_{t-1}} E[\varepsilon_t|X_{t-1}, X_{t-2}, \dots] \\ &= 0, \\ Var(X_t|X_{t-1}, X_{t-2}, \dots) &= E[X_t^2|X_{t-1}, X_{t-2}, \dots] = (\alpha_0 + \alpha_1 X_{t-1}^2) E[\varepsilon_t^2] \\ &= \alpha_0 + \alpha_1 X_{t-1}^2. \end{aligned}$$

**Definicija 2.3.** Proces  $(X_t, t \in \mathbb{Z})$  je generalizirani autoregresivni uvjetno heteroskedastični proces reda  $(p, q)$  ako vrijedi:

$$\begin{aligned} X_t &= \sigma_t \varepsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2, \end{aligned}$$

uz oznake analognima onima u ARCH procesu. Pišemo  $X_t \sim GARCH(p, q)$ .

Bavit ćemo se GARCH(1,1) procesom:

$$\begin{aligned} X_t &= \sigma_t \varepsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1} + \beta_1 \sigma_{t-1}^2, \end{aligned}$$

točnije, slično kao i u EWMA modelu, zanimat će nas model kojim ćemo procjenjivati uvjetnu volatilnost:

$$\hat{\sigma}_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1 r_{t-1} + \hat{\beta}_1 \hat{\sigma}_{t-1}^2. \quad (2)$$

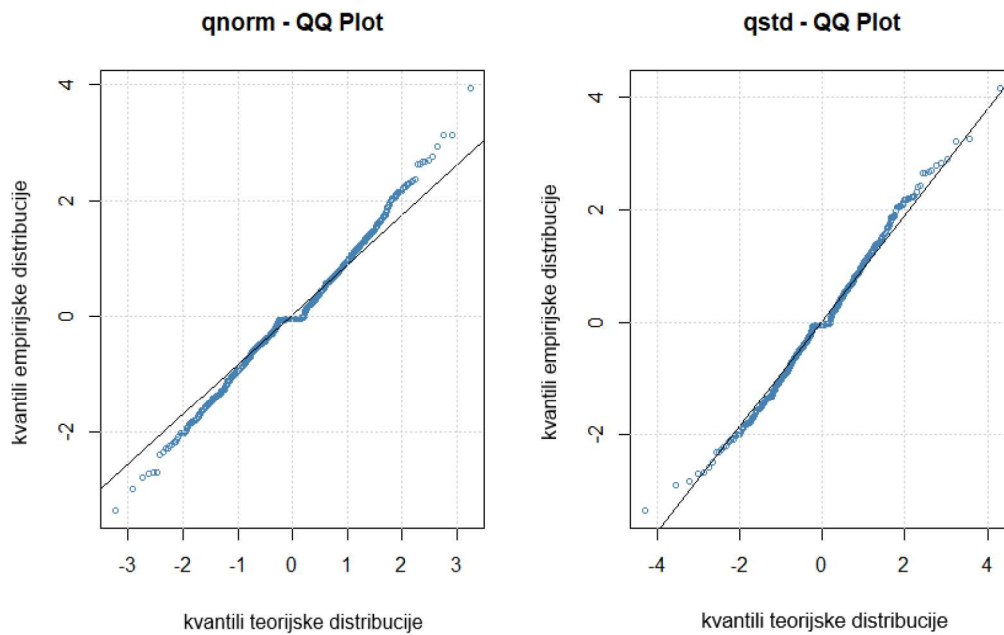
Dovoljni uvjeti za pozitivnost volatilnosti su da konstante  $\alpha_0, \alpha_1$  i  $\beta_1$  budu veće od 0 dok je za slabu stacionarnost potrebno  $\alpha_1 + \beta_1 < 1$ . Također se može pokazati da je bezuvjetna varijanca GARCH(1,1) procesa

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

Dokaze ovih tvrdnji možete pronaći u [11].

Glavna razlika između ovog i EWMA modela je to što smo u EWMA modelu imali zadanu

konstantu  $\lambda$ , dok u GARCH modelu procjenjujemo sve konstante. Prvo smo McLeod-Li statističkim testom <sup>3</sup> provjerili ima li smisla podatke modelirati GARCH procesom. Nulta hipoteza testa je da u modelu ne postoji uvjetna heteroskedastičnost. Na razini značajnosti 0.05 odbacili smo nultu hipotezu,  $p$ -vrijednosti na svim lagovima manje su od 0.05. Kako bi odredili što točniji model, bitno je odrediti i distribuciju reziduala  $\varepsilon_t$ . Najčešće se pretpostavlja da su oni  $\mathcal{N}(0, 1)$ . Nakon analize reziduala u modelu u kojem pretpostavljamo normalnu distribuciju reziduala, dolazimo do zaključka da reziduali imaju teže repove te da ih ima smisla modelirati  $t$ -distribucijom (slika 7).



Slika 7: Kvantili distribucija  $\varepsilon_t$

Nakon što smo izračunali procjene za  $\alpha$ ,  $\alpha_1$  i  $\beta$ , dolazimo do modela:

$$\hat{\sigma}_t^2 = 3.88 \cdot 10^{-5} + 0.2307739r_{t-1} + 0.7606646\hat{\sigma}_{t-1}^2.$$

Procjenu za  $\sigma_t$  računamo po prethodno navedenom izrazu, a VaR po (1).

**Primjer 2.5.** Procjenjujemo s vjerojatnošću 0.05 da dana 21.8.2021. log-povrati u jednom danu padnu ispod -0.03158, odnosno da investitor koji posjeduje 500 dionica Saponije pretrpi gubitak veći ili jednak 10819.33 kn.

<sup>3</sup>Više o testu možete saznati iz Y. Chen, *On the Robustness of Ljung-Box and McLeod-Li Q Tests: A Simulation Study*, Economics Bulletin 3 (2002), 17; 1-10

## 2.4. Semiparametarska procjena

Do sada smo se bavili vjerojatnosti  $p$  koja je jednaka 0.05, no postavlja se pitanje što ako želimo drastično veće ili manje vjerojatnosti, primjerice 0.01%? Ranije smo spomenuli da bi ovakve vjerojatnosti mogle biti od interesa mirovinskim fondovima ili brokerima na burzi. Događaji koji se realiziraju s 0.01% vjerojatnosti su ekstremni, stoga se takva procjena rizika prigodno zasniva na *teoriji ekstremnih vrijednosti (EVT)*. S obzirom da će nas zanimati ekstremi, proučavat ćemo repove distribucija, a EVT će biti alat uz pomoć kojeg ćemo ih procjenjivati.

