

**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

DIPLOMOVÁ PRÁCE

Bc. Patrik Lipták

**Výpočet Value-at-Risk s využitím teorie
extrémních hodnot**

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Radek Hendrych, Ph.D.

Studijní program: Matematika

Studijní obor: Finanční a pojistná matematika

Praha 2017

UNIVERZITA KARLOVA
Matematicko-fyzikální fakulta

Katedra pravděpodobnosti a matematické statistiky

Akademický rok: 2016/2017

ZADÁNÍ DIPLOMOVÉ PRÁCE

Jméno a příjmení: **Patrik Lipták**

Studijní program: **Matematika**

Studijní obor: **Finanční a pojistná matematika**

Děkan fakulty Vám podle zákona č. 111/1998 Sb. určuje tuto diplomovou práci:

Téma práce v českém jazyce: **Výpočet Value-at-Risk s využitím teorie extrémních hodnot**

Téma práce v anglickém jazyce: **Value-at-Risk Calculation Using Extreme Value Theory**

Zásady pro vypracování:

Uchazeč podrobně pojedná o metodách výpočtu kvantitativní rizikové míry Value-at-Risk (VaR) založených na teorii extrémních hodnot. Případně uváže jejich modifikaci či rozšíření. Následně provede porovnání jednotlivých přístupů prostřednictvím rozsáhlejší simulační studie či aplikace na reálná tržní data.

Seznam odborné literatury:

Berkowitz, J. & O'Brien, J. (2002). How Accurate are the Value-at-Risk Models at Commercial Banks, *Journal of Finance* 57, pp. 1093-1112.

Franke, J., Härdle, W. K., & Hafner, C. M. (2015). *Statistics of Financial Markets: An Introduction*, 4. vydání, New York: Springer.

Holton, G. A. (2003). *Value-at-Risk: Theory and Practice*, San Diego: Academic Press.
[druhé vydání z roku 2014 dostupné na www.value-at-risk.net]

Morgan, J. (1996). *RiskMetrics - Technical document*, 4. vydání, New York: Morgan Guaranty Trust Company.

Tsay, R. S. (2010). *Analysis of Financial Time Series*, 3. vydání, New York: Wiley.

Literatura citovaná ve výše uvedených zdrojích.


Vedoucí diplomové práce: **RNDr. Hendrych Radek, Ph.D.**

Navrhování oponenti:

Konzultanti:

Datum zadání diplomové práce: 25.10.2016

Termín odevzdání diplomové práce: dle harmonogramu příslušného akademického roku


.....
Vedoucí katedry

V Praze dne 2.2.2017

Univerzita Karlova
Matematicko-fyzikální fakulta
Studijní oddělení
121 16 Praha 2, Ke Karlovu 3
IČ: 00216208, DIČ: CZ00216208
Tel.: 951 551 412, 951 551 111


.....
Děkan

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V Praze dne 5.5.2017

Podpis autora

Název práce: Výpočet Value-at-Risk s využitím teorie extrémních hodnot

Autor: Bc. Patrik Lipták

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Radek Hendrych, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Táto diplomová práca skúma teóriu extrémnych hodnôt a jej aplikáciu vo finančnom rizikovom manažmente s hlavným zameraním na výpočet známej rizikovej miery – hodnoty v riziku (VaR). V prvej, teoretickej časti, je poskytnutý dôkladný prierez teóriou extrémnych hodnôt s dôrazom na fundamentálne zistenia a ich dôsledky a zhrnutie metód založených na tejto teórii, špecifickejšie metóda blokových maxím, Hillova metóda a metóda založená na presahoch cez medzu. Ďalej sú zdôraznené štatistické problémy, ktoré pri aplikácii týchto metód môžu vzniknúť a ako sa s nimi vysporiadať. Druhá časť je tvorená rozsiahlou empirickou štúdiou, kde sa tieto postupy spoločne s teoretickými zisteniami aplikujú na reálne tržné dáta cien akciového indexu Dow Jones Industrial Average, akcií americkej banky JPMorgan a akciového indexu Russell 2000 za účelom obecnějšího porovnania počínania jednotlivých metód založených na teórii extrémnych hodnôt spoločne s tradičnou metodológiou RiskMetrics.

Klíčové slová: backtesting, RiskMetrics, teória extrémnych hodnôt, Value-at-Risk

Title: Value-at-Risk Calculation Using Extreme Value Theory

Author: Bc. Patrik Lipták

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Radek Hendrych, Ph.D., Department of Probability and Mathematical Statistics

Abstract: This diploma thesis studies extreme value theory and its application in financial risk management, when focusing on computation of well-known risk measure – Value at Risk (VaR). The first part of the thesis reviews theoretical background. In particular, it rigorously discusses the extreme value theory when emphasizing fundamentals theorems and their consequences followed by the summary of methods based on this theory, specifically, Block Maxima method, Hill method and Peaks over Threshold method. Moreover, specific issues that may arise in such applications and ways how to deal with these problems are described. The second part of the thesis contains extensive empirical study, which together with theoretical findings applies each of the examined method to real market data of the closing prices of Dow Jones Industrial Average stock index, stocks of JPMorgan and stock index Russell 2000 in order to compare methods based on extreme value theory together with the classic methodology RiskMetrics.

Keywords: backtesting, extreme value theory, RiskMetrics, Value-at-Risk

Rád by som poďakoval všetkým, ktorí mi pri písaní tejto diplomovej práce pomohli. Zvláštne poďakovanie patrí môjmu vedúcemu RNDr. Radkovi Hendrychovi, Ph.D. za možnosť konzultovať všetky nejasnosti a hlavne za všetky cenné rady a pripomienky.

Obsah

Zoznam použitých skratiek	3
Úvod	4
1 Základné definície	6
1.1 Riziková miera	6
1.2 Value-at-Risk	6
1.3 Empirické vlastnosti finančných výnosov	9
2 Motivácia	11
2.1 Metodológia RiskMetrics	11
3 Teória extrémnych hodnôt	13
3.1 Štandardné rozdelenia extrémnych hodnôt	13
3.2 Odhady parametrov	17
3.2.1 Parametrický odhad: Blokové maximá	18
3.2.2 Výpočet VaR pomocou BM metódy	19
3.2.3 Neparametrický odhad: Hillova metóda	20
3.2.4 Výpočet VaR pomocou Hillovej metódy	21
3.3 Prístup pomocou presahov cez medzu	21
3.3.1 Metóda odhadu	23
3.3.2 Presahy cez vyššiu medzu	23
3.3.3 Výpočet VaR pomocou POT metódy	25
3.3.4 VaR pre dlhší časový horizont	25
4 Zovšeobecnenie na finančné časové rady	26
4.1 Extremálny index	26
4.1.1 Odhady extrémálneho indexu	27
4.1.2 Výpočet VaR s využitím extrémálneho indexu	28
4.2 Využitie vysvetľujúcich premenných	28
4.3 Očistenie rady výnosov	28
5 Empirická štúdia: metodológia	31
5.1 Výber dát	31
5.2 Porovnanie metód	32
5.2.1 Nepodmienené testovanie pokrytia	32
5.2.2 Test nezávislosti	33
5.2.3 Podmienené testovanie pokrytia	33
6 Empirická štúdia: výsledky	34
6.1 Úvodná dátová analýza a deskriptívne štatistiky	34
6.2 Modifikácia dát	36
6.3 Aplikácia metód na okno	38
6.3.1 RiskMetrics	38
6.3.2 BM metóda	38
6.3.3 Hillova metóda	40

6.3.4	POT metóda	41
6.3.5	Porovnanie	42
6.4	Použitie rolovacieho okna	43
6.4.1	Dow Jones Industrial Average	43
6.4.2	JPMorgan Chase & Co.	46
6.4.3	Russell 2000	49
6.5	Limitácie štúdie	52
	Záver	66
	Zoznam použitej literatúry	67
	Zoznam obrázkov	70
	Zoznam tabuliek	72
	Prílohy	73

Zoznam použitých skratiek

AR - Autoregressive Process
BCBS - Basel Committee on Banking Supervision
BM - Block Maxima
CC - Conditional Coverage
df - degrees of freedom
DJI - Dow Jones Industrial Average
ES - Expected Shortfall
EVT - Extreme Value Theory
EWMA - Exponentially Weighted Moving Average
FRTB - Fundamental Review of the Trading Book
GARCH - Generalized AutoRegressive Conditional Heteroscedasticity Process
GEVD - Generalized Extreme Value Distribution
GPD - Generalized Pareto Distribution
iid - independent and identically distributed
JB - Jarque-Bera
JPM - JPMorgan
LB - Ljung-Box
LM - Lagrange Multipliers
MDA - Maximal Domain of Attraction
POT - Peaks Over Threshold
RUT - Russell 2000
UC - Unconditional Coverage
VaR - Value-at-Risk

Úvod

Cieľom tejto práce bude skúmať teóriu extrémnych hodnôt a jej potenciál vo finančnom rizikovom manažmente. Práca poskytne dôkladný a rozsiahly priezrez teoretickými základmi teórie extrémnych hodnôt a zdôrazní štatistické problémy a limitácie tejto teórie. Ďalej bude pojednávať o tom, ako môže byť táto teória aplikovaná na finančné dáta a špecifické problémy, ktoré môžu pri týchto aplikáciách vzniknúť. Použitím rôznych metód porovná ich schopnosť presne odhadnúť veľmi známu rizikovú mieru – Value-at-Risk.

Jedným z dôvodov použitia rizikových mier je odhad možných budúcich strát portfólia finančnej inštitúcie a výpočet minimálnej kapitálovej požiadavky, pri akej je inštitúcia schopná absorbovať tieto potenciálne straty. Riziková miera, na ktorú sa bude sústrediť táto práca je hodnota v riziku (VaR z angl. Value-at-Risk), ktorú používa napr. direktíva Solvency II (ďalej to boli regulačné rámce Basel I a II) pri výpočte minimálnej kapitálovej požiadavky.¹ Následne sa práca v krátkosti zmieni o stále viac a viac použíwanejšej rizikovej miere, a to podmienenej hodnote v riziku (ES z angl. Expected Shortfall), ktorá je od januára 2016 uprednostňovaná regulačným rámcom Basel III.²

V našich diskusiách, ako aj v empirickej štúdii, sa práca bude koncentrovať výlučne len na tržné riziko. Tržné riziko je riziko zmeny hodnoty finančnej pozície z dôvodu zmien hodnôt podkladových inštrumentov ako sú napríklad akcie, dlhopisy, ceny komodít, výmenné kurzy, . . . (McNeil a kol., 2005, str. 463). Avšak, banky a ostatné finančné inštitúcie sú vystavované aj iným triedam rizika ako je napr. kreditné alebo operačné riziko. Hoci sú tieto triedy rizík rovnako dôležité, nebudú v tejto práci ďalej rozoberané. Taktiež bude v ďalšom priebehu upustené od diskusií alebo aplikácií ostatných rizikových mier ako je VaR a ES. Pre záujemcov, Cipra (2015, kap. 9 a 10) vo svojej knihe pokryl ostatné triedy rizík a poskytol širokú škálu alternatívnych rizikových mier.

Práca je rozdelená do 6 kapitol. Prvá kapitola poskytuje základné štatistické definície rizikových mier, na ktoré sa práca sústreďuje. Druhá kapitola ukazuje prečo je vhodné zaoberať sa zakomponovaním teórie extrémnych hodnôt do kvantitatívneho rizikového manažmentu pre účely výpočtu hodnoty v riziku. Tretia kapitola prezentuje teoretický rámec celej práce. Obsahuje dôkladné zhrnutie metód založených na teórii extrémnych hodnôt a ich následne použitie pre výpočet skúmanej rizikovej miery. Štvrtá kapitola pozostáva zo série postupov, ktoré sa kvôli často nereálnym predpokladom na reálne tržné dáta snažia dané metódy založené na teórii extrémnych hodnôt zovšeobecniť tak, aby ich bolo možné použiť v praxi. Posledný postup z tejto kapitoly tvorí vlastný ad hoc prístup, ktorý sa použije na rozsiahlu empirickú štúdiu v kapitolách 5 a 6. V piatej kapitole sa najprv predstaví metodologický postup, akým boli dáta vyberané a ako boli metódy medzi sebou porovnávané. Šiesta kapitola následne s využitím teoretických

¹<http://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=OJ:L:2015:012:FULL&from=CS>, navštívené 24.03.2017.

²<https://www.bis.org/bcbs/publ/d352.pdf>, navštívené 24.03.2017.

poznatkov prezentuje na troch vybraných dátových súboroch dôkladnú analýzu dát, výsledky odhadov VaR a spoločné porovnanie metód s množstvom grafických výstupov. Práca je zakončená limitáciami, implikáciami, návrhmi na ďalší výskum a použiteľnosťou týchto metód v praxi.

1. Základné definície

Táto kapitola je venovaná prevažne zavedeniu rizikovej miery VaR. V prvej sekcii zavedieme rizikové miery v širšom zmysle a budeme pojednávať o tzv. axiómoch koherencie, ktoré by mali všetky „rozumné“ rizikové miery spĺňať. V nasledujúcej sekcii poskytneme základné štatistické definície, ktoré budú nevyhnutné pre neskoršiu definíciu VaR. Následne prejdeme k samotnej definícii VaR a ES a ich spoločnému porovnaniu. Kapitulu zakončíme empirickými vlastnosťami finančných časových radov, na ktoré sa bude táto práca sústrediť.

1.1 Riziková miera

Artzner a kol. (1999) sa zhodli, že vhodná riziková miera by mala spĺňať skupinu vlastností, označované ako axiomy koherencie. Nech L je náhodná veličina reprezentujúca stratu portfólia zloženého z rizikových aktív s distribučnou funkciou $F_L(l) = P(L \geq l)$. Nech sú straty portfólia reprezentované množinou \mathcal{M} , t. z. $L \in \mathcal{M}$. Rizikové miery sú reálne funkcie $\rho : \mathcal{M} \rightarrow \mathbb{R}$. Množstvo $\rho(L)$ reprezentuje kapitál potrebný na pokrytie straty L .

Riziková miera ρ je *koherentnou* mierou rizika, ak pre $L, L_1, L_2 \in \mathcal{M}$ platia nasledujúce axiomy:

- Monotónia: $L_1 \leq L_2 \Rightarrow \rho(L_1) \leq \rho(L_2)$.
- Pozitívna homogenita: $\rho(\lambda L) = \lambda \rho(L)$, $\forall \lambda > 0$.
- Translančná invariancia: $\rho(L + a) = \rho(L) + a$, $\forall a \in \mathbb{R}$.
- Subaditivita: $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$.

Monotónia vraví, že väčšie straty vyžadujú väčšiu alokáciu kapitálu. Pozitívna homogenita pojednáva o tom, že kapitál požadovaný na pokrytie straty je proporčný veľkosti straty. Translančná invariancia vraví, že ak aj strata väčšia o deterministické množstvo a , potom kapitál potrebný na pokrytie tejto straty je väčší o a jednotiek. Predpoklad subaditivity jednoducho pojednáva o intuitívnej vlastnosti, kde by riziko malo byť redukované alebo aspoň nenavýšené pri diverzifikácii, t. z. kapitál potrebný na pokrytie straty portfólia, ktoré je tvorené kombináciou dvoch portfólií by nemal byť vyšší ako suma kapitálov potrebných na pokrytie strát každého portfólia zvlášť.

1.2 Value-at-Risk

Ako sme v úvode kapitoly spomínali, začneme zavedením nevyhnutných štatistických definícií (pre ich originálne znenie viď (McNeil a kol., 2005, str. 38–39)).

Definícia 1 (Zobecnená inverzná a kvantilová funkcia).

- Nech $h : \mathbb{R} \rightarrow \mathbb{R}$ je neklesajúca funkcia. Zobecnenú inverznú funkciu definujeme ako:

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

kde použijeme konvenciu, že infimum prázdnej množiny je ∞ .

- Pri danej distribučnej funkcii F je zobecnená inverzná funkcia F^{\leftarrow} nazývaná kvantilová funkcia distribučnej funkcie F . Pre $\alpha \in (0,1)$ budeme α kvantilom distribučnej funkcie F nazývať:

$$q_{\alpha}(F) := F^{\leftarrow}(\alpha) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}.$$

Poznámka. Ak F je spojitá a striktné rastúca, kvantilová funkcia je rovná bežnej kvantilovej funkcii $F : q_{\alpha}(F) = F^{-1}(\alpha)$.

Nech $F_L(l) = P(L \geq l)$ značí ako v predošlej sekcii distribučnú funkciu strát. Pri porovnávaní dvoch portfólií sa naskytá otázka, ktoré portfólio je rizikovejšie. Chceme nájsť ukazateľ popisujúci rôzne riziká, t. j. chceme definovať rizikovú mieru, ktorá by bola schopná tieto portfólia porovnať. Prvou intuitívnou možnosťou by bola maximálna možná strata, ktorá je rovná $\inf\{l \in \mathbb{R} : F_L(l) = 1\}$. Lenže táto miera by mohla byť použitá iba v prípade, keby straty predstavovali konečné čísla. My sa budeme pozeráť na finančné straty ako na neobmedzené vo všeobecnosti. Preto sa ďalej budeme zaoberať rizikovými mierami, ktoré dokážu poskytnúť konečný výstup aj v prípade neobmedzených strát.

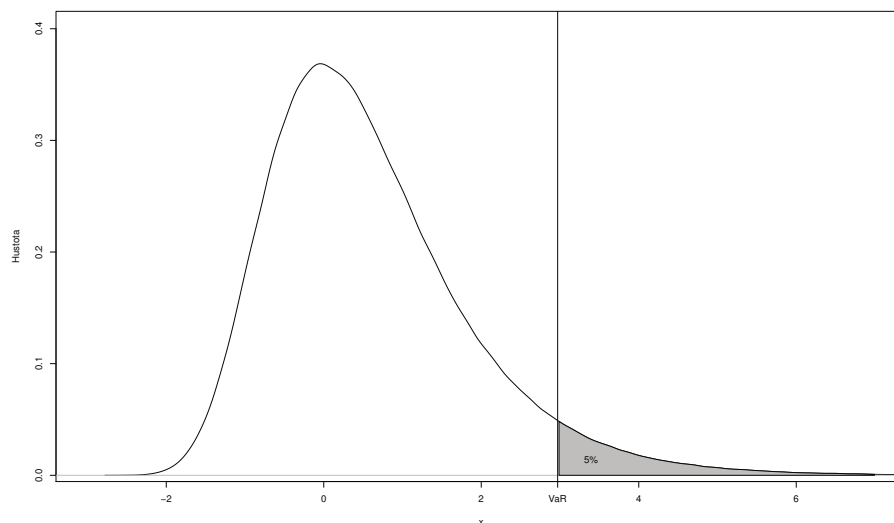
Riziková miera VaR je vhodným kandidátom. Vo všeobecnosti sa jedná o odhad maximálnej straty, ku ktorej môže dôjsť s predpísanou spoľahlivosťou v stanovenom budúcom období.

Definícia 2 (Hodnota v riziku - VaR). *VaR portfólia na konfidenčnej hladine $\alpha \in (0,1)$ je daný najmenším číslom l takým, že pravdepodobnosť, že strata L prekročí l nie je väčšia ako $(1 - \alpha)$. Formálne:*

$$VaR_{\alpha}(L) = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}.$$

Kvôli jednoduchosti budeme v celom ďalšom priebehu práce značiť VaR na konfidenčnej hladine α ako VaR_{α} . Bežne používané hodnoty α vo finančnom sektore sú 0,95 alebo 0,99.

Ukazateľ VaR je obľúbený vďaka schopnosti opísať finančné riziko portfólia jediným číslom. Jeho jednoduchosť ho robí taktiež atraktívnou mierou rizika kvôli ľahkej interpretovateľnosti a porozumeniu i pre širšiu nematematickú verejnosť v porovnaní s ostatnými mierami rizika. Navyše, je ho možné relatívne jednoducho spočítať a je jednoducho graficky interpretovateľný (viď obr. 1.1 pre 95% VaR Gumbelovho rozdelenia). Na druhú stranu je VaR vystavený určitej kritike, ktorá by mala byť zdôraznená. Po prvé, VaR nie je koherentnou mierou rizika, pretože nespĺňa predpoklad subaditivity. Po druhé, netušíme ako sa správa rozdelenie strát za hodnotou VaR. Za touto medzou môžu nasledovať signifikantne vyššie straty s ešte relatívne veľkou pravdepodobnosťou, ktoré môžu mať devastačný dopad na niektoré finančné inštitúcie.



Obr. 1.1: 95% VaR Gumbelovho rozdelenia.

Expected Shortfall

Druhá diskutovaná riziková miera, Expected shortfall (ES), sa snaží zabezpečiť opatrnejšie zachytenie „chvostového rizika“ a kapitálovej adekvátnosti počas nepriaznivých období na finančnom trhu. Basel Committee on Banking Supervision (BCBS) v dokumente „Fundamental Review of the Trading Book“ (FRTB), ňou navrhuje nahradiť vyššie diskutovanú mieru VaR.³

Definícia 3 (Podmienená hodnota v riziku - ES). *Pre stratu L s $\mathbb{E}|L| < \infty$ a distribučnú funkciu F_L , ES na konfidenčnej hladine $\alpha \in (0,1)$ je definovaná ako:*

$$ES_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 q_\varphi(F_L) d\varphi,$$

kde $q_\varphi(F_L) = F_L^{\leftarrow}(\varphi)$ je kvantilová funkcia F_L .

Poznámka. Kvôli jednoduchosti budeme podmienenú hodnotu v riziku na konfidenčnej hladine α opäť značiť len ako ES_α . Ak je stratová funkcia F_L spojitá, ES_α môže byť chápaná ako priemerná strata za predpokladu, že hodnota VaR_α je prekročená. Presnejšie:

$$ES_\alpha = \mathbb{E}(L|L \geq VaR_\alpha).$$

Z konštrukcie ES vyplýva, že ES bude stále konzervatívnejšia riziková miera ako VaR. BCBS to vzala do úvahy a 99% VaR nahrádza 97,5% ES. Na rozdiel od VaR je ES koherentnou mierou rizika a teda spĺňa aj predpoklad subaditivity. Avšak, môže to byť len teoretický benefit, pretože aj keď akademické príklady, kde VaR porušuje predpoklad subaditivity, môžu byť vytvorené jednoducho, podľa tržných praktikov má tento chýbajúci predpoklad zriedkakedy praktické dôsledky. Ďalej, odhliadnuc od matematických vlastností, v praxi je

³<http://www.bis.org/publ/bcbs265.pdf>, navštívené 24.03.2017.

skutočný benefit pri prechode k ES pre rizikový manažment limitovaný dostupnými dátami (Voigt, 2015).

Ďalej, pre výpočty ES potrebujeme pracovať s ťažko odhadnuteľnými extrémnymi udalosťami, ktoré sa vyskytnú s veľmi malou pravdepodobnosťou ale majú veľký dopad. Citlivosť v chvostoch rozdelenia strát, ktorú BCBS považuje za hlavnú výhodu ES sa môže naopak ukázať ako veľká nevýhoda. Pri výpočte VaR sa totižto používajú empirické odhady založené na minulých hodnotách obsahujúce rádovo stovky obchodných dní. Avšak do výpočtu ES vstúpi len veľmi málo pozorovaní a výpočet strednej hodnoty napr. z dvoch alebo troch hodnôt, každej s vysokým dopadom zvyšuje riziko instability. Z tohto dôvodu musí byť analyzovaná zmyslupnosť výsledného ES. Ďalšou nevýhodou je, že pri ES je prakticky nemožná možnosť spätného testovania, keďže táto miera závisí na tvare chvostu stratovej distribučnej funkcie, ktorá je vo všeobecnosti neznáma (Voigt, 2015).

Taktiež je dôležité zdôrazniť, že nahradenie VaR mierou ES navrhnutou FRTB je relevantné pre tržné riziko a neovplyvní výpočty kapitálovej požiadavky pre ostatné triedy rizík. VaR má teda stále rozsiahle použitie a je nepravdepodobné, že sa jeho dôležitosť zmenší.

1.3 Empirické vlastnosti finančných výnosov

Keďže sa v empirickej štúdii budeme zaoberať časovými radmi, ktoré zachycujú vývoj hodnoty akcií v čase, poskytneme v tejto sekcii súhrn ich typických vlastností. Vývoj cien akcií v čase väčšinou tvorí nestacionárny proces (pre stacionárny proces viď definíciu (Cipra, 2008, str. 328)), ktorý je zvyčajne integrovaný, rádu 1. Preto, je bežné modelovať relatívne zmeny cien alebo logaritmus týchto relatívnych zmien (Cont, 2001). Relatívnymi zmenami cien budeme nazývať rad $\{R_t\}_{t \in \mathbb{N}}$, kde:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}, \quad t = 2, 3, \dots, \quad (1.1)$$

a P_t je cena akcie v čase t , $P_t \neq 0$, $\forall t$. Rad $\{R_t\}_{t \in \mathbb{N}}$ budeme v celom ďalšom priebehu jednoducho označovať ako *finančné výnosy* alebo skráteno len *výnosy*. Pri spojitom úročení, ktoré je určené vzťahom $P_t = P_{t-1}e^{r_t}$ dostávame tzv. *logaritmické výnosy*:

$$r_t = \ln \frac{P_t}{P_{t-1}} = \ln(1 + R_t), \quad t = 2, 3, \dots, \quad (1.2)$$

ktoré sú častejšie analyzované ako (1.1), kvôli viacerým dôvodom, napr. jednoduchému výpočtu logaritmických výnosov pre dlhší časový horizont alebo vhodnejším štatistickým vlastnostiam. Z Taylorovho rozvoja ale pre $\ln(1+R_t)$ pre hodnoty R_t blízke nule plynie $r_t \sim R_t$.

Finančné výnosy majú tendenciu vykazovať dočasnú závislosť v druhom momente. Inak povedané, zatiaľ, čo výnosy sa zdajú byť nekorelované, štvorcové (taktiež absolútne) výnosy sa zdajú byť vysoko korelované, čo spôsobuje v čase sa

meniacu volatilitu a tzv. zhlukovanie volatilit. Zhlukovanie volatility je tendencia volatility finančných trhov objavovať sa v zhlukoch vysokých a nízkych volatilit, t. j. veľké (alebo malé) výkyvy výnosov sú očakávané skôr po väčších (resp. menších) predchádzajúcich výkyvoch (Cipra, 2008, str. 377). Taktiež, Black (1976) zistil, že poklesy výnosov akcií zvyknú navýšiť volatilitu viac ako nárasty výnosov rovnakej veľkosti. Tento fenomén je známy ako pákový efekt. Ďalej, Fama (1965) zistil, že finančné výnosy majú zvyčajne rozdelenie s ťažkými chvostami alebo tzv. leptokurtické rozdelenie. Tieto rozdelenia sú špicatejšie okolo stredu, pričom na koncoch je ich hustota väčšia a v ramenách menšia v porovnaní s normálnym rozdelením s rovnakou strednou hodnotou a rozptylom. Význačnou charakteristikou takýchto rozdelení býva kladný koeficient špicatosti (Cipra, 2008, str. 377).

2. Motivácia

2.1 Metodológia RiskMetrics

Pred tým, než začneme s výstavbou zložitejších modelov, sa pozrieme na jednoduchší prístup k výpočtu hodnoty v riziku navrhnutý J. P. Morganom. Táto metóda je označovaná ako RiskMetrics (Morgan, 1996) a svojho času sa stala veľmi populárnou a veľmi rýchlo sa rozšírila po celom svete. Aj keď spätné testovania vykonané najprv J. P. Morganom a neskôr ostatnými účastníkmi trhu odporúčali túto metódu, neskôr sa ukázalo, že jeho základné predpoklady sú otáznave z niekoľkých uhlov pohľadu (Alexander, 1996). Navyše, existencia ťažkých chvostov v reálnych tržných dátach diskutovaná v predošlom odstavci je v jasnom konflikte s predpokladom normálne rozdelených výnosov, čo vedie k veľkému podhodnoteniu rizika.

Poznámka. Metodológia RiskMetrics sa aplikuje na (kladné) výnosy aktív. My sa budeme v celom ďalšom priebehu práce sústrediť na straty, teda záporné výnosy, preto naším záujmom nebudú ľavé chvosty distribučnej funkcie výnosov, ale budú to pravé konce distribučných funkcií strát. Uvidíme ale, že hlavným predpokladom budú normálne rozdelené výnosy a vďaka symetrii normálneho rozdelenia nebude problém prevrátiť analýzu výnosov na nami požadovanú analýzu strát. Metodológiu teda popíšeme v originálnom znení a pri neskorších praktických aplikáciách budeme počítat s vyššie uvedenou modifikáciou.

Ako sme v minulej sekcii spomínali, jedna z typických vlastností výnosov je zhlukovanie volatilit. Metodológia RiskMetrics preto z tohto dôvodu modeluje volatilitu resp. podmienený rozptyl časového radu reprezentujúceho výnosy prostredníctvom MA procesu (pre definíciu viď (Cipra, 2008, str. 333)). Stredná hodnota výnosov je zvyčajne podstatne menšia ako podmienený rozptyl, preto je tento ukazovateľ v metodológii RiskMetrics zanedbaný a podmienený rozptyl zostáva ako jediný parameter podmienenej distribučnej funkcie. Ďalej metóda kladie veľmi silný predpoklad na výnosy, ktoré sú podmienené normálne rozdelené. Základom celej metodológie RiskMetrics je tzv. EWMA metóda (z angl. Exponentially Weighted Moving Average) (Morgan, 1996, str. 78–84), ktorá sa používa na reprezentáciu konečnej pamäte trhu. Presnejšie, jednodenný odhad volatility je:

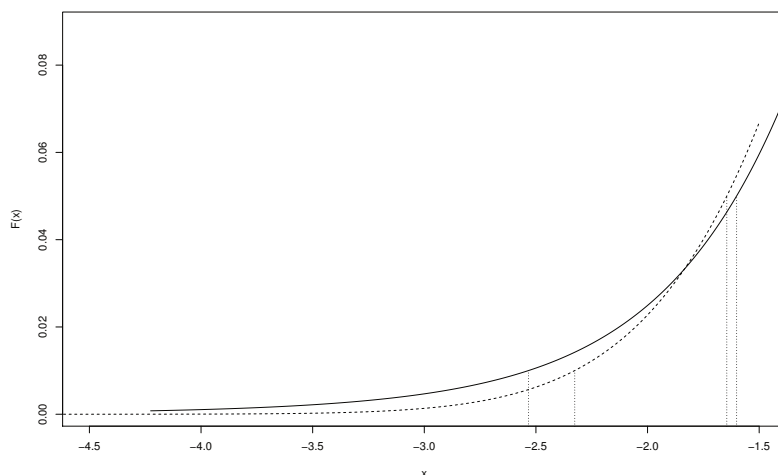
$$\sigma_{t+1|t}^2 = \frac{\sum_{\tau=0}^{\infty} \lambda^{\tau} r_{t-\tau}^2}{\sum_{\tau=0}^{\infty} \lambda^{\tau}} = (1 - \lambda) \sum_{\tau=0}^{\infty} \lambda^{\tau} r_{t-\tau}^2 = (1 - \lambda)r_t^2 + \lambda\sigma_{t|t-1}^2, \quad (2.1)$$

kde λ je parameter modelu ($0 < \lambda < 1$). Značenie $\sigma_{t+1|t}^2$ zdôrazňuje, že volatilita odhadnutá pre daný deň t je použitá pri predikcii volatility pre ďalší deň ($t + 1$). Denný VaR pre deň ($t + 1$) na konfidenčnej hladine p (napr. 95 %) sa za podmienky normality spočíta vynásobením $\sigma_{t+1|t}$ s $(1 - p)$ -tým kvantilom štandardného normálneho rozdelenia, t. j. $Var_p^{t+1} = \sigma_{t+1|t} \Phi^{-1}(1 - p)$. Hodnota parametra λ je určená vhodnou optimalizačnou procedúrou, resp. v praxi sa pre denné dáta odporúča hodnota $\lambda = 0.94$, ktorá vykazovala najlepšie výsledky pri spätnom

testovaní (Morgan, 1996, str. 97–101).

Pozrime sa bližšie na zdanlivý úspech RiskMetrics pri odhade VaR. Model stanovuje konfidenčnú hladinu 95 %. Návod, ako dostať tento 5% odhad VaR je jednoducho vynásobiť odhad podmienenej štandardnej odchýlky konštantou -1.65 (čo zhruba odpovedá 5% kvantilu štandardného normálneho rozdelenia). Dôvod, prečo takýto návod mohol uspieť v spätnom testovaní je zrejmý, keď sa pozrieme numericky na 5% kvantily niektorých leptokurtických rozdelení. Pre mnoho symetrických rozdelení s nulovou strednou hodnotou, aj napriek ťažkým chvostom, je 5% kvantil zhruba -1.65 krát štandardná odchýlka. Napríklad, pre štandarizované Študentovo t rozdelenie so 7 stupňami voľnosti (ktoré je leptokurtické a má špicatosť 5, čo je typická špicatosť výnosov vo finančných časových radoch) to je -1.60 , veľmi blízko -1.65 (viď obr. 2.1). To vysvetľuje, prečo je RiskMetrics zvyčajne prijateľným nástrojom v určovaní rizika.

Je však evidentné, že pre vyššie konfidenčné hladiny (napr. 99 %), ktoré budú v našom záujme, je efekt ťažkých chvostov omnoho silnejší, a preto VaR bude značne podhodnotený za predpokladu normality. Napríklad 1% kvantil štandarizovaného Študentovho t rozdelenia so 7 stupňami voľnosti je cca -2.54 , čo je značne väčšie (v absolútnej hodnote) ako za predpokladu normality s kvantilom -2.33 (viď obr. 2.1), kde percentil korešpondujúci -2.33 je 1.43 %. Teda dospeli sme k názoru, že uspokojivý výkon RiskMetrics pri odhadovaní VaR je zapríčinený hlavne výberom konfidenčnej hladiny 95 %. Väčšinou pri skúmaní postačiteľnosti kapitálu je avšak požadovaná hladina 99 % a viac, na ktorej RiskMetrics značne podhodnotí riziko. V ďalšom priebehu sa preto budeme sústrediť na iné metódy, ktoré budú schopné lepšie vystihnúť skutočné správanie výnosov na finančných trhoch.



Obr. 2.1: Distribučné funkcie leptokurtického štandarizovaného t -rozdelenia (plná čiara) so 7 stupňami voľnosti a štandardného normálneho rozdelenia (čiarkovaná čiara); vertikálne bodkované čiary zobrazujú 5% a 1% kvantily oboch rozdelení.

3. Teória extrémnych hodnôt

Cieľom tejto kapitoly bude sumarizovať poznatky z teórie extrémnych hodnôt (EVT z angl. Extreme Value Theory) a následne popísať jej aplikáciu na finančné dáta. Podrobne opíšeme tri prístupy bežne používané na modelovanie extrémnych udalostí. Prvý z nich je parametrický a modeluje priamo celé rozdelenie maximálnej poriadkovej štatistiky. Tento postup je označovaný ako *Blokové maximá* (BM z angl. Block Maxima). Príkladom môžu byť ročné maximálne straty v rade denných zmien ceny akcií. Hill (1975) navrhol ďalšiu metódu, označovanú ako Hillov odhad. Tá je na rozdiel od minulej metódy neparametrická a jej výhoda spočíva v tom, že v porovnaní s minulou metódou netreba nijakým spôsobom dáta pripravovať resp. ich deliť do skupín pred samotnou aplikáciou metódy. Nevýhodou je, že nemôže byť použitá vo všeobecnosti na ľubovoľné dáta. Tretia metóda skúma len dáta, ktoré prekročili určitú, vopred stanovenú medzu (POT z angl. Peaks over Threshold). Tretia spomenutá metóda sa v dnešnej dobe využíva najčastejšie, keďže využíva dáta najefektívnejšie v porovnaní s dvoma vyššie spomenutými metódami a je preto považovaná za najužitočnejšiu pre praktické aplikácie (McNeil a kol., 2005, str. 264).

3.1 Štandardné rozdelenia extrémnych hodnôt

Predpokladajme, že $\{X_i\}_{i \in \mathbb{N}}$ je postupnosť nezávislých, rovnako rozdelených (iid) náhodných veličín s neobmedzeným nosičom, ktoré reprezentujú finančné straty, resp. záporné výnosy s distribučnou funkciou F (pre jednoduchosť upustíme od indexu L). Strata bude v tejto kapitole chápaná ako kladné číslo a budeme vraviť, že *extrémna udalosť* nastala, keď strata nadobudne hodnôt v pravom chvoste distribučnej funkcie F .

Označme $M_n = \max(X_1, \dots, X_n)$ ako maximálnu stratu. V klasickej EVT je naším cieľom zistiť limitné rozdelenie M_n . Z predpokladov o $\{X_i\}_{i \in \mathbb{N}}$, plynie:

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = F^n(x). \quad (3.1)$$

V praxi, distribučná funkcia F nie je známa a tým pádom nie je známa ani distribučná funkcia M_n . Navyše $F^n(x) \rightarrow 0, \forall x \in \mathbb{R}$, pri $n \rightarrow \infty$ a teda $M_n \xrightarrow{P} \infty$. Maximum z n neobmedzených náhodných veličín rastie nad všetky medze. Toto degenerované rozdelenie nemá žiadny praktický význam a musí byť určitým spôsobom normované.

Definícia 4 (Sféra príťažlivosti). *Náhodná veličina X s distribučnou funkciou F patrí do sféry príťažlivosti (maximum domain of attraction) nedegenerovaného rozdelenia H , ak pre vhodné postupnosti $c_n > 0$ a $d_n \in \mathbb{R}, \forall n$ platí:*

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - d_n}{c_n} \leq x\right) = H(x), \quad x \in \mathbb{R}. \quad (3.2)$$

Prvky postupnosti $\{c_n\}_{n \in \mathbb{N}}$ nazývame škálovacie a prvky postupnosti $\{d_n\}_{n \in \mathbb{N}}$ lokalizačné faktory. Značíme $F \in MDA(H)$.

Ak takéto postupnosti existujú, limitné rozdelenie M_n je určené podľa nasledujúcej vety, ktorá patrí k fundamentálnym v teórii extrémnych hodnôt.

Veta 1 (Fisher a Tippett, 1928, Gnedenko, 1943). *Nech $F \in MDA(H)$ pre nejaké nedegenerované rozdelenie H , potom H patrí do jednej z troch distribučných rodín:*

$$\begin{aligned} \text{Gumbel: } \mathcal{G}(x) &= \exp\{-e^{-x}\}, \quad x \in \mathbb{R}; \\ \text{Fréchet: } \mathcal{F}_\alpha(x) &= \begin{cases} 0, & x \leq 0, \\ \exp\{-x^{-\alpha}\}, & x > 0, \alpha > 0; \end{cases} \\ \text{Weibull: } \mathcal{W}_\alpha(x) &= \begin{cases} \exp\{-(-x)^{-\alpha}\}, & x < 0, \alpha < 0, \\ 1, & x \geq 0. \end{cases} \end{aligned}$$

Dôkaz. Gnedenko (1943). □

Rozdelenia $\mathcal{G}, \mathcal{F}_\alpha, \mathcal{W}_\alpha$ sa nazývajú štandardné rozdelenia extrémnych hodnôt. Jenkinson (1955) ukázal, že je možné dostať jednoparametrickú reprezentáciu týchto troch štandardných rozdelení. Táto reprezentácia je označovaná ako zobecnené štandardné rozdelenie extrémnych hodnôt (GEVD z angl. Generalized Extreme Value Distribution).

Definícia 5. *Distribučnú funkciu zobecneného rozdelenia extrémnych hodnôt definujeme ako:*

$$\mathcal{H}_\phi(x) = \begin{cases} \exp\{-(1 + \phi x)^{-1/\phi}\}, & \phi \neq 0, \\ \exp\{-e^{-x}\}, & \phi = 0, \end{cases} \quad x \in \mathbb{R} \quad (3.3)$$

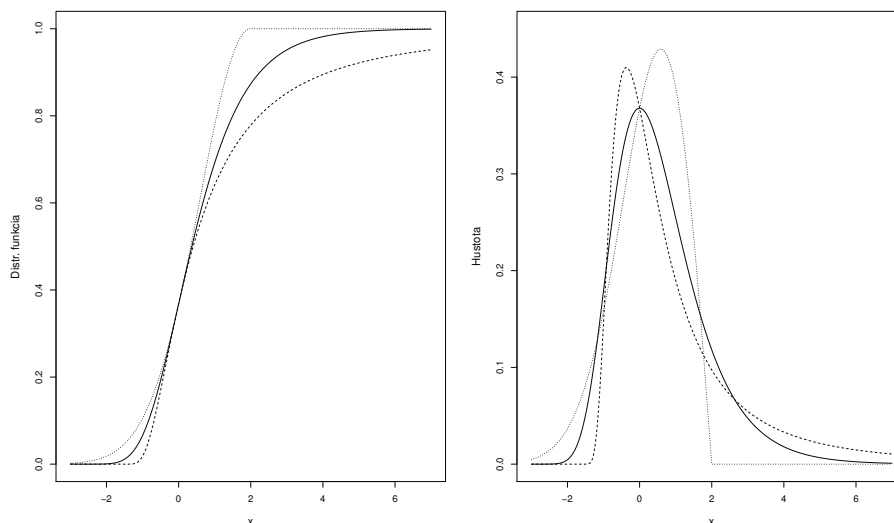
kde $\phi \in \mathbb{R}$ je paramater udávajúci tvar a $1 + \phi x > 0$.

Presnejšie pre $\phi = 0$ dostávame Gumbelovo rozdelenie. Zvyšné rozdelenia sú určené vzťahmi:

$$\left. \begin{aligned} \mathcal{H}_\phi\left(\frac{x-1}{\phi}\right) &= \mathcal{F}_{1/\phi}(x), \quad \phi > 0, \\ \mathcal{H}_\phi\left(-\frac{x+1}{\phi}\right) &= \mathcal{W}_{-1/\phi}(x), \quad \phi < 0. \end{aligned} \right\} \forall x \in \mathbb{R}$$

Úplne triedy týchto rozdelení dostaneme pridaním parametra polohy $\mu \in \mathbb{R}$ a merítka $\sigma > 0$: $\mathcal{H}_{\phi, \mu, \sigma}(x) := \mathcal{H}_\phi\left(\frac{x-\mu}{\sigma}\right)$ (viď obr. 3.1 pre príklady). Poznamenajme, že parameter μ je ľavým koncovým bodom distribučnej funkcie Fréchetovho rozdelenia a pravým koncovým bodom distribučnej funkcie Weibullovoho rozdelenia.

Poznámka. Pri konvergencii normovaných maxím je typ limitného rozdelenia určený jednoznačne až na parameter polohy a merítka, ktoré závisia na zvolených postupnostiach normujúcich konštant. Tie je vždy možné zvoliť tak, že limitou je štandardná distribučná funkcia $\mathcal{H}_{\phi, 0, 1}$ (Franke a kol., 2004, str. 313).



Obr. 3.1: Distribučné funkcie (vľavo) a hustoty (vpravo) GEV rozdelení: plná čiara odpovedá Gumbelovmu rozdeleniu s $\phi=0$; čiarkovaná čiara odpovedá Fréchetovmu rozdeleniu s $\phi=0,5$ a bodkovaná čiara odpovedá Weibullovmu rozdeleniu s $\phi=-0,5$. Všetky rozdelenia sú s parametrom polohy $\mu=0$ a parametrom merítka $\sigma=1$.

Fisherova-Tippetova-Gnedenkova veta vlastne vraví, že GEV rozdelenie je jediné možné limitné rozdelenie normalizovaného maxima M_n . Na tomto mieste sa naskytá otázka, ku ktorému z troch možných distribučných rodín M_n konverguje, ak spĺňa podmienku vety 1. Rozhodujúcim faktorom je, ako rýchlo pravdepodobnosť pre extrémne veľké hodnoty strát klesá za nejakou medzou x , keď x rastie. Keďže táto pravdepodobnosť prekročenia x hrá veľkú rolu v EVT, budeme ju bližšie skúmať.

Označme $\bar{F}(x) = P(X > x) = 1 - F(x)$ (je známa ako funkcia prežitia, z angl. survival function). U Fréchetovho rozdelenia s distribučnou funkciou \mathcal{F}_α má distribučná funkcia prežitia tvar:

$$\bar{\mathcal{F}}_\alpha = 1 - \exp\{-x^{-\alpha}\} \sim x^{-\alpha}, \quad x \rightarrow \infty.$$

V skutočnosti všetky rozdelenia patriace do sféry príťažlivosti Fréchetovho rozdelenia vykazujú rovnaké správanie. $x^\alpha \bar{F}(x)$ je takmer konštantná pre $x \rightarrow \infty$ alebo obecnjšie *pomaly sa meniaci* funkcia.

Definícia 6. (McNeil a kol., 2005, def. 7.7)

- *Kladná lebesgueovsky merateľná funkcia $s(x)$ na $(0, \infty)$ sa nazýva pomaly sa meniaci v ∞ , ak pre všetky $t > 0$:*

$$\lim_{x \rightarrow \infty} \frac{s(tx)}{s(x)} \rightarrow 1.$$

- *Kladná lebesgueovsky merateľná funkcia $r(x)$ na $(0, \infty)$ sa nazýva pravidelne*

sa meniaci v ∞ s indexom $\rho \in \mathbb{R}$, ak pre všetky $t > 0$:

$$\lim_{x \rightarrow \infty} \frac{r(tx)}{r(x)} \rightarrow t^\rho.$$

Poznámka. Pomaly sa meniace funkcie sú funkcie, ktoré sa pre veľké x menia relatívne pomaly, napr. logaritmická funkcia. Pravidelne sa meniace funkcie sú funkcie, ktoré môžu byť vyjadrené ako súčin mocnínovej funkcie s pomaly sa meniacou funkciou.

Sféru príťažlivosti Fréchetovho rozdelenia je možné charakterizovať pomocou chvostov príslušných rozdelení.

Veta 2 (Gnedenko, 1943). Pre $\alpha > 0$, pomaly sa meniacu funkciu $s(x)$, $x \in \mathbb{R}$, platí:

$$F \in MDA(\mathcal{F}_\alpha) \Leftrightarrow \bar{F}(x) = x^{-\alpha}s(x).$$

Dôkaz. Gnedenko (1943). □

Podľa vyššie uvedeného tvrdenia rozdelenia patriace do sféry príťažlivosti Fréchetovho rozdelenia sú rozdelenia s chvostmi *pravidelne sa meniacej* funkcie. Tieto rozdelenia sú v EVT najviac skúmané, keďže to sú rozdelenia s ťažkými chvostami a nekonečnými vyššími momentmi, t. j. $E(X^k) = \infty, k > \alpha$, pre nejakú nezápornú náhodnú veličinu X s $F \in MDA(\mathcal{F}_\alpha)$ (Embrechts a kol., 1997, str. 568). Tieto rozdelenia sú vhodné na modelovanie veľkých fluktuácií logaritmických výnosov,...

Príklad. Paretovo rozdelenie s distribučnou funkciou:

$$F(x) = 1 - \left(\frac{\lambda}{\lambda + x} \right)^\alpha, \quad \alpha > 0, \lambda > 0, x \geq 0,$$

patrí do sféry príťažlivosti Fréchetovho rozdelenia. To overíme jednoducho:

$$\begin{aligned} \bar{F}(x) &= \left(\frac{\lambda}{\lambda + x} \right)^\alpha = \left(\frac{\lambda x}{\lambda + x} \right)^\alpha x^{-\alpha}. \\ s(x) &= \left(\frac{\lambda x}{\lambda + x} \right)^\alpha = \left(\frac{\lambda + x}{\lambda x} \right)^{-\alpha} = \left(\frac{1}{x} + \frac{1}{\lambda} \right)^{-\alpha}. \end{aligned}$$

$s(x)$ konverguje ku konštante λ^α pre $x \rightarrow \infty$, t. z. je pomaly sa meniaci funkcia a tým sú splnené predpoklady vety 2. Ďalšími príkladmi sú t , log-gamma a Cauchyho rozdelenie.

Pre distribučnú funkciu Gumbelovho rozdelenia platí:

$$\bar{\mathcal{G}}(x) = 1 - \exp\{-e^{-x}\} \sim e^{-x}, x \rightarrow \infty.$$

Sféra príťažlivosti Gumbelovho rozdelenia pozostáva z tzv. *von Missových* distribučných funkcií a ich *chvostovo-ekvivalentných* distribučných funkcií. Definícia

týchto funkcií je pomerne náročná a technicky zložitá. McNeil a kol. (2005, def. 7.31, 7.32) vo svojej knihe podrobne opísal tieto funkcie a taktiež ukázal pomocné tvrdenie ako rozhodnúť, či daná funkcia je alebo nie je von Missova (McNeil a kol., 2005, str. 294).

$MDA(\mathcal{G})$ obsahuje množstvo rozdelení s rôznorodými chvostami od ľahkých (exponenciálne alebo Gaussovské rozdelenie), po stredne ťažké chvosty (napr. log-normálne), čo z Gumbelovej triedy, spoločne s Fréchetovou triedou, robí vhodných kandidátov na modelovanie strát vo finančnej sfére. Avšak, chvosty rozdelení v Gumbelovej sfére príťažlivosti klesajú k nule omnoho rýchlejšie ako ľubovoľná mocnina, a preto aj rýchlejšie ako rozdelenia z $MDA(\mathcal{F}_\alpha)$. Je možné dokázať, že nezáporná náhodná veličina X s distribučnou funkciou $F \in MDA(\mathcal{G})$ má konečné momenty ľubovoľného rádu, t. z. $E(X^k) < \infty$ pre $k > 0$ (Embrechts a kol., 1997, str. 148). Rozdelenia v Gumbelovej triede majú taktiež konečné aj nekonečné pravé koncové body (Embrechts a kol., 1997, kap. 3.3.3).

Sféra príťažlivosti Weibullovoho rozdelenia je tvorená distribučnými funkciami F , ktoré majú obmedzený nosič sprava, t. z. majú konečný pravý koniec, $x_F = \sup\{x \in \mathbb{R} : F(x) < 1\} < \infty$.

Veta 3 (Gnedenko, 1943). *Pre $\alpha < 0$, pomaly sa meniacu funkciu $s(x)$, $x \in \mathbb{R}$, platí:*

$$F \in MDA(\mathcal{W}_\alpha) \Leftrightarrow x_F < \infty \quad a \quad \bar{F}(x_F - x^{-1}) = x^\alpha s(x).$$

Dôkaz. Gnedenko (1943). □

Fakt, že $x_F < \infty$ robí z tejto triedy najnevhodnejšiu voľbu na modelovanie extrémnych udalostí vo finančnej sfére a preto sa nebudeme týmito rozdeleniami ďalej zaoberať. Príkladom rozdelenia, ktoré patrí do tejto sféry príťažlivosti je napríklad Beta rozdelenie definované na intervale (0,1).

3.2 Odhady parametrov

Na základe teoretických výsledkov prezentovaných v predošlých sekciách sme pripravení aplikovať BM metódu. Predpokladajme, že máme k dispozícii dáta $\{X_1, \dots, X_n\}$, ktoré sú nezávislé a rovnako rozdelené s distribučnou funkciou $F \in MDA(\mathcal{H}_\phi)$. Z predošlej sekcie a vety 1 vieme, že skutočné rozdelenie maxima M_n môže byť pre veľké n aproximované GEV rozdelením. Reálne nepoznáme skutočné rozdelenie strát a nemôžeme teda na základe toho určiť normujúce konštanty c_n a d_n . Preto použijeme trojparametrické vyjadrenie $\mathcal{H}_{\phi, \mu, \sigma}$, kde nahradíme normujúce konštanty c_n a d_n parametrami merítka a polohy, $\sigma > 0$ resp. $\mu \in \mathbb{R}$. Tieto parametre môžu byť aproximované parametrickými alebo neparametrickými metódami. Uvedieme a bližšie popíšeme vybrané metódy, ktoré použil vo svojej knihe Tsay (2010, kap. 7.5.2.2).