### 2.4.1. Teorija ekstremnih vrijednosti

Prednost teorije ekstremnih vrijednosti nad nekim drugim metodama je u tome što ne zahtjeva pretpostavku o distribuciji log-povrata, što kod EWMA i GARCH modela nije slučaj. Kod takvih modela, s obzirom da modeliramo cijelu distribuciju log-povrata, opažanja koja se nalaze u centru distribucije dominirat će procjenom pošto ih u uzorku ima više. U skladu s tim, dobivat ćemo točnije procjene za učestale, "obične" događaje, a lošije za repno ponašanje distribucije.

EVT se eksplicitno fokusira na analizu repova distribucije, dakle na vjerojatnost ekstremnih događaja, što znači da pruža točne predikcije ponašanja ekstremnih događaja. U tome se ujedno krije i nedostatak ove teorije - u većini slučajeva nemamo dovoljno ekstremnih zapažanja za procjenu što odmah u startu ograničava primjenjivost i praktičnost ove teorije.

Ova se teorija gotovo uvijek prezentira u terminima desnog repa, odnosno pozitivnih opservacija, što će se koristiti i u ovom poglavlju. Iako nas u procjeni rizika zapravo zanima lijevi rep distribucije, tj. negativne vrijednosti, što smo mogli vidjeti i prije, ovakva prezentacija neće predstavljati problem s obzirom da uvijek možemo log-povrate pomnožiti s -1.

Neka je  $X_1, X_2, \dots$  niz nezavisnih jednako distribuiranih slučajnih varijabli s funkcijom distribucije  $F$  i neka je  $M_n = \max(X_1, \dots, X_n)$ . Tada vrijedi:

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \prod_{k=1}^n P(X_k \leq x) = \prod_{k=1}^n F(x) = F^n(x).$$

Desni rep distribucije karakterizirat ćemo brojem:

$$x_D = \sup\{x \in \mathbb{R} : F(x) < 1\}.$$

Pustimo li  $n$  u beskonačnost, za svaki  $x < x_D$  vrijedit će  $P(M_n \leq x) = F^n(x) \rightarrow 0$ , dok će za  $x \geq x_D$  vrijediti  $F^n(x) = 1$ , što znači da distribucija  $F^n$  postaje degenerativna. Kako bismo to izbjegli, mora se provesti normalizacija. Do problema dolazi kada treba odrediti konstante  $a > 0$  i  $b$  takve da  $\frac{M_n - b}{a}$  ima nedegenerativnu marginalnu distribuciju kada  $n$  ide u beskonačnost, tj.

$$\lim_{n \rightarrow \infty} F^n(ax + b) = G(x).$$

Može se pokazati da ako takve konstante postoje,  $G$  dolazi iz familije *distribucija ekstremnih vrijednosti*.

**Definicija 2.4.** *Funkcija distribucije generalizirane distribucije ekstremnih vrijednosti (GEV) dana je izrazom:*

$$H_{\xi}(x) = \begin{cases} \exp(-(1 + \xi x))^{-1/\xi} & \text{za } \xi \neq 0, \\ \exp(-e^{-x}) & \text{za } \xi = 0, \end{cases}$$

pri čemu je  $x$  standardiziran, tj.  $x = \frac{z-\mu}{\sigma}$ ,  $z \in \mathbb{R}$ .

Glavni rezultat ove teorije ističe da, bez obzira kakav je oblik distribucije s kojom se susrećemo i dokle god se distribucija povrata ne mijenja kroz vrijeme, granična distribucija normaliziranih maksimuma pripada jednoj od sljedećih triju kategorija:

- Weibull - slučaj kada distribucija ima kratak rep (primjerice u osiguranjima, distribucija smrtnosti;  $\xi < 0$ ),
- Gumbel - slučaj u kojima repovi eksponencijalno padaju (normalna i log-normalna distribucija;  $\xi = 0$ ),
- Frèchet - teški repovi (t-distribucija i Pareto;  $\xi > 0$ )

Najznačajniji parametar u ovoj teoriji je *repni indeks* kojeg označavamo s  $\alpha$ , dok njegov inverz, tzv. *parametar oblika*, označavamo s  $\xi$ . Parametar  $\alpha$  Studentove distribucije odgovara broju stupnjeva slobode.

Unutar ove teorije postoje dvije glavne metode za procjenu tih parametara: metoda maksimuma po segmentima i nešto modernija metoda vrijednosti iznad praga (POT).

### Metoda maksimuma po segmentima

Ova metoda namijenjena je velikom uzorku. Temelji se na principu procjene generalizirane distribucije ekstremnih vrijednosti na način da uzorak podijelimo u segmente i zatim gledamo maksimum svakog segmenta te procjenjujemo distribuciju tako formiranog uzorka maksimuma segmenata.

Neka su  $X_1, \dots, X_n$  slučajne varijable koje dijelimo u  $k$  segmenata (blokova) duljine  $l$ . Maksimume tih segmenata  $M_{l_1}, \dots, M_{l_k}$  možemo promatrati kao nezavisan jednako distribuiran uzorak iz GEV. Za procjenu parametara koristimo metodu maksimalne vjerodostojnosti. Kvaliteta procjene ovom metodom uvelike ovisi o izboru duljine segmenata  $l$ . Odabere li se premali  $l$ , doći će do pristranosti u procjeni, a s druge strane ako se odabere prevelik  $l$ , varijanca procjene bit će velika s obzirom da se u tom slučaju procjena radi s malim brojem maksimuma. Duljina segmenta se obično bira tako da odgovara podacima s kojim se radi, nema egzaktne formule za odabir. Primjerice, ako radimo s podacima o temperaturama, prirodno je za  $l$  uzeti jednu

godinu i gledati maksimalne temperature po godinama.

VaR unutar ove metode računamo po formuli:

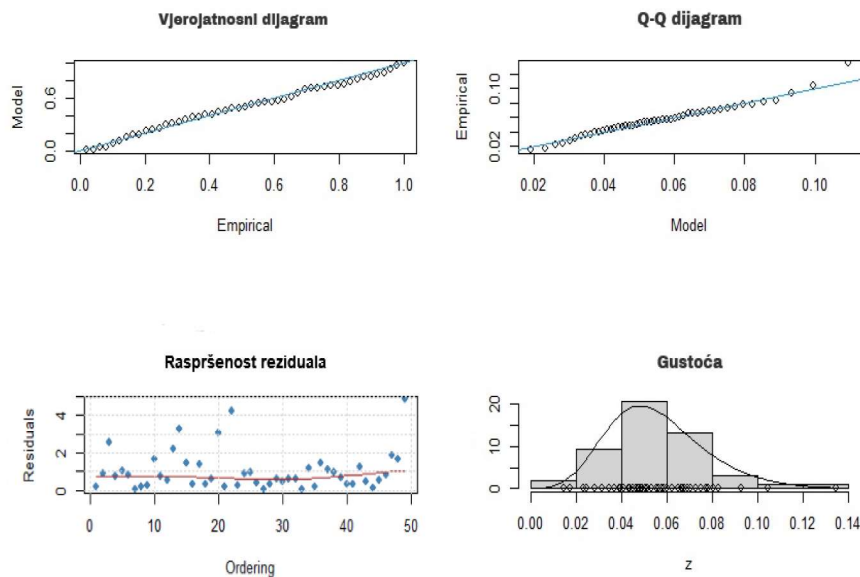
$$VaR(p) = \begin{cases} \mu - \frac{\sigma}{\epsilon} (1 - (-l \ln(1 - p))^{-\epsilon}) & \text{za } \xi \neq 0, \\ \mu - \sigma \ln(-l \ln(1 - p))^{-\epsilon} & \text{za } \xi = 0, \end{cases}$$

pri čemu je  $l$  duljina segmenata te  $\mu$  i  $\sigma$  normalizirajuće konstante iz Definicije 2.4.

**Primjer 2.6.** Unutar ovog primjera primijenit ćemo metodu maksimuma po segmentima na log-povrate dionica Saponije. Za početak treba procijeniti parametre pripadne generalizirane distribucije ekstremnih vrijednosti. To činimo uz R paket *fExtremes*. Nakon testiranja nekoliko duljina segmenata, pokazalo se da je najoptimalnija duljina  $l = 17$ .

parametar	procjena	standardna greška
$\xi$	-0.08548057	0.077158334
$\sigma$	0.01897106	0.002016592
$\mu$	0.04651000	0.0029710782

Tablica 3: Procjene parametara i standardne greške procjene



Slika 8: Grafički prikazi dijagnostike procijenjenog modela

$VaR(0.999)$  iznosi 0.158118, tj. s 99.9%-tnom vjerojatnošću procjenjujemo da log-povrati neće biti veći od 0.158118.

## Metoda vrijednosti iznad praga

Kako za metodu maksimuma po segmentima treba velik broj podataka, a na kraju se koristi razmjerno mali broj podataka za procjenu (metoda je "rasipna" s podacima), u praksi se češće koristi metoda vrijednosti iznad praga s obzirom da je efikasnija i praktičnija. Bazira se na modelima za sve velike opservacije koje prelaze određeni prag i tako bolje koristi podatke o ekstremnim vrijednostima. Unutar ove metode razlikujemo dva pristupa: potpuno parametarski modeli koji se baziraju na generaliziranoj Pareto distribuciji i semiparametarski modeli koje gradimo uz pomoć Hillovog i Pickandovog procjenitelja. Kako generalizirana Pareto distribucija ima važnu ulogu u modeliranju rizika, u okviru ovog poglavlja ćemo predstaviti i takve modele.

Neka je  $u$  prag i  $(r_t, t \in \mathbb{N})$  vremenski niz log-povrata. Ovaj model bazira se na promatranju  $r_t - u$ , ukoliko je vrijednost praga prijeđena, tj.  $r_t > u$ . Zanima nas dakle uvjetna distribucija:

$$F_u(x) := P(r - u \leq x | r > u) = \frac{P(u \leq r \leq x + u)}{P(r > u)} = \frac{P(r \leq x + u) - P(r \leq u)}{1 - P(r \leq u)}. \quad (3)$$

## Generalizirana Pareto distribucija

GPD je glavna distribucija koja se koristi za modeliranje podataka koji prelaze određeni prag, tj. funkcije distribucije prekoračenja. Njezina funkcija distribucije zadana je izrazom:

$$G_{\xi, \beta}(x) = \begin{cases} 1 - (1 + \xi x / \beta)^{-1/\xi} & \text{za } \xi \neq 0, \\ 1 - e^{-x/\beta} & \text{za } \xi = 0, \end{cases}$$

pri čemu je  $\beta > 0$  te  $x \geq 0$  kada  $\xi \geq 0$  i  $0 \leq x \leq -\beta/\xi$  kada  $\xi < 0$ .

Kako je već ranije spomenuto,  $\xi$  je parametar oblika, a  $\beta$  je dodatni parametar koji nazivamo *parametar skaliranja*. Za različite vrijednosti parametra  $\xi$  dobivamo različite poznate distribucije, primjerice za  $\xi$  jednak nuli dobivamo eksponencijalnu distribuciju (koja pripada Gumbelovoj klasi), i u tom smislu je ovo generalizirana distribucija.

Može se pokazati da za sve poznate uobičajene neprekidne funkcije distribucije postoji funkcija  $\beta(u)$  takva da

$$\lim_{u \rightarrow x_D} \sup_{0 \leq x < x_D - u} |F_u(x) - G_{\xi, \beta(u)}| = 0,$$

što znači da kako povećavamo prag  $u$ , odnosno što je prag bliži desnom repu, funkcija distribucije prekoračenja  $F_u$  konvergira generaliziranoj Pareto distribuciji (Balkema i de Haan (1974), Pickands (1975)).