3.2.1 Parametrický odhad: Blokové maximá

Pre daný náhodný výber $\{X_1, \dots, X_n\}$ existuje len jedno maximum a nie je možné odhadnúť tri parametre na základe jedného extrémneho pozorovania. Jedna z možností používaných v literatúre je rozdeliť dáta do m blokov o veľkosti g a aplikovať EVT v každom bloku. Pre jednoduchosť predpokladajme, že počet všetkých pozorovaní $n = mg$. Označme blokové maximum j -tého bloku ako M_{gj} . Toto delenie samozrejme vyžaduje, aby boli dáta rozdelené nejakou prirodzenou cestou. Napríklad ak máme k dispozícii denné straty, tak ich môžeme rozdeliť do mesačných, štvrtročných alebo ročných blokov. Ak g je dostatočne veľké, $(M_{gj} - d_g)/c_g$ by malo konvergovať ku GEV rozdeleniu a postupnosť $\{M_{gj}\}_{j=1, \dots, m}$ môže byť považovaná za náhodný výber z m pozorovaní s GEV rozdelením.

Poznámka. Keď celkový počet pozorovaní n nie je násobkom g , môžeme vziať posledný blok o menšej veľkosti alebo môžeme ignorovať prvých pár pozorovaní, aby mal každý z m blokov g pozorovaní.

Trojicu uvažovaných parametrov odhadneme metódou maximálnej vierohodnosti. Pri výpočte bude potrebná hustota GEV rozdelenia, ktorá sa dostane jednoduchým derivovaním distribučnej funkcie $\mathcal{H}_{\phi, \mu, \sigma}$ uvedenej v (3.3):

$$h_{\phi, \mu, \sigma}(x) = \begin{cases} \frac{1}{\sigma} (1 + \phi \frac{x-\mu}{\sigma})^{-1/\phi-1} \exp\{-(1 + \phi \frac{x-\mu}{\sigma})^{-1/\phi}\}, & \phi \neq 0, \\ \frac{1}{\sigma} \exp\{-\frac{x-\mu}{\sigma}\} \exp\{-e^{-(x-\mu)/\sigma}\}, & \phi = 0. \end{cases} \quad x \in \mathbb{R}$$

Predpokladajme dostatočne veľkú veľkosť g tak, že maximálne pozorovania každého z m blokov sú nezávislé bez ohľadu na to, či počiatkové dáta boli alebo neboli nezávislé. Vierohodnostná funkcia náhodného výberu M_{g1}, \dots, M_{gm} má tvar:

$$\mathcal{L}(\phi, \mu, \sigma; M_{g1}, \dots, M_{gm}) = \prod_{i=1}^m h_{\phi, \mu, \sigma}(M_{gi}).$$

Zlogaritmovaním dostávame logaritmickú vierohodnosť:

$$\begin{aligned} \ell(\phi, \mu, \sigma; M_{g1}, \dots, M_{gm}) &= \sum_{i=1}^m \ln h_{\phi, \mu, \sigma}(M_{gi}) \\ &= -m \ln \sigma - \left(1 + \frac{1}{\phi}\right) \sum_{i=1}^m \ln \left(1 + \phi \frac{M_{gi} - \mu}{\sigma}\right) \\ &\quad - \sum_{i=1}^m \left(1 + \phi \frac{M_{gi} - \mu}{\sigma}\right)^{-1/\phi}. \end{aligned} \quad (3.4)$$

Maximálne vierohodným odhadom ϕ , μ a σ je:

$$(\hat{\phi}, \hat{\mu}, \hat{\sigma}) = \arg \max_{\phi, \mu, \sigma} \ell(\phi, \mu, \sigma; M_{g1}, \dots, M_{gm}),$$

za podmienok $\sigma > 0$ a $1 + \phi(M_{gi} - \mu)/\sigma > 0, \forall i$. K odhadnutiu týchto parametrov je potrebná nelineárna metóda. Tieto odhady sú za určitých podmienok nestranné a asymptoticky normálne (McNeil a kol., 2005, str. 272).

Uviedli sme, ako odhadneme parametre GEV rozdelenia, no doteraz sme neuviedli ako určiť počet (m) resp. veľkosť (g) blokov. Veľká hodnota g totižto

vedie k presnejším odhadom rozdelenia blokových maxim GEV rozdelením a malej chybe v odhadovaných parametroch, zatiaľčo veľká hodnota m nám dáva viac BM dát pre odhad metódou maximálnej vierohodnosti a vedie k malému rozptylu v odhadovaných parametroch (McNeil a kol., 2005, str. 272). Nie je žiadna zaužívaná metóda, a preto je potrebný akýsi kompromis.

3.2.2 Výpočet VaR pomocou BM metódy

V predošlej časti sme ukázali, ako môžeme parametrickou metódou odhadnúť parametre GEV rozdelenia, ktorým vyrovnávame historické straty. V tejto kapitole ukážeme, ako pomocou týchto odhadov spočítame ukazovateľa hodnoty v riziku, na ktorý sa táto práca sústreďí. Diskusiu rozdelíme do dvoch častí.

Prvá časť

Predpokladáme, že máme k dispozícii n pozorovaní. Týchto n pozorovaní rozdelíme do m blokov o veľkosti g . Ak má posledný blok menej ako g pozorovaní, budeme ignorovať prvých pár pozorovaní, aby mal každý blok g pozorovaní. Metódou blokových maxim získame odhady $\hat{\phi}$, $\hat{\mu}$, $\hat{\sigma}$ pre blokové maximá $\{M_{gj}\}_{j=1,\dots,m}$. Nech p označuje malú pravdepodobnosť (napr. 5%), ktorá indikuje potenciálnu stratu a q_{1-p} je $(1-p)$ -tý kvantil blokových maxim, ktoré majú GEV rozdelenie s odhadnutými parametrami. Po dosadení do (3.3), kde uvažujeme aj parameter polohy a merítka máme:

$$(1-p) = \begin{cases} \exp\{-[1 + \hat{\phi} \frac{(q_{1-p} - \hat{\mu})}{\hat{\sigma}}]^{-1/\hat{\phi}}\}, & \hat{\phi} \neq 0, \\ \exp\{-\exp\{-\frac{q_{1-p} - \hat{\mu}}{\hat{\sigma}}\}\}, & \hat{\phi} = 0, \end{cases} \quad (3.5)$$

kde $1 + \hat{\phi} \frac{q_{1-p} - \hat{\mu}}{\hat{\sigma}} > 0$ pre $\hat{\phi} \neq 0$. Kvantil q_{1-p} dostaneme aplikovaním logaritmu na (3.5) a následným vyjadrením q_{1-p} z tejto rovnice:

$$q_{1-p} = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\phi}} \{1 - [-\ln(1-p)]^{-\hat{\phi}}\}, & \hat{\phi} \neq 0, \\ \hat{\mu} - \hat{\sigma} \ln[-\ln(1-p)], & \hat{\phi} = 0. \end{cases} \quad (3.6)$$

Druhá časť

Pre danú pravdepodobnosť p je kvantil z rovnice (3.6) VaR založený na EVT pre blokové maximá. Ďalším krokom je prepojenie medzi blokovými maximami a originálnymi pozorovaniami $\{X_i\}_{i=1}^n$. Blokové maximá pre dostatočne veľké dĺžky blokov sú väčšinou nekorelované (Tsay, 2010, kap. 7.6), a preto môžeme použiť vzťah (3.1) zo začiatku kapitoly:

$$(1-p) = P(M_{gj} \leq q_{1-p}) = [F(q_{1-p})]^g,$$

kde F je distribučná funkcia strát a $j = 1, \dots, m$. Tento vzťah nám umožňuje získať VaR pre originálne dáta $\{X_i\}_{i=1}^n$. Presnejšie, pre malú pravdepodobnosť p je $(1-p)$ -tý kvantil rozdelenia strát, t. j. VaR_{1-p} rovný:

$$VaR_{1-p} = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\phi}} \{1 - [-g \ln(1-p)]^{-\hat{\phi}}\}, & \hat{\phi} \neq 0, \\ \hat{\mu} - \hat{\sigma} \ln[-g \ln(1-p)], & \hat{\phi} = 0, \end{cases} \quad (3.7)$$

pričom výsledný vzorec sme získali opakovaním postupu z prvej časti, kde do rovnice (3.5) bolo miesto $(1-p)$ počítané s $(1-p)^g$.

3.2.3 Neparametrický odhad: Hillova metóda

Parameter tvaru ϕ môže byť alternatívne odhadnutý užitím neparametrických metód. Odhad sa aplikuje priamo na originálne dáta $\{X_i\}_{i=1,\dots,n}$, teda v porovnaní s predošlou metódou, nie je potrebné delenie dát do skupín.

V tejto časti predpokladajme, že dáta $\{X_1, \dots, X_n\}$ sú nezávislé a rovnako rozdelené náhodné veličiny s distribučnou funkciou $F \in MDA(\mathcal{F}_\alpha)$ pre nejaké $\alpha > 0$. Inak povedané, uvažujeme len rozdelenia patriace do Fréchetovej sféry príťažlivosti. Z vety 2 vieme, že je to v prípade $\bar{F}(x) = x^{-\alpha}s(x)$, kde s je pomaly sa meniaca funkcia. To, ako sa mení funkcia $\bar{F}(x)$ pre narastajúce x je určené prevažne koeficientom $\alpha = 1/\phi$. Hillov odhad (Hill, 1975) sa preto sústreďuje na odhad tohto koeficientu.

Definícia 7. (*Hillov odhad*) (Franke a kol., 2004, def. 17.11) Nech $X^{(1)} \geq X^{(2)} \geq \dots \geq X^{(n)}$ sú poriadkové štatistiky v klesajúcom poradí. Hillov odhad α_k parametra α je pre vhodné $k = k(n) \in \{1, \dots, n\}$:

$$\alpha_k = \left(\frac{1}{k} \sum_{i=1}^k \ln X^{(i)} - \ln X^{(k)} \right)^{-1}.$$

Tvar Hillovho odhadu je zrejmý z nasledujúceho jednoduchého špeciálneho prípadu (Franke a kol., 2004, str. 329). Vo všeobecnosti pre distribučnú funkciu $\bar{F}(x)$ platí: $\bar{F}(x) = s(x)x^{-\alpha}$. Predpokladajme ale, že pre fixné $c > 0$ je $s(x) = c^\alpha$ konštantné. Položme $V_i = \log(X_i/c)$, potom platí:

$$P(V_i > v) = P(X_i > ce^v) = \bar{F}(ce^v) = \frac{c^\alpha}{(ce^v)^\alpha} = e^{-\alpha v}, v \geq 0.$$

V_1, \dots, V_n sú teda nezávislé náhodné veličiny s exponenciálnym rozdelením s parametrom α . U veličín s exponenciálnym rozdelením platí: $1/\alpha = \mathbb{E}V_i$ a pre maximálne vierohodný odhad parametra α platí:

$$\hat{\alpha} = \frac{1}{\frac{1}{n} \sum_{i=1}^n V_i} = \left(\frac{1}{n} \sum_{i=1}^n \ln(X_i/c) \right)^{-1} = \left(\frac{1}{n} \sum_{i=1}^n \ln(X^{(i)}) - \ln(c) \right)^{-1},$$

kde v poslednej rovnosti bolo len zamenené poradie sčítancov v sume. Ako je možné vidieť, $\hat{\alpha}$ je podobný Hillovému odhadu. Vo všeobecnosti samozrejme platí len: $\bar{F}(x) \sim c^\alpha/x^\alpha$ pre dostatočne veľké x . Argument, ktorý bol použitý v tomto špeciálnom prípade je podobný aj v prípade najväčších pozorovaní $X^{(1)} \geq X^{(2)} \geq \dots \geq X^{(k)} \geq \eta$, kde η je nejaká vysoká medza. Jedine k najväčších poriadkových štatistík, ktoré sú porovnávané s k -tou poriadkovou štatistikou, vstúpi do definície Hillovho odhadu.

Hillov odhad je konzistentný, konverguje v pravdepodobnosti ku koeficientu α pre $n, k \rightarrow \infty$ také, že $k/n \rightarrow 0$. Za dodatočných predpokladov je možné dokázať, že $\sqrt{k}(\alpha_k - \alpha) \xrightarrow{D} N(0, \alpha^2)$, t. z., α_k má asymptoticky normálne rozdelenie (Franke a kol., 2004, str. 329).

Na prvý pohľad by sa mohlo zdať, že Hillova metóda je efektívnejšia ako metóda BM, pri ktorej musíme voliť veľkosť a počet blokov. Pri Hillovej metóde

avšak musíme zvoliť medzu $\eta = X^{(k)}$, pretože len pozorovania nad touto medzou vstupujú do definície Hillovho odhadu. Opäť čelíme dileme:

- Pre k príliš malé len veľmi málo pozorovaní ovplyvní α_k a rozptyl odhadu, ktorý je asymptoticky α^2/k je príliš veľký.
- Naopak, pre k príliš veľké, predpoklad, že $s(x)$ je aproximovaná konštantou pre všetky $x \geq X^{(k)}$ vo všeobecnosti nie je dodržaný a chyba $\mathbb{E}\alpha_k - \alpha$ je príliš veľká.

Ak dáta pochádzajú z distribúcie, ktorej chvost je veľmi blízky mocninatej funkcii, Hillov odhad je často dobrým odhadom α . V praxi sa na určenie konštanty k väčšinou konštruujú Hillove odhady pre rôzne hodnoty k . Tým sa dostávajú rôzne hodnoty α_k , ktoré sa vykresľujú do grafu $\{(k, \alpha_k) : k = 2, \dots, n\}$, v ktorom sa snažíme nájsť stabilnú oblasť, kde sú odhady skonštruované z rôzneho počtu poriadkových štatistík podobné.

3.2.4 Výpočet VaR pomocou Hillovej metódy

Vzhľadom k tomu, že uvažujeme len rozdelenia z Fréchetovej sféry príťažlivosti, po odhadnutí koeficientu α dostávame priamy odhad funkcie $\bar{F}(x)$. Pre rozdelenia z Fréchetovej sféry príťažlivosti totižto platí: $\bar{F}(x) = x^{-\alpha}s(x)$, pre nejakú pomaly sa meniacu funkciu s . Pre veľké hodnoty $x \geq X^{(k)}$ taktiež platí (Franke a kol., 2004, str. 330):

$$\frac{\bar{F}(x)}{\bar{F}(X^{(k)})} = \frac{s(x)}{s(X^{(k)})} \left(\frac{X^{(k)}}{x}\right)^\alpha \sim \left(\frac{X^{(k)}}{x}\right)^\alpha. \quad (3.8)$$

K odhadu $\bar{F}(x)$ nám ostáva odhadnúť $\bar{F}(X^{(k)})$. Keďže presne (k/n) -tina dát je väčšia alebo rovná ako poriadková štatistika $X^{(k)}$, zrejším odhadom $\bar{F}(X^{(k)})$ je k/n (McNeil a kol., 2005, str. 289). Vložením tohto odhadu do (3.8) dostaneme Hillov odhad pre $\bar{F}(x)$:

$$\hat{\bar{F}}(x) = \frac{k}{n} \left(\frac{X^{(k)}}{x}\right)^{\alpha_k}.$$

Z tohto vyjadrenia už jednoduchými úpravami dostaneme Hillov odhad VaR_{1-p} ako:

$$VaR_{1-p} = X^{(k)} \left(\frac{np}{k}\right)^{-1/\alpha_k}. \quad (3.9)$$

3.3 Prístup pomocou presahov cez medzu

Postupy pre výpočet VaR zmienené v predošlých kapitolách trpia určitými nedostatkami. Prvým je voľba dĺžky bloku g v BM metóde. Druhým nedostatkom je, že Hillova metóda môže byť aplikovaná len na dáta pochádzajúce z Fréchetovej sféry príťažlivosti. Aby sme sa vyhlili týmto problémom, použijeme modernú metódu založenú na teórii extrémnych hodnôt, ktorá sa namiesto extrémov sústredí na presahy dát cez vopred stanovenú medzu η . Tento prístup je v angličtine označovaný ako Peaks over Threshold (POT).

Z predchádzajúceho opisu vyplýva, že nový prístup nevyžaduje voľbu veľkosti bloku g , ale vyžaduje voľbu medze η . Rôzne voľby medze vedú k rôznym odhadom parametra tvaru ϕ . Inak povedané, voľba η je ako štatistický, tak aj finančný problém a hodnota tohto parametra nemôže byť určená čisto na základe štatistickej teórie. Napríklad rôzne finančné inštitúcie alebo investori majú rôznu averziu k riziku, . . . Voľba η tiež závisí na pozorovaných stratách. Pre stabilné rady môže byť niekoľkokrát menšia v porovnaní s radmi, kde straty majú veľkú volatilitu. Doterajšie štúdie ukazujú, že medza η by mala byť vyberaná tak, aby bol počet presahov dostatočne veľký, t. j. približne 5% výberu (Tsay, 2010, str. 284).

Opäť predpokladajme, že $\{X_i\}_{i \in \mathbb{N}}$ je postupnosť nezávislých, rovnako rozdelených náhodných veličín s neobmedzeným nosičom, ktoré reprezentujú finančné straty s distribučnou funkciou F . POT metóda sa namiesto skúmania marginálneho rozdelenia, t. j. maxima náhodného výberu, sústreďuje na podmienené rozdelenie výšky presahu, za podmienky, že k danému presahu došlo.

Definícia 8. (McNeil a kol., 2005, def. 7.17) Nech X je náhodná veličina s distribučnou funkciou F . Podmienené rozdelenie $F_\eta(x) := P(X \leq x + \eta | X > \eta) = \frac{F(x+\eta) - F(\eta)}{1 - F(\eta)}$, $x \geq 0$, budeme označovať ako distribučnú funkciu presahu cez medzu η ($\eta \in \mathbb{R}$).

Definícia 9. (McNeil a kol., 2005, def. 7.18.) Funkciu stredného presahu náhodnej veličiny X s konečnou strednou hodnotou definujeme ako:

$$e(\eta) = \mathbb{E}(X - \eta | X > \eta).$$

Nech distribučná funkcia strát F má tvar úplného GEV rozdelenia z definície (3.3), t. j. $F(x) = \mathcal{H}_{\phi, \mu, \sigma}(x)$. S využitím aproximácie $e^{-y} \sim 1 - y$ a po prevedení drobných algebraických úprav dostávame približný vzťah:

$$\begin{aligned} F_\eta(x) &= \frac{\mathcal{H}_{\phi, \mu, \sigma}(x + \eta) - \mathcal{H}_{\phi, \mu, \sigma}(\eta)}{1 - \mathcal{H}_{\phi, \mu, \sigma}(\eta)} \\ &= \frac{\exp\{-[1 + \phi \frac{x+\eta-\mu}{\sigma}]^{-1/\phi}\} - \exp\{-[1 + \phi \frac{\eta-\mu}{\sigma}]^{-1/\phi}\}}{1 - \exp\{-[1 + \phi \frac{\eta-\mu}{\sigma}]^{-1/\phi}\}} \\ &\sim 1 - \left(1 + \frac{\phi x}{\sigma + \phi(\eta - \mu)}\right)^{-1/\phi}, \end{aligned} \quad (3.10)$$

kde $x > 0$ a $1 + \phi(\eta - \mu)/\sigma > 0$. Ako budeme vidieť neskôr, táto aproximácia tvorí explicitné prepojenie medzi novým prístupom a tradičnou metódou extrémnych hodnôt z kapitoly 3.2. Prípad, kedy $\phi = 0$ je chápaný ako limita $\phi \rightarrow 0$:

$$F_\eta(x) \sim 1 - \exp\{-x/\sigma\}.$$

Definícia 10. (Tsay, 2010, kap. 7.7.1) Rozdelenie s distribučnou funkciou:

$$\mathcal{G}_{\phi, \psi(\eta)}(x) = \begin{cases} 1 - \left(1 + \frac{\phi x}{\psi(\eta)}\right)^{-1/\phi}, & \phi \neq 0, \\ 1 - \exp\{-x/\psi(\eta)\}, & \phi = 0, \end{cases} \quad (3.11)$$

kde $\psi(\eta) > 0$, $x \geq 0$ pri $\phi \geq 0$ a $0 \leq x \leq -\psi(\eta)/\phi$ pri $\phi < 0$ je nazývané zobecnené Paretovo rozdelenie (GPD z angl. Generalized Pareto Distribution). Parameter $\psi(\eta)$ budeme nazývať škálovací parameter.

Výsledok odvodenia (3.10) nám teda ukazuje, že rozdelenie $F_\eta(x)$ môže byť vhodne aproximované zobecneným Paretovým rozdelením s parametrami ϕ a $\psi(\eta) = \sigma + \phi(\eta - \mu)$. Formálne môže byť dokázané, že GPD je v skutočnosti prirodzenou limitnou distribučnou funkciou presahu pre viacero distribučných funkcií strát. Výsledok nasledujúcej vety môže byť vnímaný ako charakterizačná veta pre sféru príťažlivosti GEV rozdelenia. Na začiatku tejto kapitoly sme sa pozerali na charakterizácie separátne pre každý z troch možných prípadov pre $\phi > 0$, $\phi = 0$, $\phi < 0$. Nasledujúci výsledok nám dáva globálnu charakterizáciu $MDA(\mathcal{H}_\phi)$ pre všetky ϕ , ktorá sa týka limitného správania distribučnej funkcie presahu cez medzu η .

Veta 4. (*Balkema a de Haan, 1974*) *Existuje (kladná) funkcia $\psi(\eta)$ taká, že:*

$$\lim_{\eta \rightarrow \infty} \sup_{x \geq 0} |F_\eta(x) - \mathcal{G}_{\phi, \psi(\eta)}(x)| = 0,$$

práve vtedy, keď $F \in MDA(\mathcal{H}_\phi)$, $\phi \in \mathbb{R}$.

Inak, rozdelenia, pre ktoré normalizované maximum konverguje ku GEV rozdeleniu, stanovujú množinu rozdelení, pre ktoré s narastajúcou medzou distribučná funkcia presahu cez medzu η konverguje ku GPD. Navyše, parameter tvaru limitného GPD je rovnaký ako parameter tvaru limitného GEV rozdelenia pre maximum.

3.3.1 Metóda odhadu

Predpokladajme, že pre nejakú veľkú medzu η je $F_\eta(x) = \mathcal{G}_{\phi, \psi(\eta)}(x)$ pre nejaké $x \geq 0$, $\phi \in \mathbb{R}$ a $\psi(\eta) > 0$. Tento predpoklad je čistá idealizácia, keďže v praxi bude rozdelenie presahu nie *presne* GPD, ale tento predpoklad nám umožní previesť niekoľko užitočných odvodení v ďalších sekciách. Máme k dispozícii výšky strát $\{X_1, \dots, X_n\}$ s distribučnou funkciou F a náhodné číslo N_η , ktoré značí počet strát, ktoré prekročia medzu η . Pre prehľadnosť si preznačme dáta prekračujúce medzu η ako $\tilde{X}_1, \dots, \tilde{X}_{N_\eta}$ a pre všetky spočítajme výšku presahu $Y_j = \tilde{X}_j - \eta$. Radi by sme odhadli parametre GPD za pomoci N_η presahov. Existuje množstvo spôsobov, no najpoužívanejším a ľahko implementovateľným je odhad metódou maximálnej vierohodnosti. Na chvíľu predpokladajme, že presahujúce dáta môžeme pokladať za realizácie nezávislých náhodných veličín. Nech $g_{\phi, \psi(\eta)}(x)$ značí hustotu GPD, potom, logaritmická vierohodnostná funkcia je tvaru:

$$\begin{aligned} \ell(\phi, \psi(\eta); Y_1, \dots, Y_{N_\eta}) &= \sum_{i=1}^{N_\eta} \ln g_{\phi, \psi(\eta)}(Y_i) \\ &= -N_\eta \ln(\psi(\eta)) - \left(1 + \frac{1}{\phi}\right) \sum_{i=1}^{N_\eta} \ln \left(1 + \phi \frac{Y_i}{\psi(\eta)}\right), \end{aligned}$$

čo je maximalizované za podmienok: $\psi(\eta) > 0$ a $1 + \phi Y_i / \psi(\eta) > 0$, $\forall i$. Vyriešením tohto maximalizačného problému dostávame $\mathcal{G}_{\hat{\phi}, \hat{\psi}(\eta)}$ ako odhad F_η .

3.3.2 Presahy cez vyššiu medzu

Dôležitá vlastnosť GPD je, že rozdelenie presahu cez medzu η pri zmene medze zostáva v tvare (3.11) s rovnakým parametrom tvaru ϕ a škálovacím parametrom, ktorý rastie lineárne s medzou η .

Veta 5. (McNeil a kol., 2005, Lemma 7.22) Nech pre nejakú vysokú medzu η_0 je $F_{\eta_0}(x) = \mathcal{G}_{\phi, \psi(\eta_0)}(x)$, $x \geq 0$. Potom, pre ľubovoľnú medzu $\eta \geq \eta_0$ platí: $F_\eta(x) = \mathcal{G}_{\phi, \psi(\eta_0) + \phi(\eta - \eta_0)}(x)$.

Dôkaz. McNeil a kol. (2005, strana 279). □

Teda, rozdelenie presahu pri prechode na vyššiu medzu η ostáva GPD s rovnakým parametrom tvaru ϕ , jedine škálovací parameter rastie lineárne s medzou η . Pri predpoklade $\phi < 1$ je podľa vlastností GPD stredná funkcia presahu z definície 9:

$$e(\eta) = \frac{\psi(\eta_0) + \phi(\eta - \eta_0)}{1 - \phi} = \frac{\psi(\eta_0) - \phi\eta_0}{1 - \phi} + \frac{\phi\eta}{1 - \phi}, \quad (3.12)$$

kde $\eta_0 \leq \eta < \infty$ pri $0 \leq \phi < 1$ a $\eta_0 \leq \eta \leq \eta_0 - \psi(\eta_0)/\phi$ pri $\phi < 0$. Linearita funkcie stredného presahu z (3.12) v η tvorí základ pre nasledujúcu jednoduchú grafickú metódu pre výber vhodnej medze.