Dakle, da bi primjenjivali ovu metodu, moramo procijeniti oba parametra,  $\xi$  i  $\beta$ . Parametre procjenjujemo koristeći metodu maksimalne vjerodostojnosti. Za  $\xi \neq 0$  i  $(1 + \beta^{-1}\xi x_k) > 0, k = 1, \dots, n$  logaritam funkcije vjerodostojnosti dan je izrazom:

$$l(\xi, \beta) = -n \ln(\beta) - \left(1 + \frac{1}{\xi}\right) \sum_{k=1}^n \ln \left(1 + \xi \frac{x_k}{\beta}\right),$$

a inače je

$$l(\xi, \beta) = \infty.$$

VaR u ovom slučaju računamo prema sljedećoj formuli:

$$VaR(p) = u + \frac{\beta}{\xi} \left[ \left( \frac{1-p}{F(u)} \right)^{-\xi} - 1 \right].$$

Postavlja se pitanje kako odrediti prag  $u$ . U tu svrhu definiramo ME funkciju (Mean excess function) na sljedeći način:

$$e(u) = E[X - u | X > u].$$

U slučaju kada uzorak dolazi iz generalizirane Pareto distribucije, ME funkcija je linearna u varijabli  $u$ :

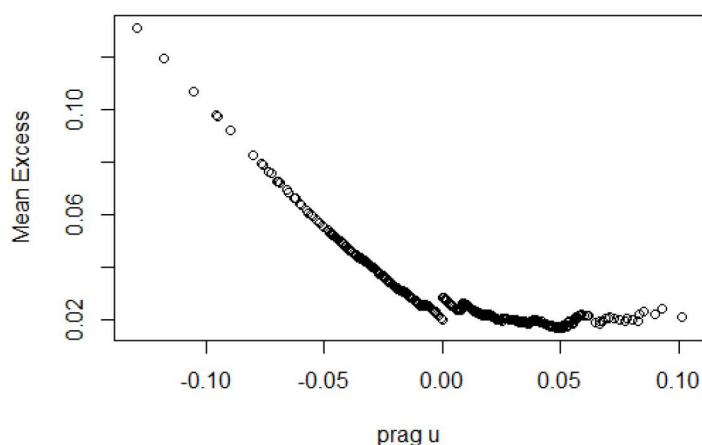
$$e(u) = \frac{\beta + \xi u}{1 - \xi}.$$

Stoga da bi odredili  $u$ , koristimo grafičku metodu na način da definiramo empirijsku ME funkciju:

$$e_T(u) = \frac{1}{N_u} \sum_{i=1}^{N_u} (r_{t_i} - u),$$

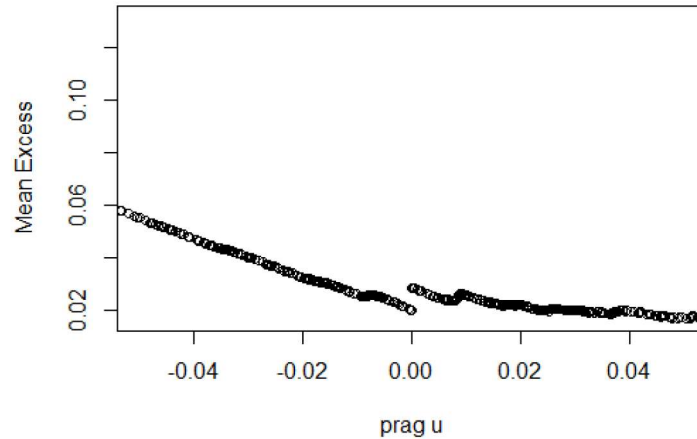
gdje je  $N_u$  broj log-povrata  $r_{t_i}$  koji prelaze prag  $u$ . Prag  $u$  se odabire tako da empirijska ME funkcija  $e_T$  bude približno linearna za  $r > u$ .

**Primjer 2.7.** *Unutar ovog primjera podatke o log-povratima Saponije pokušat ćemo modelirati generaliziranom Pareto distribucijom i zatim izračunati VaR. Prvi korak je određivanje praga  $u$ . Polazišna točka bit će nam graf ME funkcije. Koristimo R pakete evir i quantmod.*



Slika 9: ME funkcija log-povrata Saponije

Ono što možemo primijetiti jest da funkcija linearno pada do okoline točke 0 što nam sugerira da je  $\xi < 0$  i da uzorak zapravo dolazi iz Weibullove distribucije, a zatim usporava i stabilizira pad. Pogledat ćemo поближе što se događa oko 0.



Slika 10: ME funkcija log-povrata Saponije na intervalu  $[-0.05, 0.05]$

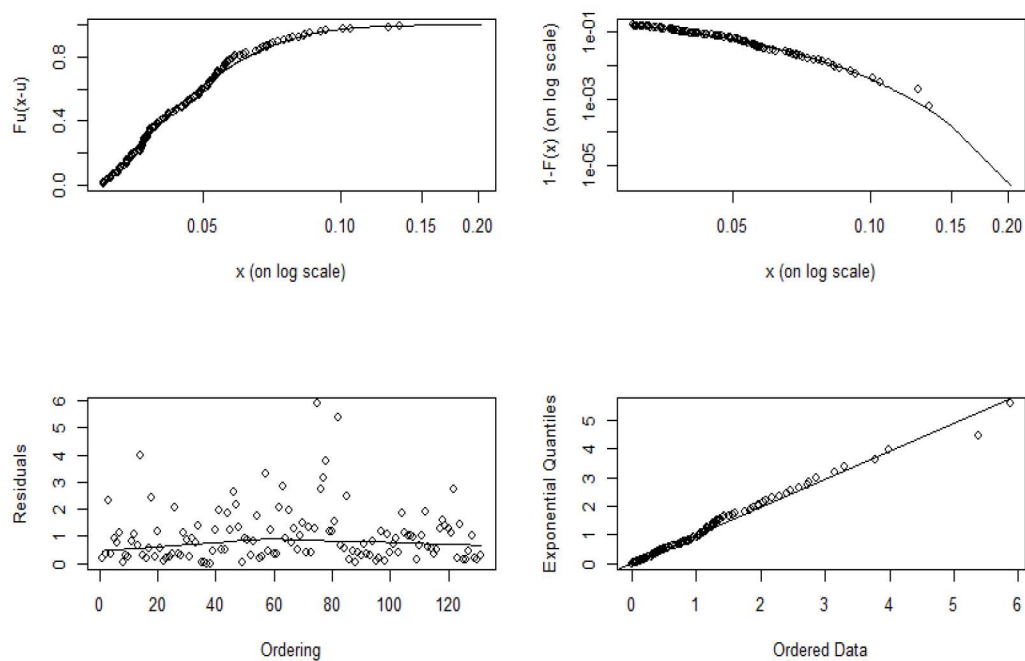
Primjećujemo da se vrijednosti stabiliziraju oko  $u = 0.03$  pa ćemo pokušati podatke modelirati s tom vrijednosti praga koristeći R funkciju `gpd`. Koristimo metodu maksimalne vjerodostojnosti. Dolazimo do sljedećih procjena:

parametar	procjena	standardna greška
$\xi$	-0.05488612	0.079345779
$\beta$	0.02088136	0.002443628

Tablica 4: Procjene parametara i standardna greška procjene

Broj log-povrata koji prelaze prag  $u$  je 113, što čini udio od 13.7% u uzorku log-povrata. Kako bismo se uvjerali jesmo li dobro procijenili parametre, grafički ćemo usporediti dobivene s teorijskim vrijednostima.





Slika 11: Grafički prikazi

Prvi graf prikazuje funkciju distribucije prekoračenja, dok na drugom grafu možemo vidjeti rep pripadne funkcije distribucije. Zaključujemo kako podaci dobro prate pretpostavljenu distribuciju s procijenjenim parametrima. Grafovi u drugom redu daju nam uvid u strukturu reziduala. Zaključujemo da nema razloga sumnjati u kvalitetu procjene parametara, pa možemo prijeći na procjenu VaR-a. Kako je ova teorija namijenjena procjenjivanju ekstremnijih VaR-ova, u sklopu ovog primjera procjenit ćemo  $VaR(0.999)$ . Procjena iznosi 0.1234796 što znači da s vjerojatnošću 99.9% možemo tvrditi da log-povrati u jednom danu neće prijeći 0.1234796.

## Hillov procjenitelj

Neka je  $(X_1, \dots, X_n)$  jednostavan slučajan uzorak iz nepoznate distribucije  $F_X$ , a  $(x_1, \dots, x_n)$  podaci, odnosno jedna realizacija slučajnog uzorka. Neka je  $t : \mathbb{R}^n \rightarrow \mathbb{R}^n$  funkcija koja sortira podatke od najvećeg prema najmanjem, tj.  $t(x_1, \dots, x_n) = (x_{(1)}, \dots, x_{(n)})$ ,  $x_{(1)} > \dots > x_{(n)}$ . Statistiku  $T = t(X_1, \dots, X_n)$  nazivamo zajednička uređajna statistika,  $X_{(1)}$  maksimalna,  $X_{(n)}$  minimalna uređajna statistika, a  $X_{(k)}$   $k$ -ta uređajna statistika,  $k = 2, \dots, n - 1$ .

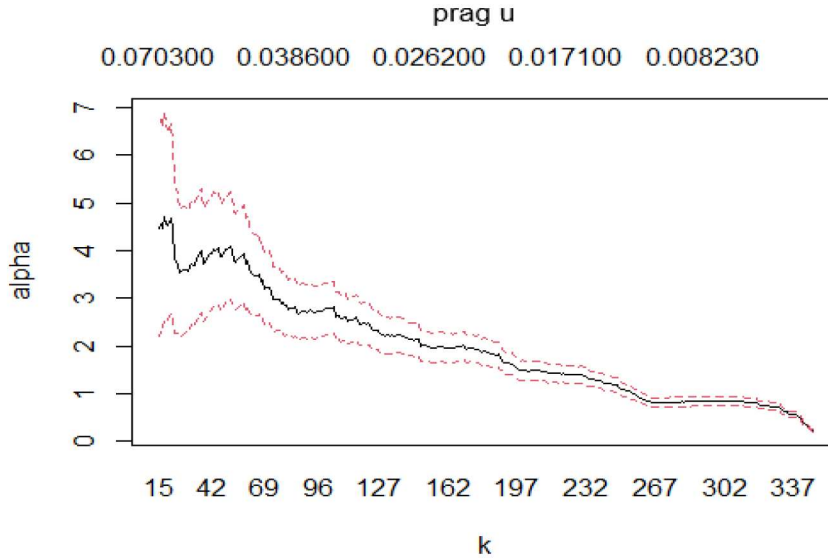
Hillov procjenitelj je semiparametarski procjenitelj repnog indeksa koji se bazira na uređajnim statistikama i kojeg definiramo s:

$$\hat{\xi} = \frac{1}{N_u} \sum_{k=1}^{N_u} \ln \left( \frac{X_{(k)}}{u} \right),$$

pri čemu je  $N_u$  broj opservacija u repu, tj. broj podataka koji prelaze prag  $u$ , a  $X_{(k)}$  predstavlja  $k$ -tu uređajnu statistiku. Ovaj procjenitelj postiže Cramer-Rao donju granicu kada podaci dolaze iz Pareto distribucije pa je ujedno i efikasan i ne može se poboljšati u tom slučaju, no s druge strane, očito je ovaj procjenitelj osjetljiv na odabir praga  $u$ .

U praksi se Hillova procjena očitava iz Hillovog grafa za  $\alpha = 1/\xi$ , tj. grafa koji prikazuje uređene parove  $(k, H_k^{-1})$ , pri čemu je  $H_k = \frac{1}{k} \sum_{i=1}^k \ln \left( \frac{x_{(i)}}{x_{(k+1)}} \right)$ . Ako dođe do stabilizacije vrijednosti  $H_k^{-1}$ , očitamo  $\alpha$  s grafa na način da uzmemo prosječnu vrijednost iz stabilnog područja, no to ponekad nije moguće. Hillov graf je obično prilično nestabilan za mali  $k$ . To je ujedno i nedostatak ovog procjenitelja, graf može pokazivati veću volatilnost i time onemogućiti iščitavanje  $\alpha$  vrijednosti. Osim toga, ovaj procjenitelj ima optimalna svojstva samo ako je distribucija podataka bliska Pareto distribuciji, tj. ako je  $\xi > 0$  i nije lokacijski invarijantan, tj. osjetljiv je na male promjene lokacijskog parametra.