Pre kladné straty X_1, \dots, X_n definujeme:

$$e_n(\eta) = \frac{\sum_{i=1}^n (X_i - \eta) \mathbb{1}_{[X_i > \eta]}}{\sum_{i=1}^n \mathbb{1}_{[X_i > \eta]}}, \quad (3.13)$$

ako empirický odhad funkcie stredného presahu z definície 9. Na štúdium tejto funkcie sa väčšinou konštruuje graf $\{(X_n^{(i)}, e_n(X_n^{(i)})) : 2 \leq i \leq n\}$, kde $X_n^{(i)}$ značí ako v predošlých sekciách i -tú poriadkovú štatistiku. Ak dáta presahujúce medzu η_0 majú GDP, tak (3.12) nám ukazuje, že pre väčšie hodnoty η by grafom mala byť priamka. Lineárny rastúci trend indikuje GPD model s kladným parametrom tvaru ϕ , naopak lineárny klesajúci trend indikuje GPD model so záporným parametrom tvaru ϕ a konečne, konštatný trend indikuje GPD model s parametrom tvaru ϕ takmer nulovým, alebo inak, exponenciálnym rozdelením presahu cez medzu η . Pri praktickej aplikácii si vykreslíme vyššie spomenutý graf, v ktorom pozorujeme bod, od ktorého je graf funkcie lineárny. Tento bod môžeme podľa práve vykonanej diskusie pokladať za vhodný výber medze.

Poznámka. V praxi avšak tieto idealizované prípady nenastávajú a dokonca aj pri dátach, ktoré sú generované z GPD dostaneme zriedkakedy perfektne lineárny graf, väčšinou k jeho pravému koncu, kde priemerujeme malý počet veľkých presahov. V skutočnosti sa často vynecháva posledných pár hodnôt, aby graf nebol príliš deformovaný (McNeil a kol., 2005, str. 280).

3.3.3 Výpočet VaR pomocou POT metódy

Rovnako ako na začiatku kapitoly 3.3.1 predpokladajme, že pre nejakú veľkú medzu η máme $F_\eta(x) = \mathcal{G}_{\phi, \psi(\eta)}(x)$ pre nejaké $x \geq 0$, $\phi \in \mathbb{R}$ a $\psi(\eta) > 0$. Pre $x \geq \eta$ potom:

$$\begin{aligned}\bar{F}(x) &= P(X > \eta)P(X > x | X > \eta) \\ &= \bar{F}(\eta)P(X - \eta > x - \eta | X > \eta) \\ &= \bar{F}(\eta)\bar{F}_\eta(x - \eta) \\ &= \bar{F}(\eta) \left(1 + \phi \frac{x - \eta}{\psi(\eta)}\right)^{-1/\phi}.\end{aligned}$$

Pomocou tohto odvodenia môžeme získať vzorec pre výpočet VaR. K tomu samozrejme potrebujeme odhady koeficientov ϕ , $\psi(\eta)$ a $\bar{F}(\eta)$. Prvé dva spomenuté odhadneme metódou z kapitoly 3.3.1. Na odhad tretieho použijeme jednoduchý empirický odhad N_η/n , kde N_η je počet dát presahujúcich cez medzu η . Pre nejakú malú pravdepodobnosť p je potom:

$$VaR_{1-p} = \eta - \frac{\psi(\eta)}{\phi} \left(1 - \left(\frac{np}{N_\eta}\right)^{-\phi}\right). \quad (3.14)$$

3.3.4 VaR pre dlhší časový horizont

Vzorce (3.7), (3.9) a (3.14) udávajú výpočet VaR pre jednu časovú jednotku s ktorou pracujeme. Vo väčšine prípadov potrebujeme odhad VaR pre viacero období. Uvažujme d období a označme ${}_dVaR_{1-p}$ ako hodnotu v riziku pre d časových jednotiek na konfidenčnej hladine $(1 - p)$. Danielsson a De Vries (1997) odvodili nasledujúce pravidlo, ktoré je označované ako „ α root of time rule“:

$${}_dVaR_{1-p} = d^{1/\alpha} VaR_{1-p} = d^\phi VaR_{1-p},$$

kde ϕ je parameter udávajúci tvar.

4. Zovšeobecnenie na finančné časové rady

Štandardná teória predošlých sekcií predpokladá nezávislú, rovnako rozdelenú postupnosť strát $\{X_i\}_{i \in \mathbb{N}}$, čo je v prípade finančných časových radov, ktoré sú skúmané v tejto práci a na ktorých bude aplikovaná reálna analýza, často ne-reálny predpoklad. V tejto kapitole preto poskytneme 3 alternatívne cesty ako sa vysporiadať s týmto problémom. V sekcii 4.1 zovšeobecníme teóriu extrémnych hodnôt pre striktné stacionárne postupnosti, čo je slabší predpoklad v porovnaní s predpokladom iid strát. V sekciiach 4.2 a 4.3 ukážeme ako finančné dáta modifikovať tak, aby na nich mohla byť EVT aplikovateľná. Najprv zakomponujeme vysvetľujúce premenné, ktoré máme v určitých situáciách k dispozícii. V ďalšej časti dáta pred aplikáciou EVT „očistíme“ (z angl. pre-whitening), na očistené dáta, o ktorých budeme predpokladať že sú iid, aplikujeme EVT a ukážeme postup ako dostať VaR pre pôvodné straty.

4.1 Extremálny index

V tejto sekcii poskytneme podmienky pre striktné stacionárny proces $\{X_i\}_{i \in \mathbb{N}}$ (pre definíciu viď (Cipra, 2008, str. 328)) s distribučnou funkciou F , ktoré zaisťujú, že maximum náhodného výberu $M_n = \max(X_1, \dots, X_n)$ a odpovedajúce maximum jeho *pridruženého procesu*, t. z. $\tilde{M}_n = \max(\tilde{X}_1, \dots, \tilde{X}_n)$ iid postupnosti $\{\tilde{X}_i\}_{i \in \mathbb{N}}$ s rovnakou distribučnou funkciou F vykazuje rovnaké limitné správanie, t. z. má rovnaký typ limitného rozdelenia.

Štandardná veličina opisujúca zhlukovanie extrémnych hodnôt procesu je tzv. *extremálny index* procesu (naďalej označovaný ako θ). Extremálny index procesu nám umožňuje charakterizovať vzťah medzi závislostnou štruktúrou dát a ich extrémnym správaním. Definícia extrémneho indexu je príliš technická, uvedieme len intuitívne odvodenie (pre formálnu definíciu viď (Embrechts a kol., 1997, str.416)). Leadbetter (1983) ukázal, že pre ľubovoľnú postupnosť reálnych čísel $\{u_n\}_{n \in \mathbb{N}}$ platí:

$$P(M_n \leq u_n) \sim P^\theta(\tilde{M}_n \leq u_n) = F^{n\theta}(u_n), \quad (4.1)$$

pre nejaké $\theta \in (0,1]$ a dostatočne veľké n . Limitne sa teda maximum z n pozorovaní stacionárnej postupnosti s extrémnym indexom θ správa ako maximum z $n\theta$ pozorovaní z príslušnej pridruženej iid postupnosti. Ako dôsledok máme nasledujúce tvrdenie pre stacionárne postupnosti.

Veta 6. (McNeil a kol., 2005, str. 270) *Nech $\{X_i\}_{i \in \mathbb{N}}$ je striktné stacionárna postupnosť s extrémnym indexom $\theta \in (0,1]$, potom:*

$$\lim_{n \rightarrow \infty} P\left(\frac{\tilde{M}_n - d_n}{c_n} \leq x\right) = H(x), \quad \forall x \in \mathbb{R},$$

pre nedegenerované rozdelenie H práve keď,

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - d_n}{c_n} \leq x\right) = H^\theta(x), \quad \forall x \in \mathbb{R}.$$

Pre procesy s extrémálnym indexom, normalizované blokové maximum konverguje v distribúcii k nedegenerovanému rozdeleniu H , za predpokladu, že $F \in MDA(\mathcal{H}_\phi)$ pre nejaké ϕ . Navyše, je možné overiť, že $\mathcal{H}_\phi^\theta(x)$ je rozdelenie rovnakého typu ako $\mathcal{H}_\phi(x)$ a limitné rozdelenie normalizovaného blokového maxima závislej postupnosti je GEV rozdelenie s rovnakým parametrom tvaru ϕ ako limita pridruženého procesu (len parameter polohy a škálovací parameter sa môžu zmeniť).

Teda, pre veľkú medzu η , rozdelenie maxima z n pozorovaní časovej rady s extrémálnym indexom θ môže byť aproximované rozdelením maxima z $n\theta < n$ pozorovaní z pridruženého procesu. Číslo $n\theta$ môže byť považované za počet takmer nezávislých zhlukov v n pozorovaniach a θ ako prevrátená hodnota priemerného počtu pozorovaní v zhlukoch. Závislosť v stacionárnej postupnosti $\{X_i\}_{i \in \mathbb{N}}$ má za efekt pomalšiu konvergenciu ku GEV rozdeleniu, z dôvodu $n\theta < n$, čo znamená, že pri odhade GEV rozdelenia maxima v iid prípade si musíme vybrať väčšie bloky.

4.1.1 Odhady extrémálneho indexu

Existuje niekoľko spôsobov, ako odhadnúť extrémálny index θ striktne stacionárnej postupnosti $\{X_i\}_{i \in \mathbb{N}}$. Každá metóda je spojená s inou interpretáciou extrémálneho indexu. V tejto sekcii poskytneme odhad založený na metóde behov. Jednoduchší odhad poskytol napr. Smith a Weissman (1994) alebo autori Ferro a Segers (2003), ktorí odvodili intervalový odhad extrémálneho indexu.

Metóda behov

Metóda je založená na limitnom výsledku O'Briena, ktorý dokázal, že za určitých podmienok platí:

$$\lim_{n \rightarrow \infty} P(M_{1n} \leq u_n | X_1 > u_n) = \theta,$$

kde $M_{1n} = \max_{2 \leq i \leq s} X_i$, pričom s je funkcia veľkosti výberu n splňujúca určité rastúce podmienky, zahrňujúcu napr. $s \rightarrow \infty$ a $s/n \rightarrow 0$ pri $n \rightarrow \infty$. Pre viac detailov a dôkaz tohto tvrdenia viď (O'Brien, 1987). Tento výsledok sa použije na odhad extrémálneho indexu θ založenom na behoch nasledovne:

$$\hat{\theta} = \frac{\sum_{i=1}^{n-k} \mathbb{1}_{[A_{in}]}}{\sum_{i=1}^n \mathbb{1}_{[X_i > u_n]}} = \frac{\sum_{i=1}^{n-k} \mathbb{1}_{[A_{in}]}}{N(u_n)},$$

kde $N(u_n)$ je počet presahov výberu $\{X_1, \dots, X_n\}$ cez medzu u_n , k je funkcia n a $A_{in} = \{X_i > u_n, X_{i+1} \leq u_n, \dots, X_{i+k} \leq u_n\}$. A_{in} teda označuje udalosť, kedy za presahom nasleduje beh k pozorovaní, ktoré sú pod medzou u_n .

4.1.2 Výpočet VaR s využitím extrémálneho indexu

Vzťah medzi distribučnou funkciou maxima striktne stacionárneho procesu a distribučnou funkciou maxima jeho pridruženého procesu z vety 6 bude použitý na výpočet VaR. Predpokladáme, že skúmané straty tvoria striktne stacionárnu postupnosť. Z (4.1) vyplýva, že $(1-p)$ -tý kvantil distribúcie $F(x)$ je $(1-p)^{n\theta}$ -tým kvantom limitného GEV rozdelenia M_n . Presnejšie, VaR môže byť spočítaný miernou modifikáciou vzťahu (3.7) ako:

$$VaR_{1-p} = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\phi}} \{1 - [-n\theta \ln(1-p)]^{-\hat{\phi}}\}, & \hat{\phi} \neq 0, \\ \hat{\mu} - \hat{\sigma} \ln[-n\theta \ln(1-p)], & \hat{\phi} = 0. \end{cases} \quad (4.2)$$

Z výsledného vzorca (4.2) vidieť, že pri ignorovaní podmienky iid strát, resp. prehliadnutí extrémálneho indexu, riskujeme podhodnotenie VaR.

4.2 Využitie vysvetľujúcich premenných

Často máme k dispozícii vysvetľujúce premenné, ktoré môžu ovplyvniť správanie $\{X_i\}_{i \in \mathbb{N}}$. Predpokladajme, že $\mathbf{x} = (x_{1t}, \dots, x_{\nu t})^T$ je vektor vysvetľujúcich premenných, ktorý je dostupný do času t . Jednoduchá metóda ako použiť vysvetľujúce premenné je predpokladať, že tri skúmané parametre ϕ , μ , σ sa menia v čase a sú lineárne funkcie vysvetľujúcich premenných, presnejšie:

$$\begin{aligned} \phi_t &= \gamma_0 + \gamma_1 x_{1t} + \dots + \gamma_\nu x_{\nu t} = \gamma_0 + \boldsymbol{\gamma}^T \mathbf{x}_t, \\ \sigma_t &= \delta_0 + \delta_1 x_{1t} + \dots + \delta_\nu x_{\nu t} = \delta_0 + \boldsymbol{\delta}^T \mathbf{x}_t, \\ \mu_t &= \beta_0 + \beta_1 x_{1t} + \dots + \beta_\nu x_{\nu t} = \beta_0 + \boldsymbol{\beta}^T \mathbf{x}_t. \end{aligned}$$

Ak $\boldsymbol{\gamma} = \mathbf{0}$, potom je parameter tvaru $\phi_t = \gamma_0$, teda je invariantný v čase. Takže, testovanie štatistickej signifikantnosti $\boldsymbol{\gamma}$ nám môže poskytnúť informácie o príspevku vysvetľujúcich premenných do parametra tvaru. Rovnaké metódy sa môžu aplikovať na škálovací a lokačný parameter. Taktiež, v predchádzajúcich rovniciach sme využili rovnaké vysvetľujúce premenné pre všetky tri parametre ϕ_t, σ_t, μ_t . V praxi môžu byť použité rôzne vysvetľujúce premenné pre rôzne parametre. Pre ilustráciu, Tsay (2010, kap. 7.7.8) vo svojej knihe poskytol návrhy na rôzne vysvetľujúce premenné.

4.3 Očistenie rady výnosov

V tejto sekcii už upustíme od obecného prípadu, kedy sme uvažovali straty portfólia. Od tejto časti sa zameriame výhradne len na záporné výnosy portfólia, ktoré budú použité v empirickej štúdiu a celý prístup očistenia dát vysvetlíme na tomto prípade. Úlohou tejto sekcie bude príprava na nasledujúcu empirickú štúdiu, kde bude potreba odhadnúť rádovo tisíce odhadov VaR za účelom porovnania metód. Je zrejmé, že by bolo nepraktické a časovo náročné sa zaoberať každým odhadom zvlášť. Preto sa pokúsime navrhnúť vlastný zautomatizovaný prístup, ktorý nám umožní rutinným spôsobom pripraviť dáta tak, aby na nich

mohla byť EVT aplikovateľná.

Ako už bolo spomenuté v kapitole 1.3, rad reprezentujúci finančné výnosy $\{r_t\}_{t \in \mathbb{N}}$ vykazuje podmienenú heteroskedasticitu. My sa teraz pokúsime zmierniť tento problém tzv. „očistením dát“ reprezentujúcich výnosy použitím niektorého z GARCH modelov. Obdobný postup použili vo svojom článku aj McNeil a Frey (2000). Tí na očistenie dát ale použili GARCH(1,1) model a VaR odhadli len na základe POT metódy. Tento postup v tejto práci značne zobecníme a miesto modelu GARCH(1,1) budeme podmienenú strednú hodnotu výnosov modelovať pomocou AR procesu a podmienený rozptyl pomocou GJR-GARCH procesu, ktorý dokáže zohľadniť pákový efekt u výnosov. Takisto ukážeme ako následne aplikovať na očistené dáta všetky metódy založené na EVT za účelom výpočtu VaR.

Definícia 11. Modelom $AR(p)$ -GJR-GARCH(m, s) rádov p, m, s budeme nazývať model, ktorý modeluje výnosy r_t stochastickým procesom:

$$\begin{aligned} r_t &= \varphi_1 r_{t-1} + \dots + \varphi_p r_{t-p} + e_t, \\ e_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m (\alpha_i e_{t-i}^2 + \gamma_i \mathbb{1}_{[t-i]} e_{t-i}^2) + \sum_{j=1}^s \beta_j \sigma_{t-j}^2, \end{aligned}$$

kde $\{\epsilon_t\}$ sú iid náhodné veličiny s nulovou strednou hodnotou a jednotkovým rozptylom, $\mathbb{1}_{[t]}$ nadobúda hodnotu 1 pre $e_t < 0$ a 0 inak; $\varphi_1, \dots, \varphi_p \in \mathbb{R}$, $\varphi_p \neq 0$, $\alpha_0 > 0$, $\alpha_i \geq 0$ a $\alpha_i + \gamma_i \geq 0$ pre $i = 1, \dots, m$ a $\beta_j \geq 0$ pre $j = 1, \dots, s$ sú neznáme parametre.

Diskutovaná metóda sa aplikuje nasledovne:

1. Uvažujme rad n pozorovaní a označme súčasný okamih ako čas t , teda máme k dispozícii výnosy $\{r_v\}_{v=t-n+1}^t$. Na vyrovnanie týchto výnosov použijeme AR(1)-GJR-GARCH(1,1) model popísaný definíciou 11. Teoreticky by bolo možné odhadnúť viac modelov resp. modely rôznych rádov a z nich pomocou určitého kritéria (napr. Akaikeho, ...) vybrať najvhodnejšie rády. Kvôli už teraz dosť veľkej výpočetnej náročnosti sa obmedzíme výhradne na modely 1. rádu. Tento predpoklad nepovažujeme za nijako obmedzujúci, keďže aj renomovaní autori (McNeil a Frey, 2000) vo svojich štúdiách úspešne použili najjednoduchšie verzie GARCH modelov. Výsledné modely nebudeme v našom zautomatizovanom postupe nijako verifikovať (z dôvodu časovej náročnosti). V našej analýze ale ukážeme vývoj parametrov modelu v čase a budeme diskutovať o ich stabilite.

Z odhadnutého modelu vyexportujeme reziduá, ktoré zároveň štandardizujeme, explicitnejšie, vycentrujeme ich a predelíme podmienenou štandardnou odchýlkou. Tým dostaneme rad $\{\epsilon_v = \hat{\sigma}_v^{-1}(r_v - \hat{\mu}_v) : v = (t - n + 1), \dots, t\}$. Prvky tohto radu budeme pokladať za náhodný výber iid pozorovaní s t rozdelením.

2. Potrebujeme určiť hodnoty parametrov v skúmaných metódach, špecificky veľkosť bloku g v BM metóde, konštantu k v Hillovom odhade a medzu η

v POT metóde. Nanešťastie neexistuje žiaden automatický spôsob, ako ich určiť. V BM metóde budeme dáta deliť do skupín tak, aby sa jednotlivé blokové maximá dali považovať za nezávislé a zároveň, aby delenie do skupín bolo v istom zmysle „rozumné“. Tsay (2010, kap. 7.2.2) odporúča deliť denné výnosy do mesačných resp. štvrtročných blokov. Ako sa ďalej v praktickej časti ukáže, my budeme deliť denné výnosy do mesačných blokov a týždenné výnosy do polročných blokov. Pri odhade konštanty k v Hillovom odhade nedisponujeme žiadnou empirickou štúdiou, ktorá by sa zaoberala vhodným výberom tejto konštanty, preto bude v ďalšej časti práce komentovaná vlastná ad hoc voľba. V prípade odhadu medze η v POT metóde, McNeil a kol. (2005, kap. 7.2) diskutuje o rôznych grafických technikách, ktoré boli bližšie rozobrané v sekcii 3.3.2. Pre naše účely v praktickej časti, kde budeme potrebovať odhadnúť rádovo tisícky hodnôt VaR, by ale bolo prakticky neuskutočniteľné vyšetriť jednotlivo každý prípad zvlášť. Namiesto toho použijeme podobný postup ako McNeil a Frey (2000), a zafixujeme počet presahov cez medzu N_η , kde N_η bude predstavovať 10% z počtu n .

3. Konečne, podmienený 1 denný odhad VaR pre čas $(t + 1)$ spočítame ako:

$$VaR_{1-p}^{t+1} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} VaR_{1-p}, \quad (4.3)$$

kde $\hat{\mu}_{t+1}$ resp. $\hat{\sigma}_{t+1}$ sú predikcie podmienenej strednej hodnoty, resp. podmienenej štandardnej odchýlky o jeden krok dopredu, t. j. pre čas $(t + 1)$ v odhadnutom modeli AR(1)-GJR-GARCH(1,1) a VaR_{1-p} je odhad VaR pre štandarizované reziduá v tomto modeli spočítaný na základe vzťahov (3.7), (3.9) a (3.14) s dosadenými parametrami z kroku 2.

5. Empirická štúdia: metodológia

V tejto kapitole predstavíme metodológiu, ktorá bude použitá v empirickej štúdií. Po prvé, v sekcii 5.1 popíšeme výber a charakter dát a v sekcii 5.2 ukážeme, ako jednotlivé metódy medzi sebou porovnáme.

5.1 Výber dát

K empirickej štúdií sme si na analýzu vybrali tri dátové súbory. Voľba nižšie zmienených dátových súborov bola jednak z dôvodu jednoduchej dostupnosti dát, kvôli veľkému počtu dostupných dát, a ďalej kvôli dobrej reprezentácii trhu. Prvý, na ktorom bude predstavená detailná analýza všetkých diskutovaných metód, je jeden z najznámejších akciových indexov a to Dow Jones Industrial Average (DJI), ktorý zachycuje vývoj akcií 30 veľkých amerických spoločností.⁴ Ďalej, za účelom všeobecnejšieho porovnania zopakujeme v krátkosti prevedené postupy na denné výnosy akcií americkej banky JPMorgan, aby sme vyšetrili počínanie metód aj na výnosoch akcií tvorených jedinou spoločnosťou.⁵ V poslednej časti analýzy budeme uvažovať *týždenné* výnosy ďalšieho známeho akciového indexu Russell 2000, ktorý v porovnaní s DJI meria počínanie približne 2000 menších tzv. „small-cap“ amerických spoločností.⁶

Ďalej sa budeme zaoberať výlučne indexom DJI. Pri DJI sme uvažovali denné uzatváracie ceny akciového indexu očistené od dividend počas obchodných dní od 1.1.1997 do 11.11.2016, dohromady 5001 pozorovaní. Na základe nich sme spočítali **záporné** logaritmické výnosy, skrátene ďalej len *log výnosy*, zavedené v (1.2). K dispozícii teda máme 5000 log výnosov, čo predstavuje približne 20 obchodných rokov (uvažujeme 252 dňový obchodný rok).

Nech n_T označuje počet všetkých pozorovaných log výnosov, $n_T = 5000$. Analýzu rozdelíme do dvoch častí. V prvej časti vyberieme vzorku prvých n log výnosov, ktoré budeme naďalej označovať ako *okno*. Na tomto okne otestujeme všetky predpoklady, ktoré sú potrebné na to, aby sme mohli aplikovať EVT. Ak dáta nebudú spĺňať predpoklady, pokúsime sa ich niektorým z postupov zmienených v kapitole 4 modifikovať tak, aby bola EVT aplikovateľná.

V ďalšej časti sa pokúsime tieto metódy porovnať s metodológiou RiskMetrics EWMA a následne medzi sebou. Uvedomme si, že po aplikácii EVT na pozorované okno bude výstupom odhad **jedného** jednodenného VaR, ktorý použijeme ako odhad výšky extrémnej straty na skúmanej konfidenčnej hladine pre čas $n + 1$. Na základe jednej hodnoty nie sme schopní zrovnáť tieto štyri metódy medzi

⁴<http://chart.finance.yahoo.com/table.csv?s=~DJI&a=0&b=1&c=1997&d=10&e=11&f=2016&g=d&ignore=.csv>, navštívené 24.03.2017.

⁵<http://chart.finance.yahoo.com/table.csv?s=JPM&a=0&b=1&c=1997&d=10&e=11&f=2016&g=d&ignore=.csv>, navštívené 24.03.2017.

⁶<http://chart.finance.yahoo.com/table.csv?s=~RUT&a=8&b=10&c=1987&d=10&e=14&f=2016&g=w&ignore=.csv>, navštívené 24.03.2017.

sebou. Preto použijeme metódu „rolovacieho okna“, t. j. okno vybrané v prvej časti použijeme pre odhad VaR pre čas $n + 1$, potom vezmeme log výnosy z časov $2, \dots, n + 1$ a tie použijeme pre odhad jednodenného VaR pre čas $n + 2, \dots$. Takto pokračujeme až po okno zložené z výnosov $n_T - n, \dots, n_T - 1$. Celkovo dostaneme $n_T - n$ jednodenných odhadov VaR. Na základe týchto výstupov už budeme schopní porovnať metódy medzi sebou štatistickými testmi bližšie popísanými v nasledujúcej sekcii.

5.2 Porovnanie metód

Zatiaľ sme sa zaoberali len rôznymi metódami odhadov VaR. V tejto sekcii budeme diskutovať, ako porovnať jednotlivé metódy medzi sebou. Využijeme bežne používanú techniku nazývanú spätné testovanie (z angl. backtesting), ktoré prevedieme pre všetky nami skúmané konfidenčné hladiny, ktoré budú 95 %, 97,5 %, 99 % a 99,5 %. Jeho aplikácia je nasledovná. Opakovane budeme implementovať diskutované metódy na rolovacie okno o veľkosti n a výsledného odhadu VaR porovnáme so skutočnými pozorovanými log výnosmi. Počas vyhodnocovania budeme porovnávať pozorovaný počet presahov odhadu VaR, nazvime ich ako *porušenia* s očakávaným počtom presahov na danej konfidenčnej hladine.

Pri testovaní sa budeme sústrediť na dve kritéria:

1. Zhoduje sa počet porušení VaR (označme π) s očakávaným počtom presahov?
2. Sú porušenia VaR náhodne rozmiestnené v čase?

Pre danú konfidenčnú hladinu $(1 - p)$ je očakávaný počet presahov cez VaR_{1-p} z definície p . Ak π je väčšie ako p , metóda podceňuje riziko, čo môže viesť k nedostatočnej kapitálovej požiadavke. Naopak, ak π je menšie ako p , metóda preceňuje skutočné riziko, čo spôsobuje zbytočne veľkú alokáciu kapitálu. Pre skúmaný súbor log výnosov a odhadov VaR, by mali byť počty π a p približne rovnaké.

Pre 1 denné odhady VaR budeme formálne testovať tieto dve kritéria najskôr samostatne a neskôr združené. Aplikujeme testy založené na pomere vierohodností (LR z angl. Likelihood Ratio), ktoré navrhol Christoffersen (1998). Christoffersen špecifikoval LR test správneho nepodmieneného pokrytia (Unconditional Coverage Test), LR test nezávislosti porušení (Independence Test) a združený LR test správneho pokrytia a nezávislosti (Conditional Coverage Test).

5.2.1 Nepodmienené testovanie pokrytia

V UC (z angl. Unconditional Coverage) teste testujeme, či skutočný počet porušení π je signifikantne odlišný od očakávaného počtu porušení p . Matematicky testujeme:

$$\begin{aligned} H_0 &: \pi = p, \\ H_1 &: \pi \neq p. \end{aligned}$$

Testová štatika je:

$$LR_{uc} = -2 \ln \left[\frac{(1-p)^{m_0} p^{m_1}}{(1-\hat{\pi})^{m_0} \hat{\pi}^{m_1}} \right] \stackrel{as.}{\sim} \chi^2(1),$$

kde m je celkový počet odhadov VaR, m_1 je počet porušení a $m_0 = m - m_1$. Počet porušení $\hat{\pi}$ je jednoducho odhadnutý ako $\hat{\pi} = m_1/m$.

5.2.2 Test nezávislosti

Ďalej chceme testovať druhé kritérium, t. j. či je počet porušení rozmiestnený náhodne v čase. Testujeme nulovú hypotézu nezávislosti, t. j. pravdepodobnosť porušenia nasledujúci deň nezávisí od toho, či nastalo porušenie v terajšom čase. Nech i, j predstavujú udalosti, označené ako 0 a 1, kde 0 značí neporušenie a 1 značí porušenie, m_{ij} je počet a π_{ij} je pravdepodobnosť, že po udalosti i bude nasledovať udalosť j . Testujeme:

$$H_0 : \pi_{01} = \pi_{11},$$

$$H_1 : \pi_{01} \neq \pi_{11}.$$

Testová štatistika je:

$$LR_{ind} = -2 \ln \left[\frac{(1-\hat{\pi})^{m_0} \hat{\pi}^{m_1}}{(1-\hat{\pi}_{01})^{m_{00}} \hat{\pi}_{01}^{m_{01}} (1-\hat{\pi}_{11})^{m_{10}} \hat{\pi}_{11}^{m_{11}}} \right] \stackrel{as.}{\sim} \chi^2(1).$$

Pravdepodobnosť dňa porušenia, pred ktorým nenastalo porušenie $\hat{\pi}_{01}$ odhadneme jeho maximálne vierohodným odhadom $\hat{\pi}_{01} = m_{01}/(m_{00} + m_{01})$. Analogicky odhadneme $\hat{\pi}_{11} = m_{11}/(m_{10} + m_{11})$.

5.2.3 Podmiené testovanie pokrytia

V CC (z angl. Conditional Coverage) teste testujeme obe predošlé kritéria združene. Pri tomto teste sa jednoducho sčítajú dve predošlé testové štatistiky: $LR_{cc} = LR_{uc} + LR_{ind}$. Táto testová štatistika má asymptoticky $\chi^2(2)$.

Poznámka. Tieto tri testy nie sú jedinou možnosťou ako porovnať dané metódy. Existuje sada alternatív, ktoré by mohli byť tiež aplikované. Avšak my tieto tri testy pokladáme za akýsi základ, jednak z pohľadu jednoduchej implementovateľnosti v porovnaní so zložitejšími testami, ktoré napr. vyžadujú vykonávať opakovaný výpočet VaR, čo je dosť výpočetne a časovo náročné. Na druhej strane, tieto štatistické testy ponúkajú omnoho sofistikovanejší spôsob ako porovnať metódy v porovnaní s jednoduchým prístupom založeným len na numerickom porovnaní skutočného a očakávaného počtu porušení.

6. Empirická štúdia: výsledky

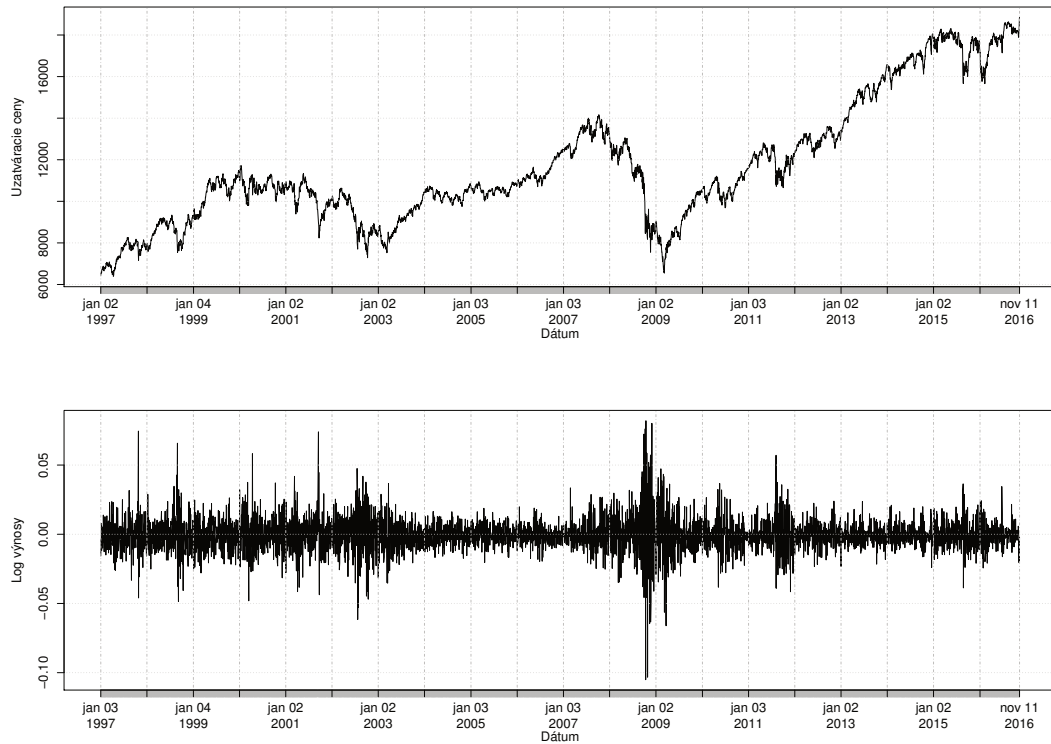
Všetky výpočty v tejto kapitole budú vykonané za pomoci softvéru `RStudio` (RStudio Team, 2015) a sčasti za pomoci softvéru `Mathematica 10` (Wolfram Research, 2015). Prvotná analýza bude vykonaná na okne s dĺžkou $n = 1500$ log výnosov. Všetky testy budú vykonané na 5% hladine spoľahlivosti. Výsledky budú prevažne zaokrúhlené na 4 desatinné miesta (v prostredí `RStudio` sa bude počítať s nezaokrúhlenými hodnotami), v prípade potreby detailnejšieho porovnania čísel bude počet miest rozšírený.

6.1 Úvodná dátová analýza a deskriptívne štatistiky

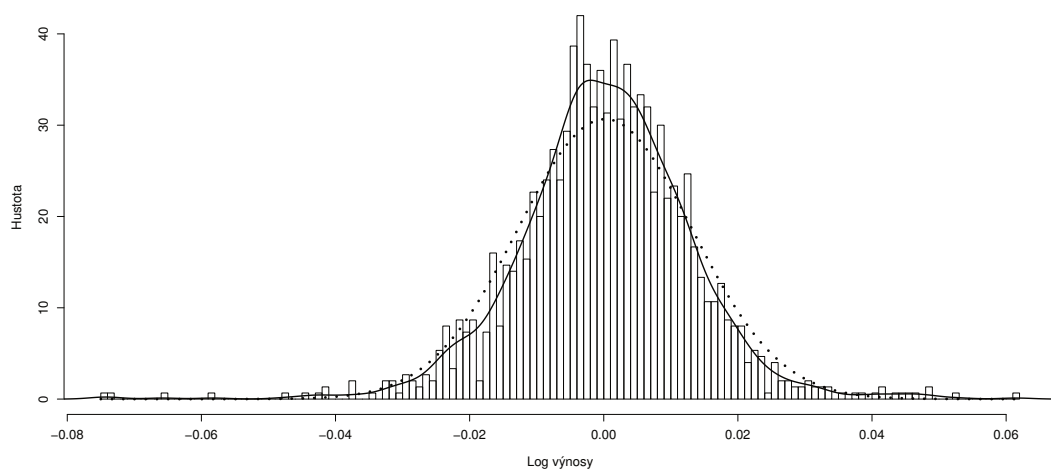
Časový rad reprezentujúci ceny akciového indexu DJI spolu s log výnosmi týchto cien $\{r_t\}_{t=1}^{5000}$ je zobrazený na obr. 6.1. Tento graf zobrazuje jednu zo základných čít finančných časových radov, ktoré boli diskutované na začiatku v sekcii 1.3, a to zhukovanie volatilit. Tento jav je vidieť najlepšie počas obdobia krízy medzi rokmi 2008–2009. Ďalšie základné charakteristiky sú potvrdené deskriptívnymi štatistikami (tab. 6.1) a histogramom log výnosov (obr. 6.2). Obidva výbery, okno a celkový výber, majú vysoký ukazovateľ špicatosti. Predpoklad normálne rozdelených log výnosov je formálne zamietnutý Jarqueovým-Berovým (JB) testom (Jarque a Bera, 1980). Ljungovej-Boxov (LB) test (Ljung a Box, 1978) spočítaný pre 21 autorokorelácií, čo odpovedá približne jednému mesiacu v obchodnom roku, zamietá hypotézu o nekorelovanosti log výnosov v celkovom výbere, ale nie v skúmanom okne (tab. 6.2). Avšak ako bolo spomenuté v sekcii 1.3, štvorcové log výnosy sú silno korelované (obr. 6.3), čo implikuje dočasnú závislosť. Tento fakt je potvrdený testom na dodatočný ARCH efekt založenom na Lagrangeových multiplikátoroch (LM) opísanom Englom (Engle, 1982), ktorý zamietá hypotézu o podmienenej homoskedasticite v oboch prípadoch, pre skúmané okno a pre celkový výber (tab. 6.2). Teda, v našich dátach je jasne viditeľné, že log výnosy nie sú výberom z iid náhodných veličín, čo nás nabáda k modifikácii dát pred použitím jednej z metód založených na teórii extrémnych hodnôt.

Štatistika	Okno	Celý výber
Minimum	-0,0615	-0,1051
1.kvartil	-0,0080	-0,0059
Medián	-0,0003	-0,0005
3. kvartil	0,0072	0,0051
Maximum	0,0745	0,0820
Rozptyl	0,0002	0,0001
Št. odchýlka	0,0130	0,0117
Šikmosť	0,2249	0,1424
Špicatosť	3,0224	7,4267

Tabuľka 6.1: Deskriptívne štatistiky okna a celého výberu log výnosov indexu DJI.



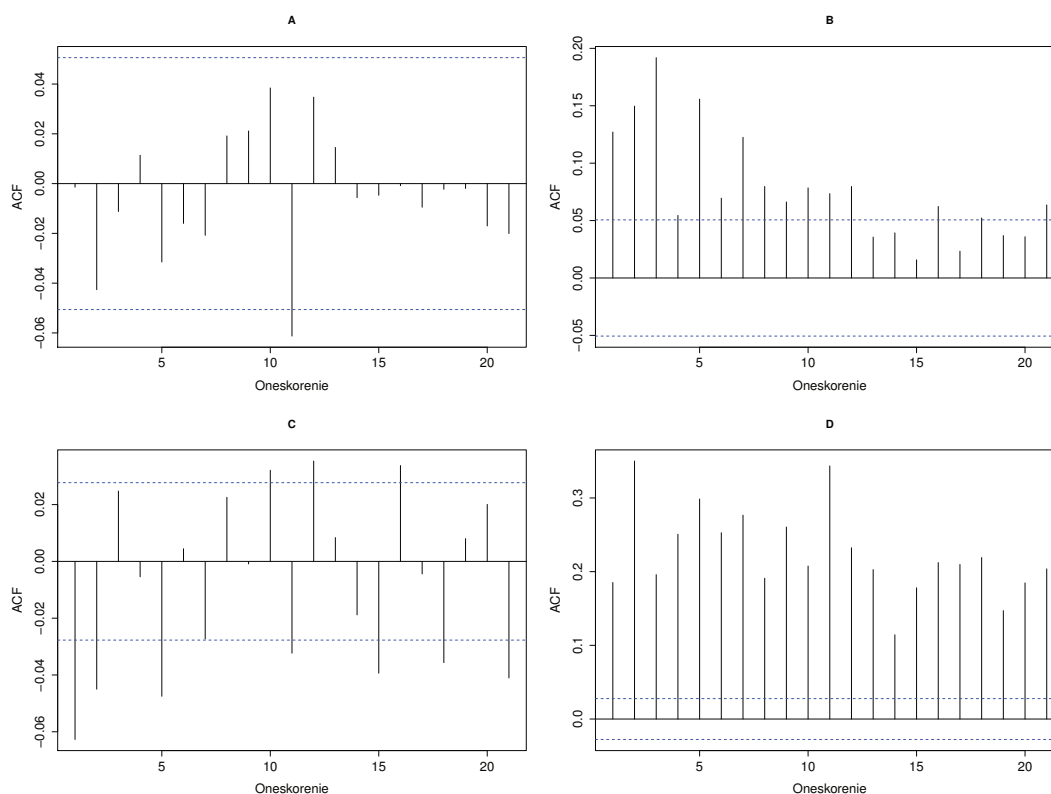
Obr. 6.1: Uzatváracie ceny a negatívne logaritmicke výnosy akciového indexu DJI.



Obr. 6.2: Histogram log výnosov okna s odhadom hustoty (plná čiara) v porovnaní s normálnym rozdelením so strednou hodnotou a rozptylom okna (bodkovaná čiara).

Test	Okno	Celý výber
JB	583,56 (0,0000)	11508,00 (0,0000)
LB(21)	18,23 (0,6346)	100,31 (0,0000)
LB ₂ (21)	253,43 (0,0000)	5651,00 (0,0000)
LM(21)	125,82 (0,0000)	1318,50 (0,0000)

Tabuľka 6.2: Hodnoty testových štatistík a príslušné p-hodnoty v zátvorkách Jarqueovho-Berovho testu normality (JB), Ljungovej-Boxovho testu pre pôvodné výbery (LB(21)) a štvorcové výbery (LB₂(21)) spolu s testom na dodatočný ARCH efekt (LM(21)).



Obr. 6.3: Korelogramy log výnosov okna (A) a ich štvorcov (B) spoločne s korelogram log výnosov celkového výberu (C) a ich štvorcov (D).

6.2 Modifikácia dát

Log výnosy pred aplikáciou metód založených na EVT modifikujeme posledným postupom zmieným v kapitole 4. Výber tohto prístupu má niekoľko odôvodnení. Zobecnenie pomocou extrémálneho indexu je extrémne citlivé na výber medze. Po vykonaní vlastnej analýzy sme nedospeli k žiadnej rozumnej ceste ako ju nastaviť vo všeobecnosti a domnievame sa, že to musí byť skúmané od prípadu k prípadu. Ďalej, priamo nedisponujeme žiadnymi vysvetľujúcimi premennými. V prípade, ak by čitateľ chcel využiť tento prístup, odporúčame zahrnutie napr. inflácie za skúmané obdobie, alebo inšpirovať sa vysvetľujúcimi premen-

Par.	Odhad	Št. chyba	t hod.	p-hod.
φ_1	0,001294	0,024998	0,0518	0,9587
α_0	0,000006	0,000000	90,7524	0,0000
α_1	0,139478	0,019748	7,0630	0,0000
β_1	0,905700	0,007929	114,2262	0,0000
γ_1	-0,163992	0,023620	-6,9430	0,0000
df	10,641124	2,499459	4,2574	0,0000

Tabuľka 6.3: Odhady koeficientov v modeli AR(1)-GJR-GARCH(1,1) pre log výnosy okna.

nými, ktoré navrhol Tsay (2010, kap. 7.7.8). My sme hľadali spôsob, ktorý by sme mohli aplikovať rutinne, bez použitia externých premenných a ktorý by bol aplikovateľný na ľubovoľný rad s finančnými výnosmi. Preto sme sa rozhodli pre tretí prístup z kapitoly 4, ktorý je podľa nás najefektívnejší a najrobustnejší.

Teraz sa zameriame výhradne na skúmané okno 1500 log výnosov, na ktorom ukážeme vzorový postup, ktorý bude neskôr, kvôli testovaniu a zrovnaniu metód zautomatizovaný a bude použité rolovacie okno na získanie dostatku dát pre účely spätného testovania.

Log výnosy vyrovnáme AR(1)-GJR-GARCH(1,1) modelom s predpokladom t rozdelených štandarizovaných reziduí. Podotknime ešte raz, že vhodnejším postupom by bolo vyskúšanie širšej škály modelov, z ktorých by sa neskôr napr. za pomoci Akaikeho kritéria vybral ten najvhodnejší model. My sa kvôli veľkému počtu odhadovaných hodnôt zameriame výhradne na model AR(1)-GJR-GARCH(1,1), ktorý je podľa nás v našom prípade najvhodnejším kandidátom z triedy ARFIMA-GARCH modelov, jednak vďaka schopnosti vysporiadať sa s podmienenou heteroskedasticitou a ďalej kvôli schopnosti zohľadniť pákový efekt. Parametre boli v R odhadnuté za pomoci podmienenej metódy maximálnej vierohodnosti, hybridnou optimalizačnou procedúrou, t. z. pri zlyhaní prvej optimalizačnej procedúry sa pokračuje na ďalšiu, . . . Pre viac informácií o optimalizačných procedúrach viď knižnicu (Ghalanos, 2015, str. 72–73). Označenie odhadnutých parametrov (tab. 6.3) je zhodné so značením z definície 11, teda prvý koeficient odpovedá AR modelu pre podmienenú strednú hodnotu a zvyšné koeficienty odpovedajú GJR-GARCH modelu pre podmienený rozptyl. Posledný parameter df (z angl. degrees of freedom) odpovedá odhadu stupňov voľnosti t rozdelenia reziduí. Odhady parametrov týkajúce sa podmieneného rozptylu sú všetky štatisticky signifikantné, parameter φ_1 odpovedajúci podmienenej strednej hodnote je štatisticky nesignifikantný, čo odpovedá faktu, že log výnosy okna sú takmer nekorelované (obr. 6.3A). Na tomto mieste by sme mohli byť vystavení kritike voči použitiu AR procesu na modelovanie strednej hodnoty, keď to zjavne nie je potrebné. Je to hlavne kvôli väčším výkyvom, ktoré nastávajú v neskorších obdobiach, hlavne v období krízy.

Z modelu vyexportujeme štandarizované reziduá, špecifickejšie $\{\epsilon_v = \hat{\sigma}_v^{-1}(r_v - \hat{\mu}_v) : v = 1, \dots, n\}$. Z grafu týchto štandarizovaných reziduí (obr. 6.4 hore) už ne-

pozorujeme porušenie podmienenej heteroskedasticity, ani porušenie predpokladu nekorelovanosti (obr. 6.4 v strede). LM test pre ARCH efekt je nesignifikantný s hodnotou testovej štatistiky 15,78 (0,7820) a LB test s hodnotou testovej štatistiky 16,85 (0,7199) neindikuje signifikantnú korelovanosť reziduí ani štvorcov reziduí s hodnotou testovej štatistiky 16,18 (0,7592) pre 21 oneskorených hodnôt. Výsledky testov indikujú, že modelom AR(1)-GJR-GARCH(1,1) sa nám podarilo očistiť pôvodné log výnosy. Poznamenajme, že normalita je opäť s hodnotou testovej štatistiky 107,75 a takmer nulovou p-hodnotou formálne zamietnutá. Stále pozorujeme vysokú špicatosť. Ďalej, teoretické kvantily normálneho rozdelenia sú na koncoch výrazne odlišné od empirických kvantilov a skôr sa prikláňajú ku kvantilom t rozdelenia (obr. 6.4 dole), čo indikuje, že reziduá majú rozdelenie s ťažkými chvostami a použitie EVT bude opodstatnené.

Na tomto mieste takisto spočítame odhad podmienenej strednej hodnoty a podmienenej štandardnej odchýlky z modelu (6.3) o jeden krok dopredu, teda pre 1501. obchodný deň, ktoré budú potrebné pre výpočet VaR v nasledujúcej sekcii. Po výpočte v R sme dostali $\hat{\mu}_{1501} = 0,000013$ a $\hat{\sigma}_{1501} = 0,0129$.

6.3 Aplikácia metód na okno

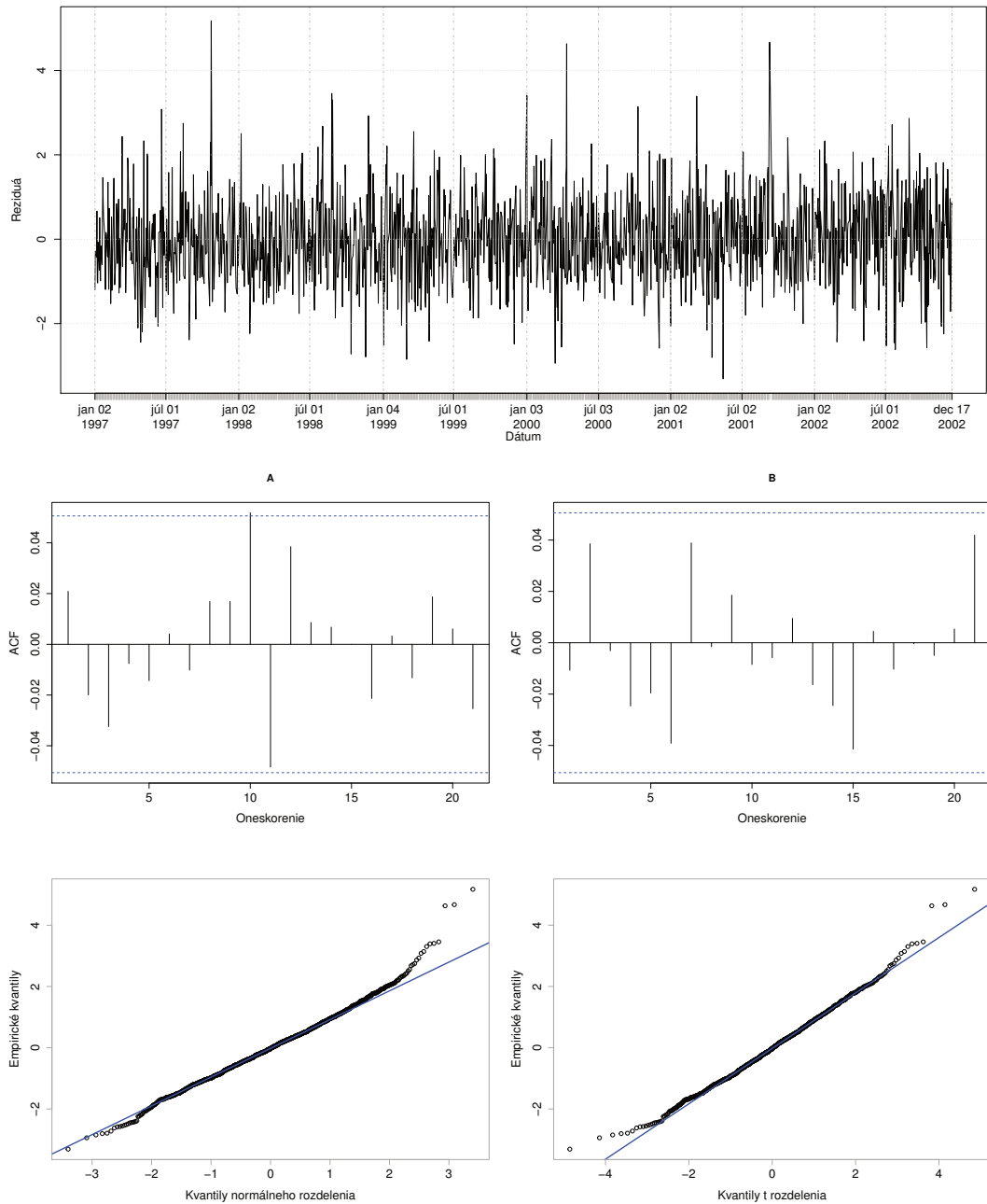
Táto sekcia bude prípravou pre hlavnú časť empirickej štúdie. Detailne popíšeme aplikáciu všetkých skúmaných metód pre výpočet VaR okna, ktorá bude v nasledujúcej sekcii zautomatizovaná za účelom zrovnania metód.

6.3.1 RiskMetrics

Pri odhade VaR metodológiou RiskMetrics sme vychádzali z rovnice (2.1), kde sme za λ dosadili hodnotu 0,94, ktorá je odporúčaná pri denných dátach (Morgan, 1996, str. 97–101), za r_t^2 sme dosadili štvorec log výnosu v čase 1500 a odhad $\sigma_{1501|1500}^2$ sme spočítali rekurzívne za pomoci spomínaného vzťahu (2.1). Počiatočnú hodnotu σ_{10}^2 sme zvolili ako 0 (v čase 0 nemáme žiadnu informáciu) s následnou aplikáciou schématu (2.1) pre rekurentný výpočet $\sigma_{2|1}^2, \sigma_{3|2}^2, \dots$. Týmto sme dostali odhad $\sigma_{1501|1500}^2 = 0,0002$. Odmocninu z tohto odhadu sme následne pre získanie VaR_{1-p} vynásobili $(1-p)$ -tým kvantilom štandardného normálneho rozdelenia, teda: $VaR_{1-p}^{1501} = 0,0145\Phi^{-1}(1-p)$.

6.3.2 BM metóda

Pri aplikácii BM metódy je potrebné určiť veľkosť bloku tak, aby boli maximá v jednotlivých blokoch nezávislé a zároveň, aby sme mali dostatok blokových maxim, ktoré vstupujú do rovnice (3.4) pre odhad metódou maximálnej vierohodnosti. Z priebehu celej práce je jasné, že denné log výnosy nie sú nezávislé. Tsay (2010, kap. 7.5.2) a viaceré empirické štúdie odporúčajú použitie mesačných až štvrtročných výnosov. My sa vzhľadom k veľkosti nášho okna rozhodneme pre mesačné bloky, kde za mesiac budeme pokladať obdobie $g = 252/12 = 21$ za sebou idúcich obchodných dní.