**Primjer 2.8.** *Kako smo u prethodnom primjeru zaključili da je procjena  $\xi$  za podatke kojim baratamo unutar ovog rada manja od 0, tj. da repovi nisu bliski Pareto distribuciji, naslućujemo kako Hillov procjenitelj u ovom slučaju neće moći primijeniti i da iz grafa neće moći iščitati procjenu za  $\alpha$ .*



Slika 12: Hilov graf

Možemo primijetiti veću volatilnost za mali  $k$  kako je i očekivano, a također možemo primijetiti da graf ne pokazuje stabilnost, gotovo da linearno pada. Hillov procjenitelj stoga dalje nećemo koristiti, no služi nam kao dodatna sugestija da je procjena za  $\xi$  bila točna.

### Pickandov procjenitelj

Pickandov procjenitelj za  $\xi$  predstavlja poboljšanje Hillovog procjenitelja, s obzirom da posjeduje bolja svojstva. Bazira se na uređajnim statistikama  $X_{(1)}, \dots, X_{(n)}$  i definiran je na sljedeći način:

$$\hat{\xi}_P = \frac{1}{\ln 2} \ln \frac{X_{(n-k)} - X_{(n-2k)}}{X_{(n-2k)} - X_{(n-4k)}}.$$

Poboljšanja koja posjeduje u odnosu na Hillovog procjenitelja:

- lokacijski je invarijatan,
- konzistentan je procjenitelj za  $\xi \in \mathbb{R}$  (ne zahtjeva  $\xi > 0$ ),
- može ga se koristiti kao indikator koji pokazuje da je modeliranje distribucijom s teškim repovima neprimjereno jer je u tom slučaju  $\xi \leq 0$ .

Za oba procjenitelja VaR računamo tako da uzimamo desni rep funkcije distribucije pa nas zanima sljedeće:

$$1 - p = 1 - P(X > VaR(p)) = P(X \leq VaR(p)) = F(VaR(p)), \quad (4)$$

pri čemu je  $F$  vjerojatnost da slučajna varijabla  $X$  prelazi prag  $u$ . Označimo s  $G$  repnu distribuciju povrata. Kako pretpostavljamo da je to distribucija koja dolazi iz generalizirane Pareto distribucije znamo da vrijedi:

$$G(x) = 1 - \left(\frac{x}{u}\right)^{-\alpha}.$$

Zapišemo li (3) u terminima funkcije distribucije  $F$ , dolazimo do izraza:

$$F_u(x) = \frac{F(u+x) - F(u)}{1 - F(u)} \implies F(u+x) = F_u(x)(1 - F(u)) + F(u).$$

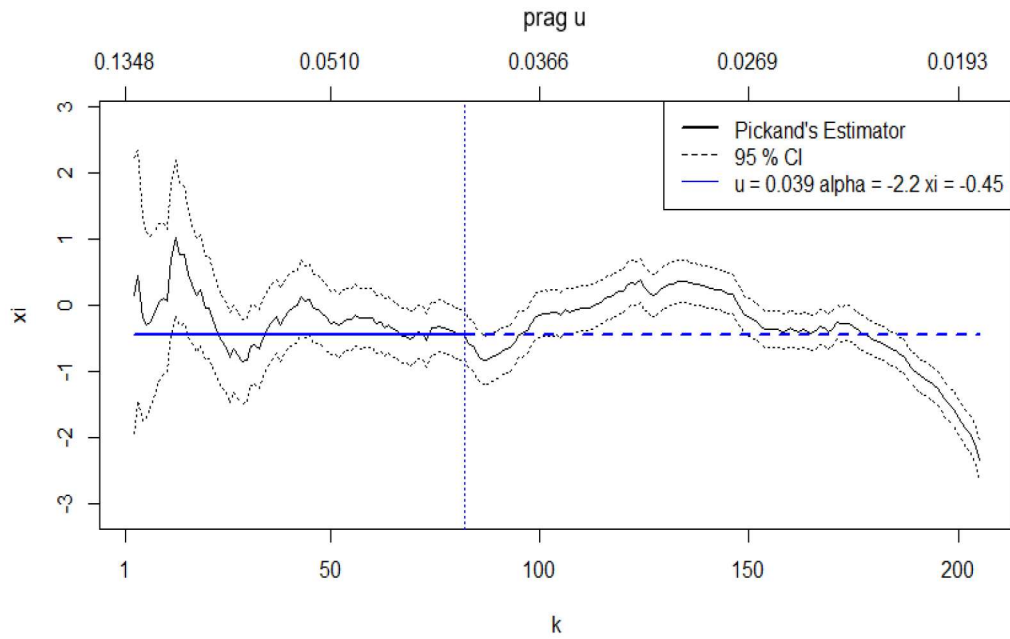
Kako smo već konstatirali, za dovoljno velik  $u$   $F_u$  možemo dobro aproksimirati s  $G_{\xi,\beta}$ , i uzmemo li  $y = u + x$ , dolazimo do izraza:

$$F(y) = G_{\xi,\beta}(x)(1 - F(u)) + F(u).$$

Vjerojatnost da podatak ne prelazi prag  $u$  možemo neparametarski aproksimirati s  $\frac{n-k}{n}$ , gdje  $n$  označava duljinu uzorka, a  $k$  broj onih koji prelaze prag  $u$  te kada iskoristimo relaciju (4), procjenu VaR-a izračunavamo:

$$VaR(p) = u \left(\frac{k}{np}\right)^{\frac{1}{\alpha}}.$$

**Primjer 2.9.** *S obzirom na nešto bolja svojstva Pickandovog procjenitelja u odnosu na Hillov za koji smo u prethodnom primjeru vidjeli da na ovim podacima nije primjenjiv, pokušat ćemo unutar ovog primjera repni indeks, a zatim i VaR procijeniti malo prije opisanom metodom. Postupak se provodi analogno Hillovom procjenitelju, crtamo Pickandov graf na osnovu dostupnih podataka i tražimo područje stabilizacije statistike  $\hat{\xi}_P$ . Za grafički prikaz koristimo R paket `evmix`.*



Slika 13: Pickandov graf za procjenu parametra  $\xi$

*Pickandov procjenitelj daje nam u blizak već procijenjenom u prethodnom primjeru, točnije  $u = 0.039$ . Broj log-povrata koji prelaze ovaj prag je 84 ili 10,2% u promatranom uzorku log-povrata. Procjena za  $\xi$  jednaka je -0.45, a  $VaR(0.999)$  prema dobivenoj formuli je 0.1088613.*

### 3. Problemi u primjeni VaR-a

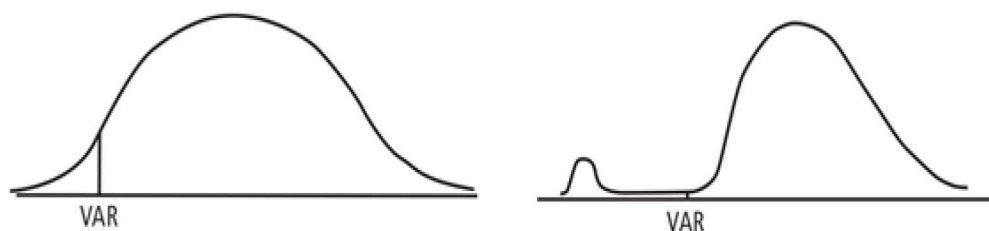
Posljednje poglavlje ovog rada posvetit ćemo kratkoj diskusiji nedostataka VaR-a i problemima u njegovoj primjeni. Objasnit ćemo dva glavna problema: činjenicu da je VaR na kraju ipak samo kvantil distribucije, da je njime lako manipulirati te na kraju ono što se navodi kao glavna zamjerka jest to da nije *koherentna* mjera rizika.

#### 3.1. VaR kao kvantil

Kao što smo vidjeli u samoj definiciji VaR-a, to je mjera koja predstavlja minimalni potencijalni gubitak portfelja, što znači da nam daje najbolji od svih najgorih slučajeva. Samim time znači da zapravo podcjenjuje stvarni, odnosno potencijalni rizik koji investitor ima uz određenu vjerojatnost. Također, čim fiksiramo vjerojatnost uz koju računamo VaR, znači da "ignoriramo", odnosno ne uzimamo u obzir događaje koji se ostvaruju s manjom vjerojatnošću, a koji su ekstremniji i znače veći gubitak.

**Primjer 3.1.** *Usporedimo dva portfelja A i B. Portfelj A neka se sastoji od  $n$  dionica koje u vremenskom periodu od 100 dana nose maksimalni gubitak od 500 kn i on se realizira s vjerojatnošću od 5%. Portfelj B neka se sastoji od  $m$  dionica koje u periodu od 100 dana nose minimalni gubitak od 500 kn koji se također realizira s vjerojatnošću od 5%. Prema  $VaR(0.05)$  ova dva portfelja imali bi jednak rizik, ali portfelj B može donijeti i višestruko veći gubitak od portfelja A.*

Kako je  $VaR(p)$   $p$ -ti kvantil funkcije distribucije log-povrata, oblik distribucije koji dolazi prije ili poslije tog kvantila za VaR zapravo nije bitan, ali u praksi je bitan jer nam može otkriti veću rizičnost i potencijalne probleme s ulaganjem u takvu imovinu.



Slika 14: Distribucije s jednakim VaR-om

Na primjeru slike 14 možemo vidjeti dvije distribucije s jednakim VaR-om, no očito je rizičnost ulaganja u ove imovine različita. Mnogi zato kao dodatnu informaciju VaR-u predlažu promatranje *Expected Shortfalla (ES)*, tj. *očekivanog gubitka*. Dok VaR daje odgovor na pitanje koliki je minimalni gubitak koji mogu pretrpjeti, ES daje odgovor na pitanje koliki je očekivani gubitak, ako se realizirao gubitak veći od VaR-a.

**Definicija 3.1.** *Expected shortfall ili očekivani gubitak definiramo s:*

$$ES = E[Q|Q \leq VaR(p)] = \int_{-\infty}^{VaR(p)} \frac{x}{p} f_Q(x) dx,$$

pri čemu je  $Q$  slučajna varijabla kojom modeliramo gubitak.

Distribucije sa slike 14, iako imaju jednak VaR, imaju različit ES.

### 3.2. Koherentnost

Navedimo prvo definiciju koherentnosti.

**Definicija 3.2.** *Za mjeru  $\varphi : \Omega \rightarrow \mathbb{R}$ , pri čemu je  $\Omega$  familija slučajnih varijabli i  $X, Y \in \Omega$  slučajne varijable, kažemo da je koherentna ako zadovoljava sljedeća svojstva:*

1. *Za  $X \leq Y$  je  $\varphi(X) \geq \varphi(Y)$  (monotonost),*
2.  *$\varphi(X + Y) \leq \varphi(X) + \varphi(Y)$  (subaditivnost),*
3.  *$\varphi(cX) = c\varphi(X)$  za  $c > 0$  (pozitivna homogenost),*
4.  *$\varphi(X + c) = \varphi(X) - c$  za  $c \in \mathbb{R}$  (invarijantnost na translacije).*

Svojstvo monotonosti garantira da ako je  $X$  portfelj koji ne prelazi vrijednost portfelja  $Y$ , onda rizik portfelja  $Y$  ne može biti veći od rizika portfelja  $X$ . Pozitivna homogenost kaže da ako se vrijednost portfelja primjerice udvostruči, udvostručit će se i rizik, no u praksi češće zapravo vrijedi  $\varphi(cX) > c\varphi(X)$ . Četvrto je svojstvo koje kaže da ako dodamo neku konstantu slučajnoj varijabli, što bi se moglo gledati kao da portfelju  $X$  dodamo nerizičnu imovinu, odnosno novac, rizik takvog portfelja jednak je riziku portfelja  $X$  umanjenom za iznos dodatno uloženog novca. Ono što VaR ne zadovoljava i zbog čega nije koherentna mjera rizika je drugo svojstvo subaditivnosti. To svojstvo garantira da je rizik jednog velikog zajedničkog portfelja manji ili jednak sumi rizika pojedinačnih portfelja što potiče diversifikaciju portfelja i u skladu je s modernom teorijom portfelja. Kršenje ovog svojstva demonstrirat ćemo na primjeru.