Obr. 6.4: Graf štandardizovaných reziduí z modelu 6.3 (hore) spolu s korelogramom reziduí (A) a štvorcov reziduí (B) a QQ grafmi, ktoré zobrazujú empirické kvantily voči kvantilom štandardného normálneho rozdelenia (vľavo dole) a kvantilom t rozdelenia s 10 stupňami voľnosti (vpravo dole).

Po výpočte v \mathbb{R} dostávame odhady parametrov $\hat{\phi} = 0,1183$, $\hat{\mu} = 1,6770$ a $\hat{\sigma} = 0,5281$. Na odhad VaR pre štandardizované reziduá použijeme vzťah (3.7):

$$VaR_{1-p} = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\phi}} \{1 - [-g \ln(1-p)]^{-\hat{\phi}}\},$$

ktorý následne pomocou rovnice (4.3):

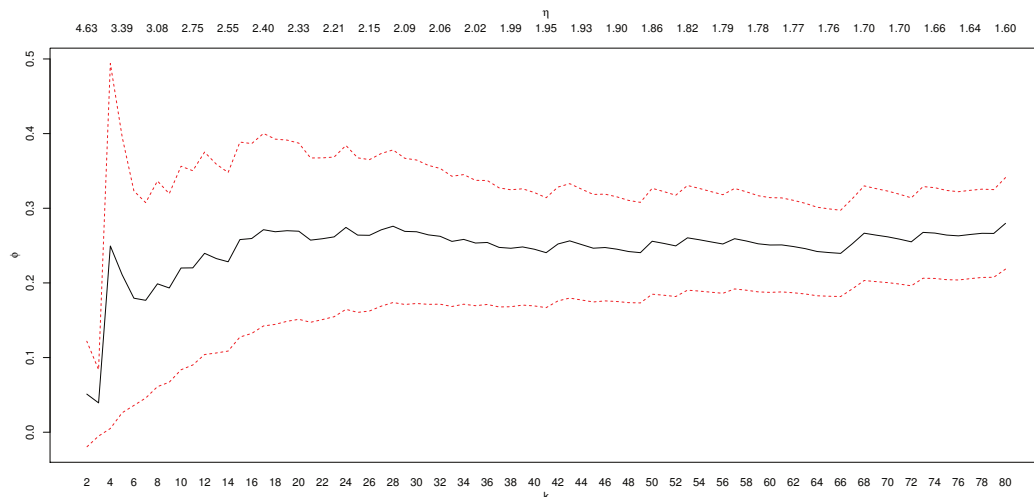
$$VaR_{1-p}^{1501} = \hat{\mu}_{1501} + \hat{\sigma}_{1501} VaR_{1-p},$$

prevedieme na VaR pre pôvodné log výnosy okna.

6.3.3 Hillova metóda

Pri Hillovom odhade čelíme inému problému. Našou úlohou je nájsť vhodnú konštantu k , resp. medzu $\epsilon^{(k)}$, kde $\epsilon^{(k)}$ značí k -tú poriadkovú štatistiku štandardizovaných reziduí. K tomuto účelu si skonštruujeme Hillov graf diskutovaný na konci sekcie 3.2.3 s tým rozdielom, že miesto parametra tvaru α_k budeme uvažovať rovno jeho prevrátenú verziu $\phi_k = 1/\alpha_k$, teda skonštruujeme graf $\{(k, \hat{\phi}_k) : k = 2, \dots, n\}$. V tomto grafe sa snažíme nájsť stabilnú oblasť, kde sú odhady skonštruované z rôzneho počtu poriadkových štatistík podobné.

Poznámka. Podotkneme, že Hillova metóda je aplikovateľná jedine v prípade, ak skúmané dáta pochádzajú z Fréchetovej sféry príťažlivosti alebo inak, majú kladný parameter tvaru ϕ . Ako vidíme, pri BM metóde a neskôr ako uvidíme pri POT metóde, v oboch prípadoch dostaneme kladný parameter tvaru a tento predpoklad môžeme teda pokladať za legitímny.



Obr. 6.5: Hillov graf. Na vodorovných osiach je vyobrazená konštantu k resp. medza $\epsilon^{(k)}$ voči parametru tvaru ϕ_k na zvislej osi.

Po vykreslení grafu (obr. 6.5) pozorujeme stabilnú oblasť približne za konštantou $k = 45$. Pri tejto voľbe dostávame v \mathbb{R} odhad parametra tvaru $\hat{\phi}_{45} = 0,2465$.

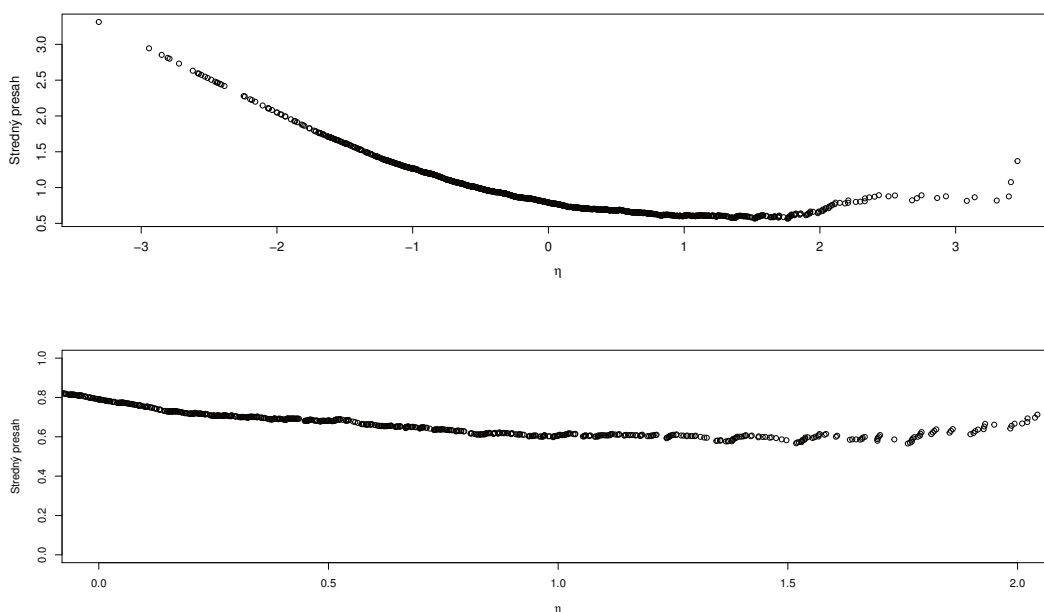
Na odhad VaR pre štandardizované reziduá použijeme vzťah (3.9):

$$VaR_{1-p} = \epsilon^{(k)} \left(\frac{np}{k} \right)^{-\hat{\phi}_k},$$

ktorý následne pomocou rovnice (4.3) prevedieme na VaR pre pôvodné log výnosy okna.

6.3.4 POT metóda

Pri POT metóde musíme analogicky ako pri Hillovom odhade určiť medzu, alternatívne počet presahov cez túto medzu, ktoré využijeme pri odhadoch parametrov. Pre stanovenie tejto medze použijeme graf funkcie stredného presahu cez medzu η zavedený v kapitole 3.3.2. V rovnici (3.13) sme odvodili empirický odhad tejto funkcie. Po skonštruovaní grafu $\{(\epsilon^{(i)}, e_n(\epsilon^{(i)})) : 2 \leq i \leq n\}$, (obr. 6.6), kde $\epsilon^{(i)}$ značí analogicky ako pri Hillovom odhade i -tú poriadkovú štatistiku štandardizovaných reziduí pozorujeme, že od medze približne 1,0 je graf funkcie lineárny a to mierne rastúci (na detailnejšom grafu boli kvôli prehľadnosti useknuté pravé krajné hodnoty, kde sa priemeruje malý počet hodnôt s veľkým presahom). Tento bod môžeme pokladať za vhodný výber medze η a kvôli mierne rastúcemu trendu očakávame odhad parametra tvaru ϕ , ktorý bude mierne väčší ako 0. Po výpočte v R dostávame odhady $\hat{\phi} = 0,0379$ a $\hat{\psi}(\eta) = 0,5829$.



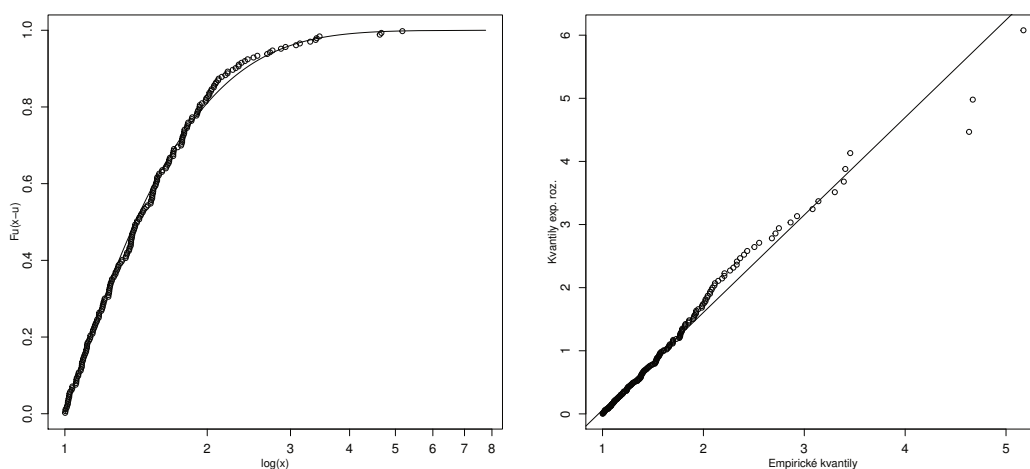
Obr. 6.6: Graf stredného presahu cez medzu η , hore pre celé okno reziduí, dole detailnejší pohľad pre reziduá z intervalu (0,2).

Na odhad VaR pre štandardizované reziduá použijeme vzťah (3.14):

$$VaR_{1-p} = \eta - \frac{\hat{\psi}(\eta)}{\hat{\phi}} \left(1 - \left(\frac{np}{N_\eta} \right)^{-\hat{\phi}} \right),$$

ktorý následne opäť pomocou rovnice (4.3) prevedieme na VaR pre pôvodné log výnosy okna.

Poznámka. Zobecneným Paretovým rozdelením s týmito parametrami sa nám podarilo slušne odhadnúť pravý chvost distribučnej funkcie presahu cez medzu η (obr. 6.7). Na obr. 6.7 taktiež vidíme QQ graf, kde porovnáваме teoretické kvantily exponenciálneho rozdelenia s parametrom $\hat{\phi}$ a empirickými kvantilmi zobecneného Paretoho rozdelenia s odhadnutými parametrami. Ten tiež potvrdzuje, že skúmané štandardizované reziduá majú mierne ťažšie konce ako tie v Gumbelovej sfére príťažlivosti.



Obr. 6.7: Naľavo graf distribučnej funkcie presahu cez medzu η spolu s vyrovnávanými hodnotami, napravo QQ graf empirických kvantilov vs. kvantilov exponenciálneho rozdelenia.

6.3.5 Porovnanie

Výsledky odhadov VaR jednotlivými metódami sme pre prehľadnosť zhrnuli v tabuľke 6.4. Pri BM metóde a POT metóde pozorujeme veľmi podobné výsledky. Zaujímavý fakt je, že Hillova metóda má všetky odhady VaR menšie v porovnaní s týmito metódami, až na VaR na hladine 95 %, kedy sa prekvapivo jeho hodnota razantne zväčšila a prevýšila odhady pomocou BM a POT metódy. Ak porovnáme metódy založené na EVT s metodológiu RiskMetrics, tak odhady VaR na hladinách menších než 99,5 % sú väčšie (až na 99% VaR odhadnutý POT metódou), zatiaľ, čo na hladine 99,5 % sa podľa očakávaní situácia preklápa a metodológia RiskMetrics podhodnocuje VaR a tým pádom aj celkové riziko v porovnaní s EVT metódami.

Hladina	RM	BM	Hill	POT
95 %	2,3844	2,1129	2,1909	2,1089
97,5 %	2,8412	2,6099	2,5989	2,6586
99 %	3,3723	3,3261	3,2571	3,4077
99,5 %	3,7340	3,9195	3,8637	3,9919

Tabuľka 6.4: Odhady jednodenných VaR v percentách pre 1501. obchodný deň metódami RiskMetrics (RM), metódou blokových maxím (BM), Hillovou metódou (Hill) a metódou založenou na presahoch cez medzu (POT).

6.4 Použitie rolovacieho okna

V nasledujúcej časti budeme chcieť nejakým spôsobom porovnať doposiaľ použité metódy. Na základe jediných hodnôt z tab. 6.4 to nie je možné, a preto musíme vykonať dodatočný počet odhadov jednodenných VaR pre dostatočne dlhý časový horizont. Ako už bolo diskutované v sekcii 4.3, vykonanie vyššie uvedených postupov pre tisícky odhadov je časovo veľmi náročné a nemôžeme sa zaoberať každým prípadom zvlášť. Špecificky máme na mysli určenie konštanty k pri Hillovom odhade a medze η pri POT metóde. Použijeme preto vlastný zautomatizovaný postup, ktorý nastaví hodnoty parametrov, ktoré sa budú používať počas celej doby.

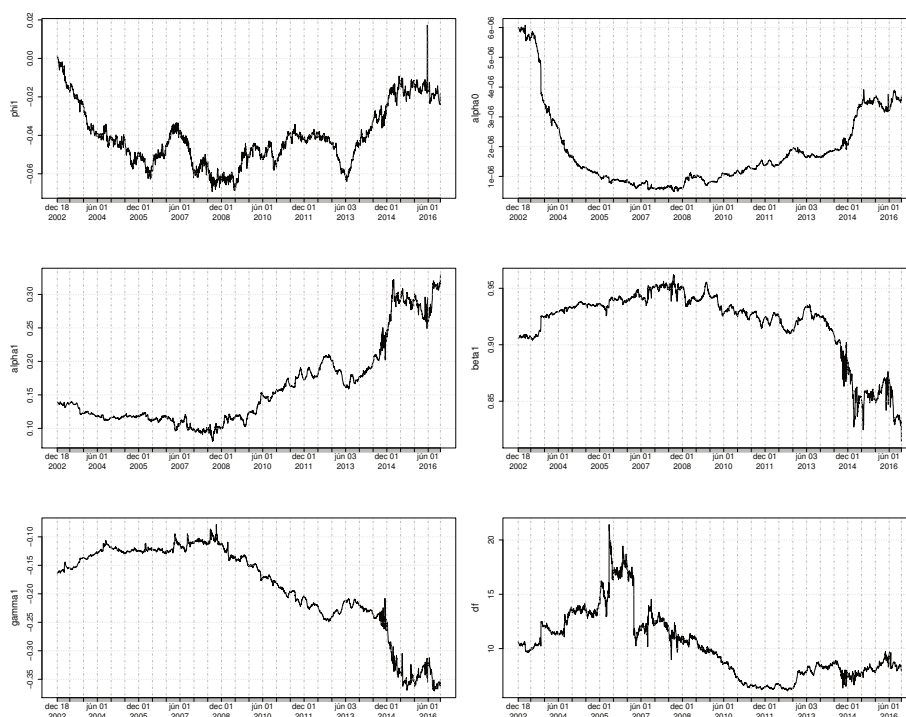
Cieľom tejto sekcie bude aplikovať metódy na rôzne typy dát pri rôznych voľbách parametrov a tým ukázať ich dobré počínanie a výsledky v širšom zmysle. Analýzu začneme doposiaľ rozoberaným indexom DJI. Potom v krátkosti aplikujeme rovnakú štúdiu na denné log výnosy jednej z najväčších amerických bánk JPMorgan, kde budeme uvažovať menšie okno v porovnaní s DJI. Analýzu zakončíme aplikáciou metód na akciový index Russell 2000, kde budeme uvažovať týždenné dáta a okno s ešte menšou dĺžkou ako v prípade banky JPMorgan.

6.4.1 Dow Jones Industrial Average

Pri BM metóde budeme uvažovať mesačné bloky, keďže pri mesačných blokoch je možné predpokladať nezávislosť. Pri Hillovej metóde nedisponujeme žiadnou empirickou štúdiou, ktorá by odporúčala vhodnú voľbu konštanty k len na základe vzorky dát, teda bez opakovanej konštrukcie Hillových grafov. Pri pohľade na Hillove grafy prvých 10 okien sa nám javí ako rozumný predpoklad vziať konštantu $k = 45$, teda zahrnúť do výpočtu odhadu parametra tvaru 45 najväčších hodnôt skúmaného okna. Zdôraznime však ešte raz, že ide o ad hoc voľbu, ktorá nie je podložená žiadnymi materiálmi a bola vybraná len na základe povahy skúmanej dátovej sady. Pri POT metóde potrebujeme určiť medzu η . Voľba bola inšpirovaná viacerými simulačnými štúdiami. Po prvé, Tsay (2010, str. 284) odporúča určiť medzu η tak, aby minimálne 5 % dát presahovalo cez túto medzu. McNeil ďalej ukázal, že pre okno o 1000 pozorovaniach, fixovaný počet 100 presahov vedie k dobrým odhadom VaR a ES (McNeil a Frey, 2000). Kuester a kol. (2006) navyše zistili, že vhodný interval pre iid dáta s t rozdelením s rôznymi stupňami voľnosti je 80–150 presahov. Navyše McNeil a Frey (2000) taktiež zistili, že GDP metóda je robustná vzhľadom k voľbe N_η . Vzhľadom k týmto zisteniam

a veľkosti nášho okna nastavíme počet presahov cez medzu η na $N_\eta=150$, čo odpovedá 10% počtu pozorovaní okna. Pri tejto voľbe by medza η z kapitoly 6.3.4 mala byť 1,2. Pri pohľade na graf (6.6) môžeme túto voľbu pokladať za adekvátnu. Pri RiskMetrics metodológii opäť vychádzame z rovnice (2.1) s počiatočnými parametrami diskutovanými v minulej sekcii. Odhad ďalších hodnôt VaR je priamy, stačí využiť rovnicu (2.1) a odmocninu z výsledku vynásobiť príslušným kvantilom štandardného normálne rozdelenia na skúmanej konfidenčnej hladine.

Ako sme už neraz spomínali, výber modelu AR(1)-GJR-GARCH(1,1) bol ad hoc voľba. Bol napevno určený pre všetky rolovacie okná a žiadna verifikácia modelu z dôvodu časovej a výpočetnej náročnosti pri odhadoch neprebíha. Na tomto mieste by sa mohla naskytnúť otázka, či je náš výber opodstatnený a či je vhodný pre každé rolovacie okno. Obr. 6.8 zobrazuje vývoj koeficientov modelov AR(1)-GJR-GARCH(1,1) v čase, alebo inak, zobrazuje odhady koeficientov pre každé rolovacie okno. Ako je vidieť, vývoj koeficientov je v čase stabilný a nedochádza k signifikantným výkyvom, preto si myslíme, že náš výber modelu bol skutočne opodstatnený.



Obr. 6.8: Vývoj koeficientov AR(1)-GJR-GARCH(1,1) modelov v čase v prípade indexu DJI.

Poznámka. Pôvodná myšlienka bola použiť ešte obcejšiu štruktúru ARMA-GJR-GARCH modelu. Avšak pri použití ARMA(1,1)-GJR-GARCH(1,1) dochádzalo k veľkej nestabilite koeficientov v ARMA modeli pre podmienenú strednú hodnotu log výnosov. Kvôli tomuto faktoru bol model ARMA(1,1) nahradený modelom AR(1), kde už vývoj koeficientu u AR(1) členu je v čase stabilný.

Po vykonaní výpočtov v R s parametrami diskutovanými v minulých odstavcoch sme dostali $n_T - n = 3500$ odhadov jednodenných VaR pre každú zo 4 metód. Tieto hodnoty použijeme na spätné testovanie bližšie opísané v kapitole 5.2. Z tab. 6.5 je zrejmé, že metodológia RiskMetrics zlyhala v UC teste na všetkých konfidenčných hladinách. Pri CC teste je situácia podobná až na 95% hladinu, kde nemôžeme zamietnuť hypotézu CC testu. BM a POT metóda vykazujú úplne identické správanie, čo sa týka výsledkov štatistických testov pri odhade 99,5% a 99% VaR. Na týchto hladinách ich v UC a CC teste mierne predčila Hillova metóda, pri ktorej nastal mierne vyšší počet porušení. V prípade 97,5% VaR zamietame pri BM metóde hypotézu testu nezávislosti a v prípade 95% odhadu VaR pri Hillovej metóde je štatistická signifikantnosť UC testu tak veľká, že zamietame hypotézu UC i CC testu.

Metóda	p	π	LR_{uc}	LR_{ind}	LR_{cc}	\overline{VaR}
RM	5,0	5,77	*4,1865	0,6187	4,8052	1,5647
	2,5	3,69	*17,6555	0,4009	*18,0564	1,8645
	1,0	2,00	*27,3953	3,4391	*30,8344	2,2130
	0,5	1,49	*44,6034	1,4388	*46,0422	2,4503
BM	5,0	4,49	2,0157	0,0921	2,1078	1,6635
	2,5	2,31	0,5077	*3,8863	4,3939	2,0815
	1,0	0,94	0,1177	0,6473	0,7650	2,6184
	0,5	0,40	0,7555	0,1205	0,8760	3,0202
Hill	5,0	3,86	*10,4113	0,6849	*11,0962	1,7755
	2,5	2,49	0,0029	0,0636	0,0666	2,0663
	1,0	0,97	0,0291	0,6868	0,7159	2,5266
	0,5	0,46	0,1331	0,1562	0,2892	2,9430
POT	5,0	4,71	0,6127	0,1856	0,7983	1,6489
	2,5	2,49	0,0029	0,0636	0,0666	2,0499
	1,0	0,94	0,1177	0,6473	0,7650	2,5852
	0,5	0,40	0,7555	0,1205	0,8760	2,9975

Tabuľka 6.5: Výsledky spätného testovania v prípade akciového indexu DJI. Hviezdička značí štatistickú signifikantnosť na 5% konfidenčnej hladine. Kritické hodnoty pre 5% konfidenčnú hladinu sú 5,9915 pre LR_{cc} a 3,8415 pre LR_{uc} a LR_{ind} , p je očakávaný podiel porušení v percentách a pozorovaný podiel porušení v percentách je označený π . LR_{uc} , LR_{ind} , LR_{cc} sú testové štatistiky testov 5.2.1, 5.2.2, resp. 5.2.3. \overline{VaR} je priemerný VaR v percentách počas testovanej periódy.

Dospeli sme k názoru, že najvhodnejšou metódou pre skúmanú radu log výnosov by bola jednoznačne POT metóda. Tradičnú metodológiu RiskMetrics by sme neodporúčali využiť v žiadnom z prípadov. To všetko potvrdzujú aj grafické výstupy (obr. 6.13, 6.14, 6.15, 6.16), kde sú zobrazené skutočné log výnosy počas obchodných dní od 18.12.2002 do 11.11.2016 spolu s dennými predikciami VaR pre každú skúmanú hladinu. Vizualne najlepšie je to viditeľné pri 99% a 99,5% hladine, kde skutočné log výnosy pri RiskMetrics prekračujú VaR vo viacerých dňoch v porovnaní s metódami založenými na teórii extrémnych hodnôt.

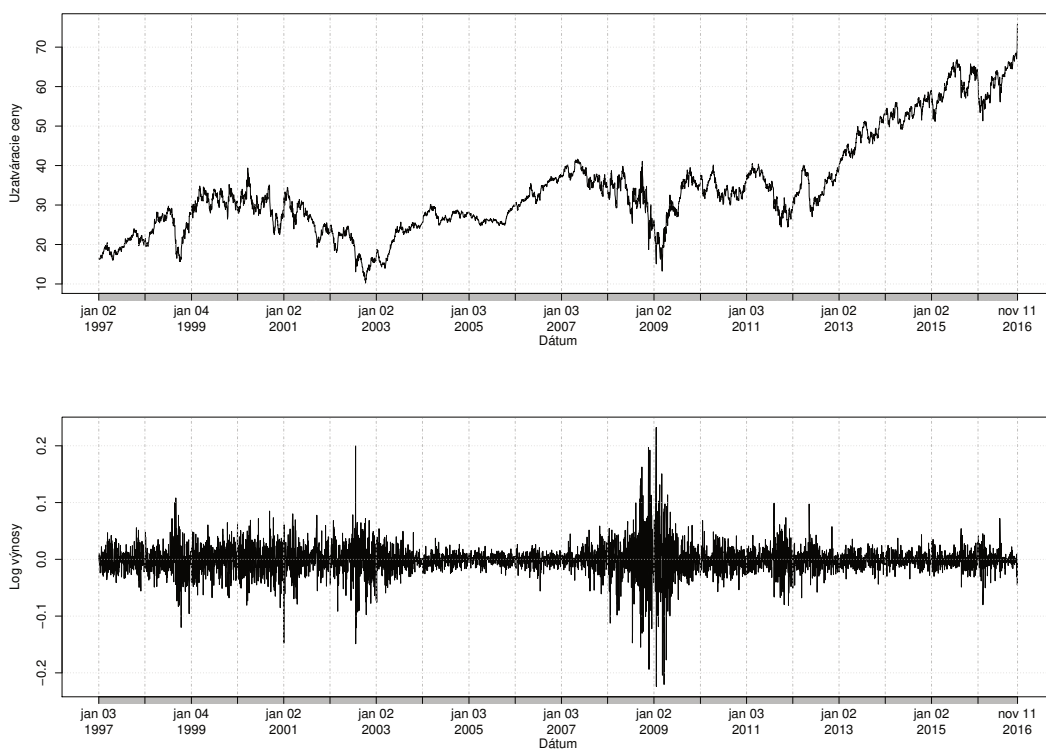
6.4.2 JPMorgan Chase & Co.