**Primjer 3.2.** *Neka su  $X$  i  $Y$  dva portfelja s jednakom distribucijom:*

$$X, Y \sim \begin{pmatrix} -1000 & 0 \\ 0.049 & 0.951 \end{pmatrix}.$$

*U ovom slučaju je  $VaR(0.05)=0$  za portfelje  $X$  i  $Y$ , s obzirom da je vjerojatnost da se realizira -1000 manja od 5%.  $VaR(0.01)$  jednak je 1000. Formirajmo zajednički portfelj  $X+Y$ :*

$$X + Y \sim \begin{pmatrix} -2000 & -1000 & 0 \\ 0.002401 & 0.093198 & 0.904401 \end{pmatrix}.$$

*Sada je  $VaR(0.05)=1000$ . Dolazimo do relacije:*

$$VaR(X + Y) = 1000 > VaR(X) + VaR(Y) = 0 + 0 = 0.$$

Može se pokazati da su ES i volatilitnost dvije mjere rizika koje zadovoljavaju svojstvo subaditivnosti. Isto tako, VaR zadovoljava subaditivnost u slučajevima kada imamo normalno distribuirane povrate. Ipak, može se pokazati da je VaR subaditivan kada je repni indeks veći od 2, za detalje vidjeti [3].

U prethodnom primjeru već se moglo naslutiti kako se može manipulirati VaR-om u praksi. Dovoljno je konstruirati portfelj koji će velike gubitke ostvarivati s vjerojatnošću nižom od one s kojom računamo VaR i rizik će proći nezamijećeno. Manageri koji upravljaju portfeljima i koji su u svom poslovanju ograničeni na način da moraju poštovati određenu granicu VaR-a, mogu ga zaobići i izmanipulirati trgovanjem opcijama. Opcija je ugovor koji svom vlasniku daje pravo, ali ne i obvezu da u unaprijed dogovorenom trenutku kupi ili proda neku financijsku imovinu. Kupovanjem i prodavanjem put i call opcija moguće je portfelj dekomponirati u manje cjeline u kojima će biti mogući veći gubici ali s manjom vjerojatnošću.



## Literatura

- [1] V. O. ANDREEVA, S. E. TINYKOV, O. P. OVCHINNIKOV, G. P. PARAHIN, *Extreme Value Theory and Peaks Over Threshold Model in the Russian Stock Market*, Journal of Siberian Federal University, Engineering & Technologies 1 (2012), 5; 111-121.
- [2] J. CEROVIĆ, M. LIPOVINA-BOŽOVIĆ, S. VUJOŠEVIĆ, *A Comparative Analysis of Value at Risk Measurement on Emerging Stock Markets: Case of Montenegro*, Business Systems Research 6(2015), 1; 36–55.
- [3] J. DANIELSSON, *Financial risk forecasting*, John Wiley & Sons Ltd, West Sussex, United Kingdom, 2011.
- [4] P. EMBRECHTS, C. KLUPPELBERG, T. MIKOSCH, *Modelling Extremal Events for insurance and finance*, Springer, New York, 1997.
- [5] V. FERNANDEZ, *Extreme Value Theory: Value at Risk and returns dependence around the world*, University of Chile, Department of Industrial Engineering, Santiago, 2003.
- [6] S. MANGANELLI, R. F. ENGLE, *Value at Risk models in finance*, European Central Bank (2001), Working paper no. 75.
- [7] A. J. MCNEIL, *Extreme Value Theory for Risk Managers*, ETH Zentrum, Departement Mathematik, Zurich, 1999.
- [8] T. NOVAK, *Semiparametarska procjena VaR-a*, Sveučilište J. J. Strossmayera u Osijeku, Odjel za matematiku, 2017.
- [9] D. RUPPERT, *Statistics and Finance*, Springer, New York, 2004.
- [10] D. RUPPERT, D. S. MATTESON, *Statistics and Data Analysis for Financial Engineering*, Springer, New York, 2015.
- [11] B. WILLIAMS, *GARCH(1,1) models*, Ruprecht-Karls-Universität at Heidelberg, Fakultät für Mathematik und Informatik, 2011.

**Sažetak:** Cilj rada je upoznati se s osnovnim metodama procjene Value at Risk-a. VaR je mjera rizika koja je na financijskom tržištu stekla široku primjenu i predstavlja minimalni gubitak tijekom određenog vremena uz određenu vjerojatnost  $p$ . Postoje 3 metode procjene: neparametarska, parametarska i semiparametarska. Neparametarska procjena bazira se na povijesnim vrijednostima log-povrata gdje VaR procjenjujemo kao  $p$ -ti kvantil tih vrijednosti, dok se parametarska procjena provodi uz izgradnju EWMA i GARCH modela. Semiparametarsku procjenu gradimo uz pomoć teorije ekstremnih vrijednosti i procjene repnog indeksa. Svaka od metoda demonstrirana je na primjeru cijena dionica domaće firme.

**Ključne riječi:** Value at Risk, VaR, povijesne simulacije, EWMA, GARCH, teorija ekstremnih vrijednosti, generalizirana Pareto distribucija

## Methods for estimating VaR

**Abstract:** The goal of this paper is to learn about basic methods for estimating Value at Risk. VaR is a highly applicable financial risk measure and represents minimum loss over a period of time with probability  $p$ . There are 3 estimation methods: nonparametric, parametric and semiparametric. Nonparametric estimation is based on historical values of log-returns, while the parametric estimation is based on building EWMA and GARCH models. The semiparametric estimation is built upon the extreme value theory and estimating the tail index. Each method is demonstrated on the stock prices of a local firm.

**Key words:** Value at Risk, VaR, historical simulation, EWMA, GARCH, extreme value theory, generalized Pareto distribution

## Životopis

Zovem se Vedrana Tokić i rođena sam 18.4.1997. u Osijeku. Nakon završetka osnovne škole upisujem I. gimnaziju u Osijeku. Tijekom osnovnoškolskog i srednjoškolskog obrazovanja sudjelovala sam na županijskim i državnim natjecanjima iz hrvatskog jezika. Nakon uspješnog završetka škole, 2015. godine upisujem Preddiplomski studij matematike na Odjelu za matematiku u Osijeku. Završavam ga 2019. godine s temom završnog rada *Elementarne metode za rješavanje diofantskih jednadžbi* i iste godine upisujem diplomski studij, smjer Financijska matematika i statistika. U trenutku obrane diplomskog rada zaposlena sam u Žito d.o.o. kao pripravnik u financijama.