V tejto časti budeme znova uvažovať denné log výnosy v období od 1.1.1997–11.11.2016, čo je dohromady $n_T = 5000$ log výnosov s tým rozdielom, že tentokrát sa bude jednať o akcie jedinej finančnej inštitúcie a to americkej banky JPMorgan a miesto okna o veľkosti 1500 použijeme okno o veľkosti $n = 1000$, aby sme ukázali, ako robustné sú metódy pri zmene veľkosti okna. Rovnako ako v predošlej časti, pri BM metóde budeme uvažovať mesačné bloky, pri Hillovej metóde konštantu k zmenšíme proporčne na hodnotu $k = 30$ a pri POT metóde budeme pre výpočet používať presahy tvorené 10 % pozorovaní z celkového počtu log výnosov okna, teda $N_\eta=100$.

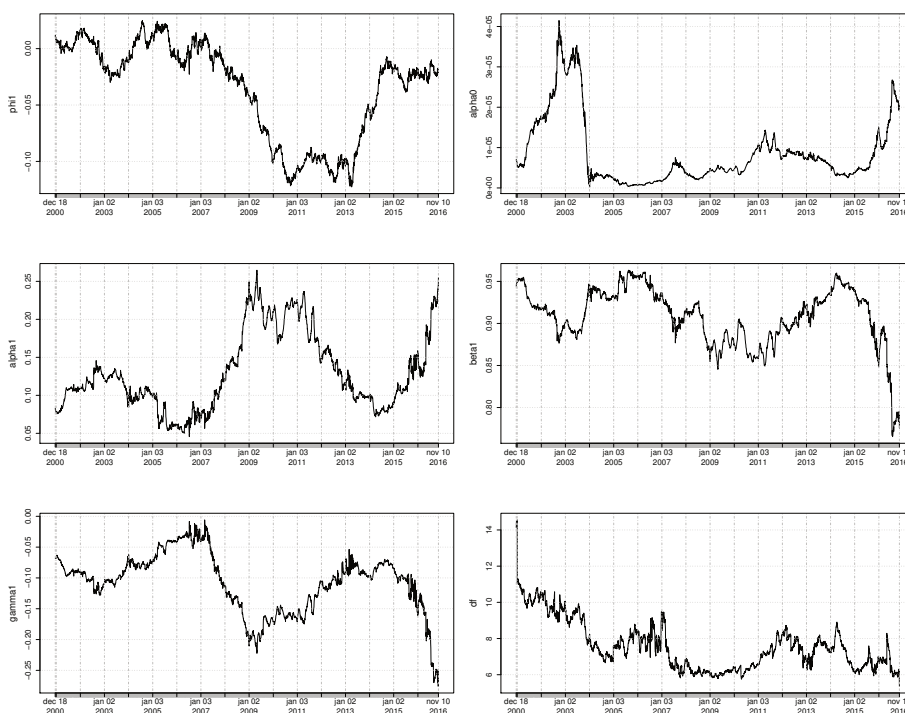
Na začiatok uvedieme základné deskriptívne štatistiky log výnosov celého súboru a deskriptívne štatistiky uzatváracích cien akcií očistených od dividend (tab. 6.6) a graf týchto veličín v čase (obr. 6.9). Pozorujeme opäť vysokú špicatosť. Log výnosy podobne ako v predošlom prípade očistíme AR(1)-GJR-GARCH(1,1) modelmi. Vývoj koeficientov týchto modelov v čase sa zdá byť opäť stabilný (obr. 6.10). Pokiaľ ide o výsledky (viď tab. 6.7), metodológia RiskMetrics opäť zlyhala na všetkých skúmaných konfidenčných hladinách, tentokrát dokonca aj v testoch nezávislosti. Jediná výnimka je len v nepodmienenom teste pokrytia na 95% hladine. Pri metódach založených na EVT zamietame hypotézu testu nezávislosti vo všetkých prípadoch na 95% konfidenčnej hladine. V POT metóde bola hodnota testovej štatistiky tak signifikantná, že zamietame hypotézu aj združeného CC testu. Odhliadnuc od toho sú metódy na vyšších hladinách spoľahlivé. Hypotézy UC, CC testu a testu nezávislosti nezamietame pri žiadnej z metód v prípadoch 97,5%, 99% a 99,5% VaR. V tomto prípade je ťažké rozhodnúť, ktorú metódu môžeme pokladať za najvhodnejšiu, keďže každá z metód je „lepšia“ v inom aspekte. Grafické porovnanie je zobrazené na obrázkoch 6.17, 6.18, 6.19, 6.20.

Štatistika	Výnosy	Ceny
Minimum	-0,2239	10,2303
1.kvartil	-0,0108	25,9264
Medián	-0,0002	31,3493
3. kvartil	0,0105	37,7646
Maximum	0,2323	75,8323
Št. odchýlka	0,0256	12,4269
Šikmosť	-0,2392	0,9584
Špicatosť	11,5682	0,3086

Tabuľka 6.6: Deskriptívne štatistiky negatívnych logaritmickejých výnosov a cien akcií JPM.



Obr. 6.9: Uzatváracie ceny očistené od dividend a negatívne logaritmicke výnosy akcií JPM.



Obr. 6.10: Vývoj koeficientov AR(1)-GJR-GARCH(1,1) modelov v čase v prípade akcií JPM.

Metóda	p	π	LR_{uc}	LR_{ind}	LR_{cc}	\overline{VaR}
RM	5,0	4,75	0,5348	*7,9233	*8,4581	3,3789
	2,5	3,15	*6,4139	*7,0961	*13,5100	4,0262
	1,0	1,73	*17,4542	*4,3491	*21,8033	4,7788
	0,5	1,30	*35,6312	*4,5331	*40,1643	5,2913
BM	5,0	5,10	0,0837	*5,2361	5,3468	3,2940
	2,5	2,48	0,0103	2,2193	2,2296	4,1476
	1,0	0,90	0,4181	0,9457	1,3637	5,3062
	0,5	0,50	0,0000	0,2111	0,2111	6,2151
Hill	5,0	4,65	1,0552	*4,3794	5,4346	3,4533
	2,5	2,70	0,6400	1,3623	2,0023	4,1012
	1,0	1,07	0,2219	2,9177	3,1395	5,1518
	0,5	0,52	0,0494	0,2322	0,2817	6,1254
POT	5,0	5,25	0,5182	*5,5015	*6,0197	3,2856
	2,5	2,68	0,4915	2,8525	3,3440	4,1092
	1,0	0,88	0,6591	1,0245	1,6836	5,2617
	0,5	0,52	0,0494	0,2322	0,2817	6,1867

Tabuľka 6.7: Výsledky spätného testovania v prípade akcií JPM. Hviezdička značí štatistickú signifikantnosť na 5% konfidenčnej hladine. Kritické hodnoty pre 5% konfidenčnú hladinu sú 5,9915 pre LR_{cc} a 3,8415 pre LR_{uc} a LR_{ind} , p je očakávaný podiel porušení v percentách a pozorovaný podiel porušení v percentách je označený π . LR_{uc} , LR_{ind} , LR_{cc} sú testové štatistiky testov 5.2.1, 5.2.2, resp. 5.2.3. \overline{VaR} je priemerný VaR v percentách počas testovanej periódy.

6.4.3 Russell 2000

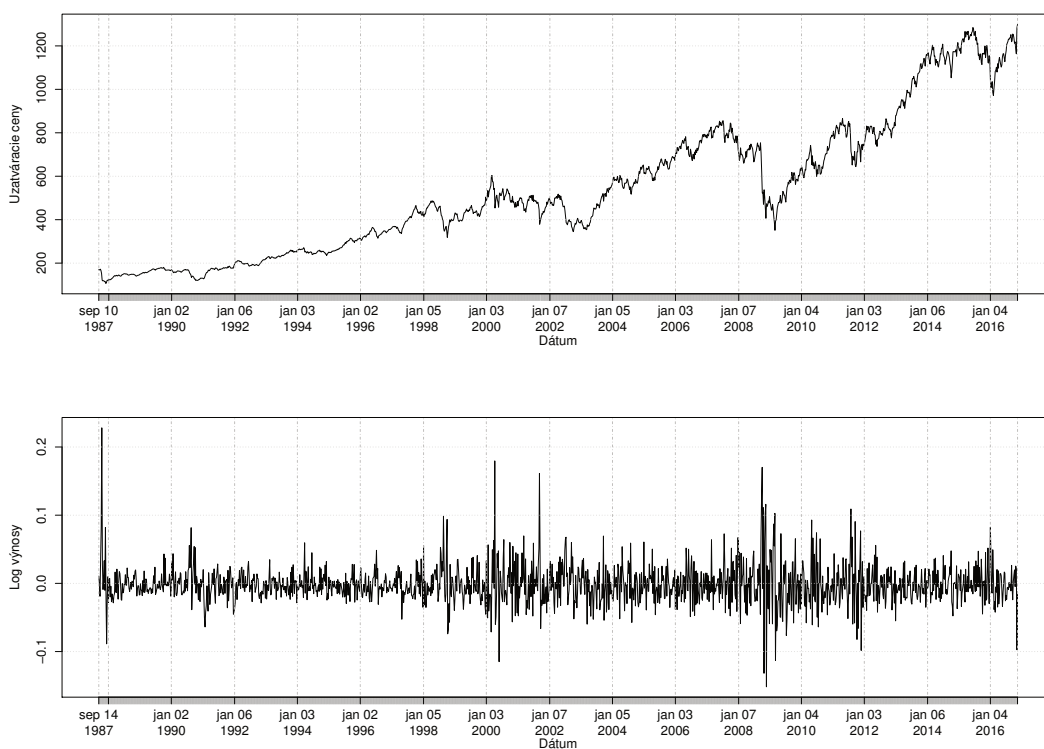
V poslednej časti budeme uvažovať *týždenné* log výnosy ďalšieho známeho akciového indexu Russell 2000, ktorý meria počínanie približne 2000 menších tzv. „small-cap“ amerických spoločností. Keďže uvažujeme týždenné dáta, budeme uvažovať dlhšie časové obdobie od prvej nájdenej historickej ceny až po súčasný okamih, t. j. od 10.9.1987 do 14.11.2016, čo je dohromady $n_T = 1522$ log výnosov. Naším cieľom v tejto sekcii bude zodpovedať otázky, či sa nám dátami s menšou frekvenciou podarí odstrániť závislostnú štruktúru z dát a či metódy založené na EVT poskytnú spoľahlivé výsledky pri relatívne malom rolovacom okne. Budeme uvažovať okno o veľkosti 600 (čo predstavuje približne 11,5 roka), menšie okno nevolíme z dôvodu už teraz pomerne malého počtu pozorovaní, ktoré budú vstupovať napr. pri BM metóde do odhadu parametrov pomocou metódy maximálnej vierohodnosti.

Opäť na začiatok uvedieme základné deskriptívne štatistiky log výnosov celého súboru a deskriptívne štatistiky uzatváracích cien akciového indexu (tab. 6.8) a graf týchto veličín v čase (obr. 6.11). Tentokrát sa pokúsime pre zmenu overiť predpoklad nezávislosti resp. nekorelovanosti pozorovaní v rámci každého okna a vykonáme test na dodatočný ARCH efekt v každom okne, kvôli overeniu, či je pri týždenných dátach potrebné postupné vyrovnanie log výnosov AR(1)-GJR-GARCH(1,1) modelmi, alebo či je možné predpokladať ich nezávislosť a EVT aplikovať priamo na neočistené dáta. Po opakovanom prevedení Ljungovej-Boxovho testu 1. rádu pre log výnosy okna a ich štvorcov a testu na dodatočný ARCH efekt 1. rádu sme zistili, že až v 921 prípadoch z celkového počtu 922 okien, sme zamietli hypotézu o nekorelovanosti 1. rádu štvorcových log výnosov (v 87 prípadoch pre pôvodné log výnosy) v rámci jedného okna a podľa výsledkov ARCH testov je signifikantný ARCH efekt 1. rádu vo všetkých oknách. Z tohto dôvodu aplikujeme rovnaký postup, ako v predošlých dvoch sekciiach, t. j. najprv dáta očistíme AR(1)-GJR-GARCH(1,1) modelmi a až potom aplikujeme EVT. Po odhade parametrov v modeloch pozorujeme stabilný vývoj koeficientov v čase (obr. 6.12).

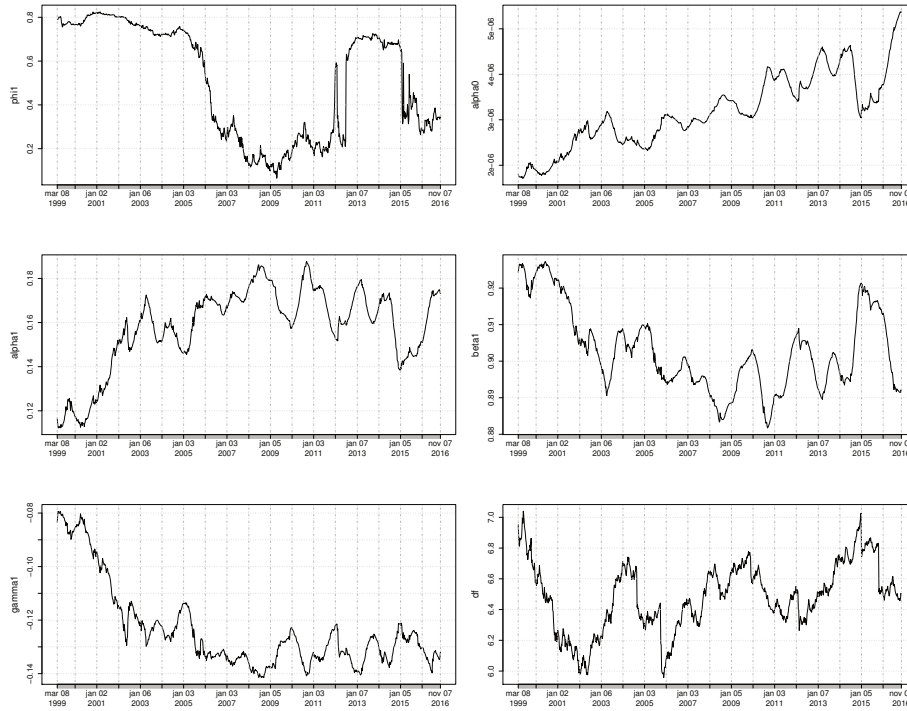
Pri BM metóde budeme uvažovať bloky o veľkosti 24, jednak z dôvodu vhodnej rovnováhy medzi počtom blokových maxim, ktoré vstúpia do odhadov metódou maximálnej vierohodnosti a počtom pozorovaní v každom z blokov, a ďalej kvôli faktu, že 24 týždňov predstavuje približne polovicu roka. Pri Hillovej metóde konštantu k zmenšíme na hodnotu $k = 18$ a pri POT metóde budeme pre výpočet opäť používať presahy tvorené 10% pozorovaní z celkového počtu log výnosov okna, teda $N_\eta=60$. Čo sa týka výsledkov (viď tab. 6.9), metodológia RiskMetrics podobne ako v minulom prípade zlyhala v UC teste na všetkých hladinách vyšších ako 95%. Na druhú stranu, hypotézy o nezávislosti nemôžeme zamietnuť ani na jednej z hladín až na 97,5% hladinu. Metódy založené na EVT si počínali výborne pokiaľ ide o výsledky testov na všetkých skúmaných konfidenčných hladinách. Hypotézy testov nemôžeme ani v jednom z prípadov zamietnuť. Opäť je ťažké určiť dominantnú metódu. Môžeme si ale všimnúť, že na 99,5% hladine si na základe štatistických testov počínali všetky metódy úplne identicky. Grafické porovnanie je na obrázkoch 6.21, 6.22, 6.23, 6.24.

Štatistika	Výnosy	Ceny
Minimum	-0,1517	106,5000
1.kvartil	-0,0160	259,4900
Medián	-0,0036	479,5100
3. kvartil	0,0115	736,5950
Maximum	0,2281	1298,6000
Št. odchýlka	0,0283	310,1932
Šikmosť	1,0018	0,6611
Špicatost'	7,4097	-0,4223

Tabuľka 6.8: Deskriptívne štatistiky negatívnych logaritmických výnosov a cien akciového indexu RUT.



Obr. 6.11: Uzatváracie ceny očistené od dividend a negatívne logaritmické výnosy akciového indexu RUT.



Obr. 6.12: Vývoj koeficientov AR(1)-GJR-GARCH(1,1) modelov v čase v prípade indexu RUT.

Metóda	p	π	LR_{uc}	LR_{ind}	LR_{cc}	\overline{VaR}
RM	5,0	5,53	0,5308	3,2454	3,7762	4,8800
	2,5	3,80	*5,4973	*3,9895	*9,4868	5,8149
	1,0	2,06	*8,0216	3,6337	*11,6553	6,9019
	0,5	1,52	*12,4197	1,6433	*14,0629	7,6421
BM	5,0	6,18	2,5314	1,7962	4,3276	4,8665
	2,5	3,47	3,1859	0,6954	3,8813	6,2229
	1,0	1,41	1,3886	0,4007	1,7893	7,9722
	0,5	0,54	0,0323	0,0655	0,0977	9,2871
Hill	5,0	5,42	0,3384	3,5392	3,8776	5,1498
	2,5	3,25	1,9658	0,9596	2,9254	6,1339
	1,0	1,63	3,0769	0,5295	3,6065	7,7358
	0,5	0,54	0,0323	0,0655	0,0977	9,2258
POT	5,0	5,86	1,3539	2,4559	3,8098	4,9079
	2,5	3,25	1,9658	0,9596	2,9254	6,2386
	1,0	1,19	0,3269	0,2900	0,6169	7,9471
	0,5	0,54	0,0323	0,0655	0,0977	9,2082

Tabuľka 6.9: Výsledky spätného testovania v prípade akciového indexu RUT. Hviezdička značí štatistickú signifikantnosť na 5% konfidenčnej hladine. Kritické hodnoty pre 5% konfidenčnú hladinu sú 5,9915 pre LR_{cc} a 3,8415 pre LR_{uc} a LR_{ind} , p je očakávaný podiel porušení v percentách a pozorovaný podiel porušení v percentách je označený π . LR_{uc} , LR_{ind} , LR_{cc} sú testové štatistiky testov 5.2.1, 5.2.2, resp. 5.2.3. \overline{VaR} je priemerný VaR v percentách počas testovanej periódy.

6.5 Limitácie štúdie

V tejto sekcii v krátkosti slovné zhrnieme porovnanie počínania metód, zdôraznime potenciálne limitácie, ktorým môže byť naša štúdia vystavená a podáme možné návrhy pre ďalší výskum.

Z teoretického uhla pohľadu sme poskytli stručný, ale dôkladný prierez EVT, ktorý pozostával z prehľadu hlavných viet, modelov a štatistických problémov teórie. Prirodzene sme museli upustiť od všetkých dôkazov a detailov, miesto toho boli však uvedené zdroje, kde si čitateľ v prípade záujmu môže tieto poznatky dohľadať. Snažili sme sa dosiahnuť vhodnú rovnováhu medzi teoretickými zisteniami a implementáciou teórie. Z empirického uhla pohľadu sme poskytli oboje, kvantitatívne aj kvalitatívne počínanie metód založených na EVT. Všetky metódy sme porovnali z pohľadu schopnosti odhadnúť VaR pomocou troch kritérií. Tento rámec poskytuje systematickejšie a rozsiahlejšie vyhodnotenie metód v porovnaní s jednoduchým porovnaním metód na základe počtu porušení. Ďalej sme tieto metódy porovnali s tradičnou metodológiou RiskMetrics.

Na základe vykonanej štúdie sme zistili, že metódy založené na EVT si vo všeobecnosti počínajú lepšie ako tradičná metodológia RiskMetrics. Touto štúdiou avšak nechceme úplne zatradiť tradičnú metodológiu. Síce výsledky poskytuje podstatne horšie, no v praxi treba zväziť komplexitu výpočtu nad ich presnosťou. Uvedomme si totiž, že pri RiskMetrics si vystačíme s predošlou hodnotou, ku ktorej váženým spôsobom pripočítame minulý štvorec log výnosu. Pri metódach založených na EVT musíme najprv na očistenie pôvodného radu log výnosov použiť pomerne komplikovaný model, ktorého odhad parametrov nie je jednoduchý. Následne musíme použiť metódu založenú na EVT, ktorá je výpočetne podstatne zložitejšia a až tak môžeme na základe týchto odhadov získať odhad VaR použitím vzorca (4.3). Navyše, ak by sme sa rozhodli zmeniť parametre, museli by sme celý postup znova zopakovať.

Naša empirická štúdia tiež čelí určitým potenciálnym limitáciám, ktoré by mali byť zdôraznené. Prvou potenciálnou limitáciou by mohli byť dáta použité v štúdiu. Čo sa týka použitých dátových súborov, uvažovali sme ceny jedného z najznámejších akciových indexov DJI, ktorý zobrazuje vývoj akcií veľkých amerických spoločností, v kontraste s týmto indexom sme použili ďalší známy index Russell 2000, ktorý zastrešuje 2000 menších „small cap“ spoločností indexu Russell 3000. Taktiež sme uvažovali ceny akcií jedinej spoločnosti, a to americkej banky JPMorgan. Z dôvodu vykonania analýzy na viacerých dátových súboroch, kde každý vykazoval typické vlastnosti finančných časových radov, ako napr. vysoká špicatosť a zhlukovanie volatilit si myslíme, že naše výsledky dávajú jasnú indikáciu o obecné dobrom počínaní metód založených na EVT v prípade finančných časových radov.

Druhým potenciálnym obmedzením by mohlo byť, že naša štúdia je limitovaná počtom pozorovaní použitými v rámci jedného okna. Niekoľko štúdií (napr. Lauridsen, 2000) a (Danielsson a De Vries, 2000)) pri jednoduchších analýzach použilo okno o veľkosti 1500 denných pozorovaní, zatiaľčo ostatní ((Marimoutou a kol.,

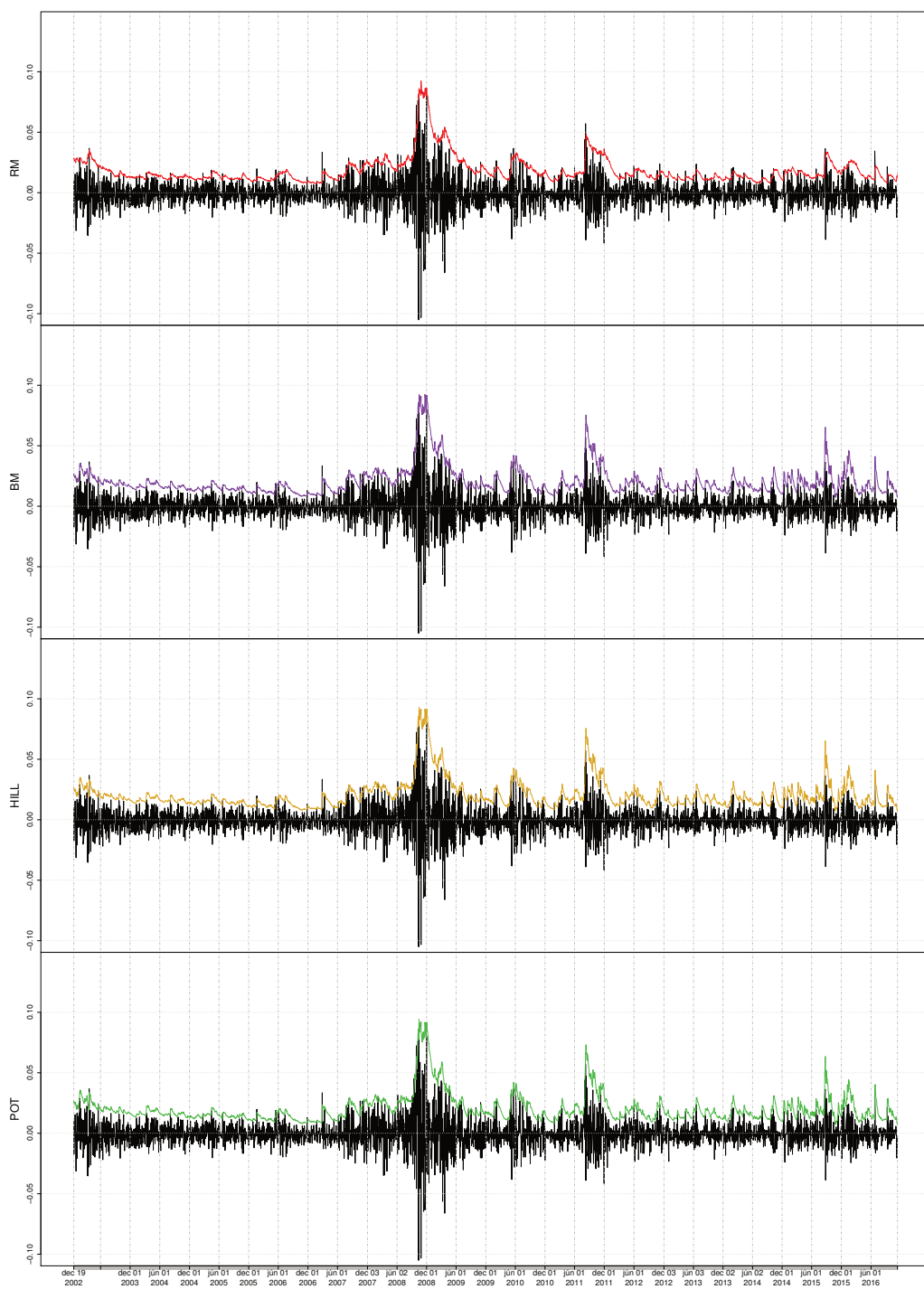
2009) a (McNeil a Frey, 2000)) použili okná s 1000 pozorovaniami. Iní (Kuester a kol., 2006) zas použili omnoho menšie okná s 250 pozorovaniami. Pri odhade rizika výskytu extrémnej udalosti s malou pravdepodobnosťou je možné odsúhlasiť, že by mali byť použité najväčšie možné okná, ktoré máme k dispozícii, aby sme zaistili zahrnutie extrémneho pozorovania. Danielsson a De Vries (2000) sa zhodli, že 250 pozorovaní je nedostatočných pre odhad 99% VaR a odporúčajú použitie väčšieho okna, aby sa zaistila vyššia precíznosť odhadov VaR. Na druhej strane, relevantné a synchronne dáta nemusia byť dostupné pre všetky rizikové faktory. Z týchto dvoch protichodných argumentov sme zvolili okná o veľkosti 600, 1000 a 1500 v skúmaných dátových súboroch tak, aby sme jednak spravili akýsi kompromis medzi počtom dát pre odhad jedného VaR a celkovým počtom odhadnutých VaR, a aby sme ukázali všeobecne dobré počínanie metód pri rôznych veľkostiach okien. Z vlastnej skúsenosti neodporúčame využívať okná s veľkosťou menšou ako 500, keďže pri týchto veľkostiach môžu nastávať problémy s prvotným vyrovnaním výnosov AR(1)-GJR-GARCH(1,1) modelmi.

Ďalšou limitáciou tejto štúdie je, že sme dobré výsledky metód založených na EVT porovnávali len s metodológiu RiskMetrics. To ale neznamená, že neexistuje iná metóda, ktorá si počína ešte lepšie v porovnaní so skúmanými metódami. Ďalej, súbor použitých metód je široký, no niekto by mohol skritizovať nezahrnutie iných metód, ktoré používajú EVT. To by síce túto prácu obohatilo, ale aj značne predĺžilo. Naším hlavným cieľom bol ale popis základných metód, pri ktorých pochopení nemá čitateľ problém s použitím prípadných zovšeobecnení ako je napríklad použitie dvojdimenzionálneho Poissonovho procesu v POT metóde (Tsay, 2010, kap. 7.7.3). Ďalej, niekto by mohol skritizovať našu štúdiu kvôli použitiu GJR-GARCH modelu pri modelovaní volatility a nezahrnutí pokročilejších modelov. Avšak náš výber GJR-GARCH modelu bol motivovaný empirickými vlastnosťami finančných výnosov. Snažili sme sa nájsť vhodného kandidáta z triedy GARCH modelov, ktorý by bol schopný zachytiť podmienenú heteroskedasticitu, zhlukovanie volatilit a pákový efekt.

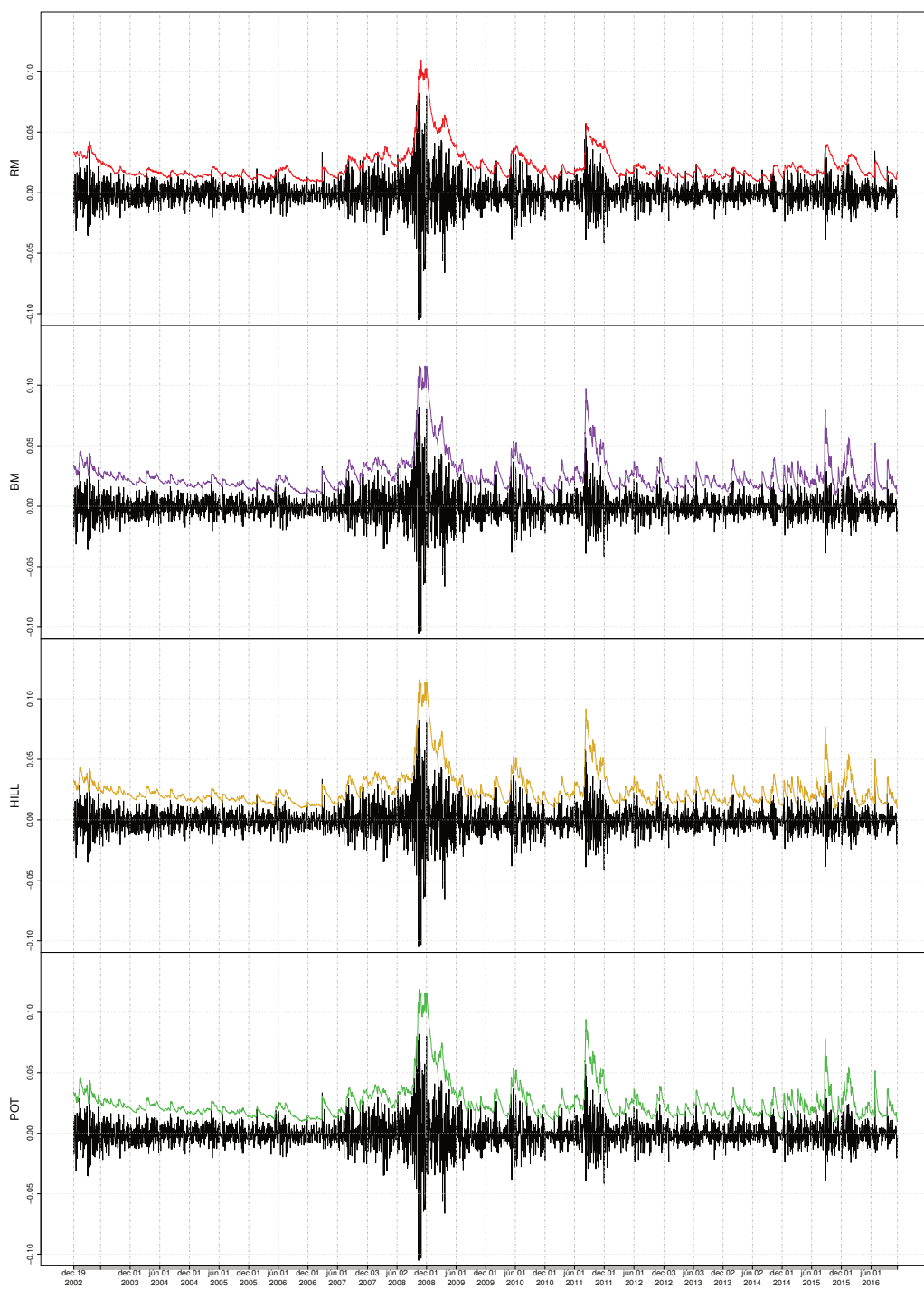
Čo sa týka rozšírenia našej štúdie, z empirického uhla pohľadu by mohla byť rozšírená a doplnená v niekoľkých oblastiach. Po prvé, mohli by sme metódy aplikovať na ceny akcií spoločnosti z iných krajín resp. na ceny ostatných tried aktív, ako sú napríklad opcie, futures, komodity, . . . Ďalej, mohli by sme preskúmať efekt použitia vysokofrekvenčných dát alebo naopak, dát s ešte nižšou frekvenciou ako je týždenná. Použitie dát s nízkou frekvenciou by mohlo pomôcť vyriešiť problém so závislosťou dát, no na druhú stranu je ich dostupnosť u rôznych inštitúcií otázná. Ďalším rozšírením by bolo použitie alternatívnej špecifikácie podmieneného rozptylu. Dobré zhrnutie potencionálnych modelov poskytol Hansen a Lunde (2005). V ich článku môže čitateľ nájsť množstvo iných teoretických štruktúr volatilit. Konečne, počínanie metód by mohlo byť porovnané pomocou iných kritérií ako nami troch použitých. Možná je alternatíva založená na stratovej funkcii, ktorú navrhol Lopez (1998). Z teoretického uhla pohľadu, možná cesta k ďalšiemu výskumu zahŕňa prieskum alternatívnych metód, ktoré by sa dokázali vysporiadať so závislými dátami inak, ako počiatočným očistením alebo za pomoci extrémálneho indexu.



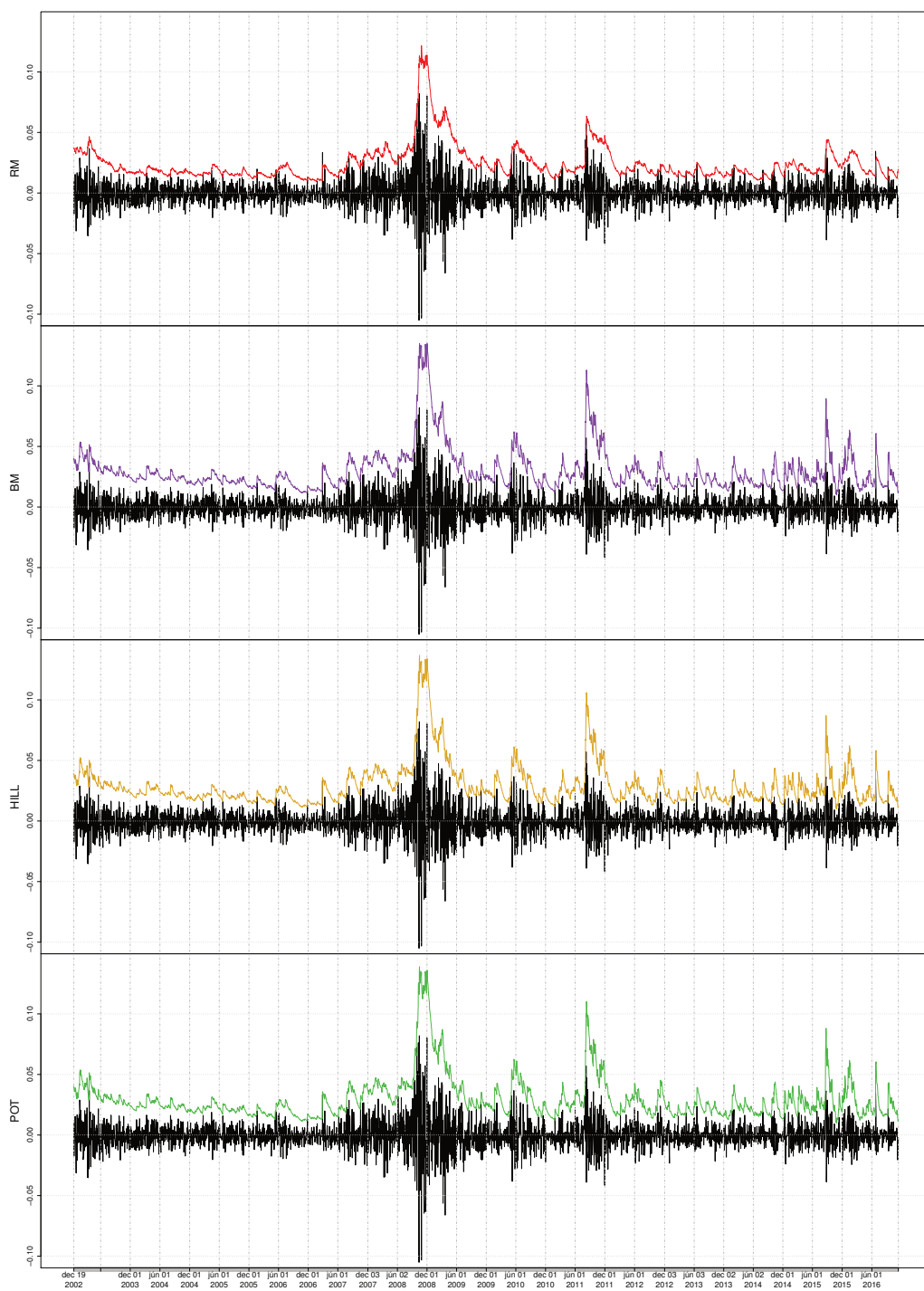
Obr. 6.13: Denné 95% odhady VaR indexu DJI v porovnaní so skutočnými log výnosmi.



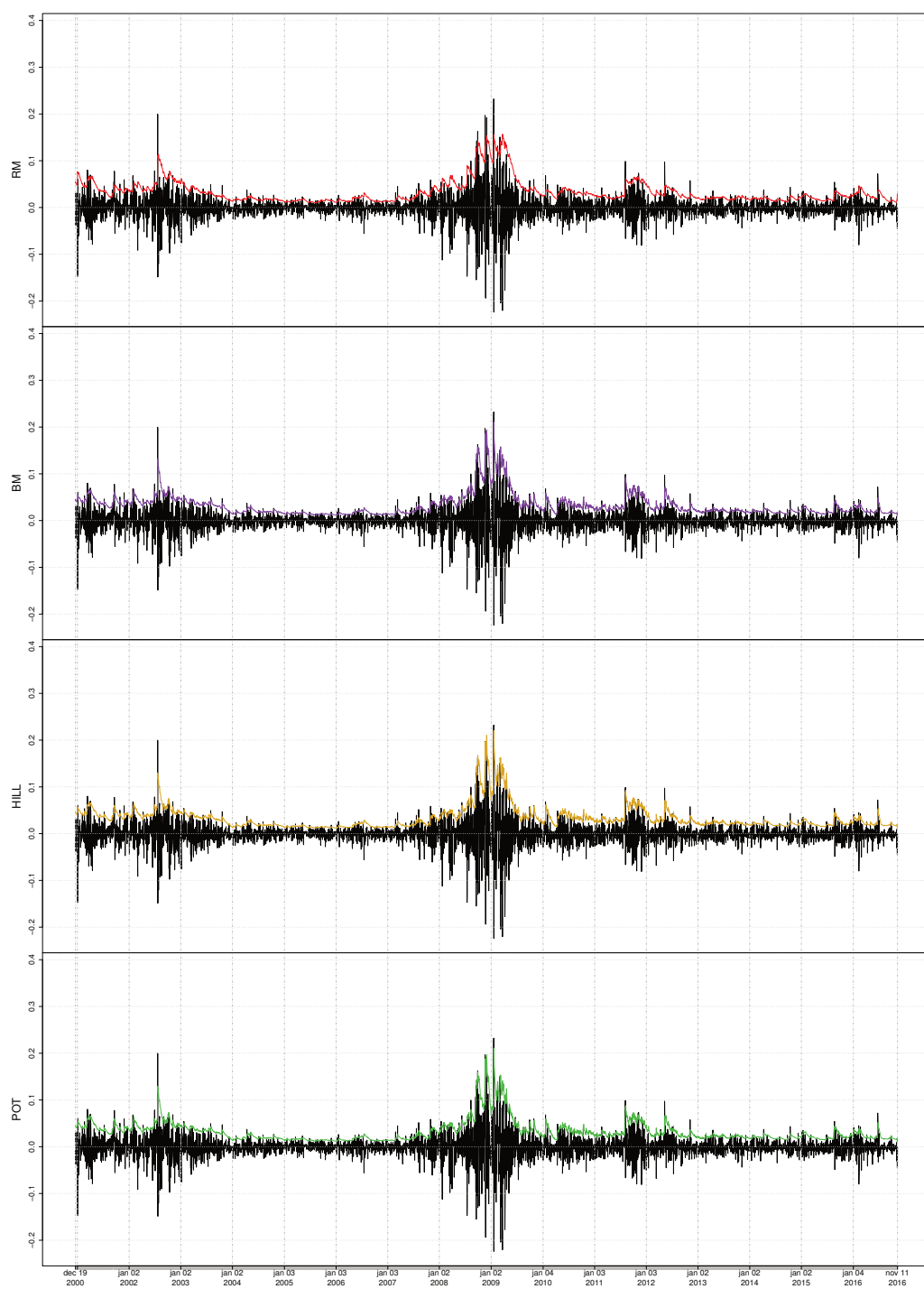
Obr. 6.14: Denné 97,5% odhady VaR indexu DJI v porovnání so skutočnými log výnosmi.



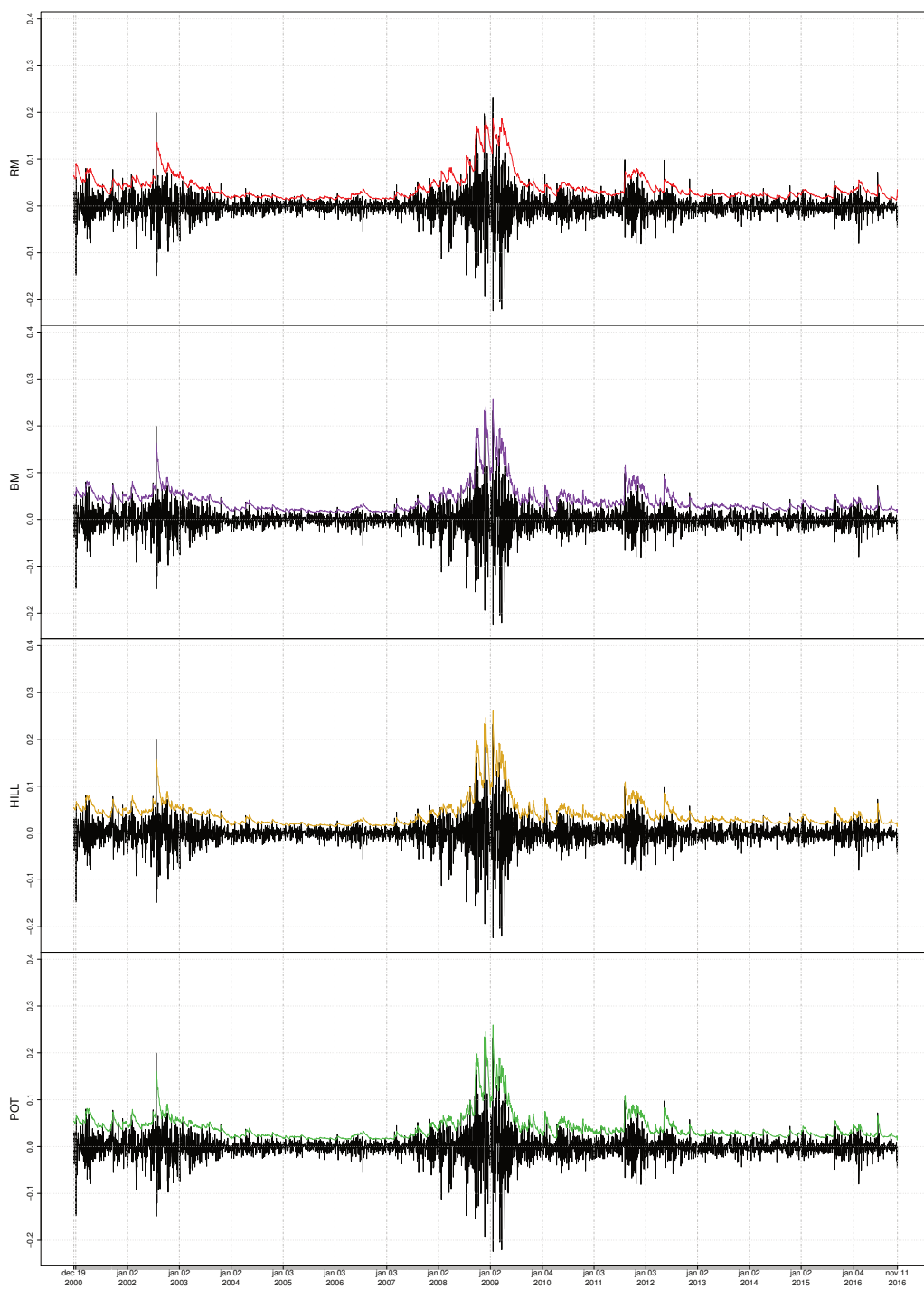
Obr. 6.15: Denné 99% odhady VaR indexu DJI v porovnaní so skutočnými log výnosmi.



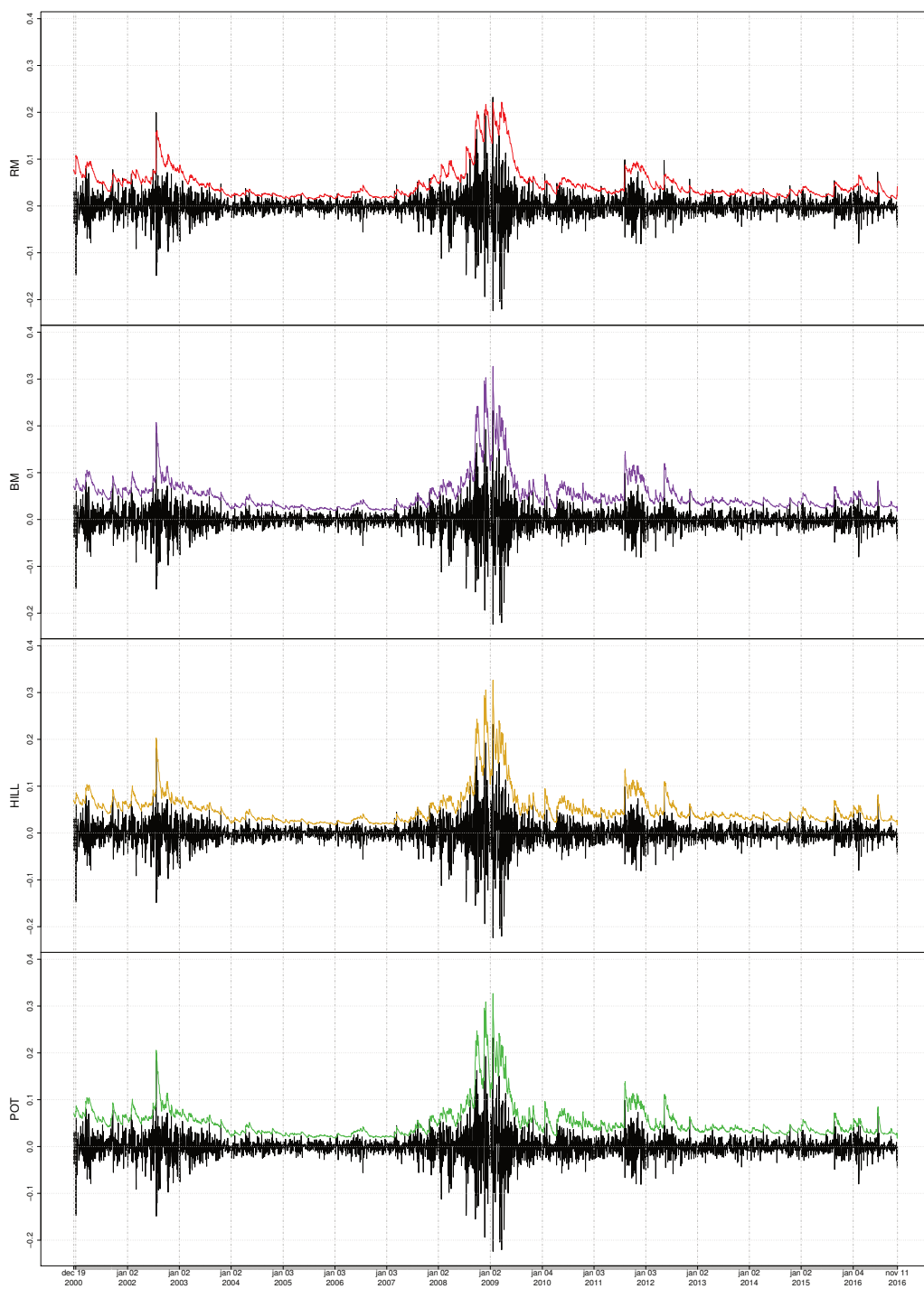
Obr. 6.16: Denné 99,5% odhady VaR indexu DJI v porovnaní so skutočnými log výnosmi.



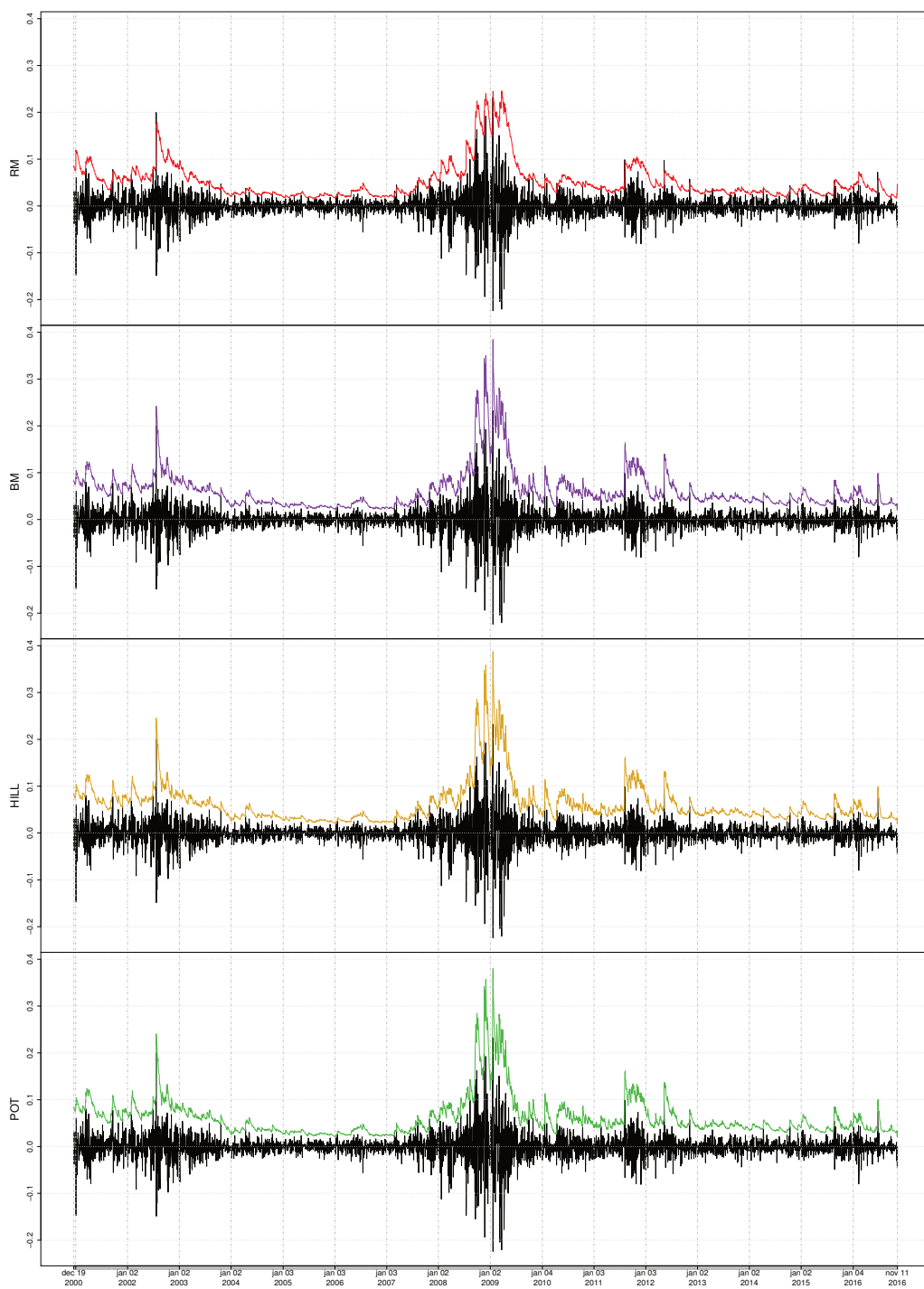
Obr. 6.17: Denné 95% odhady VaR JPMorgan v porovnání so skutočnými log výnosmi.



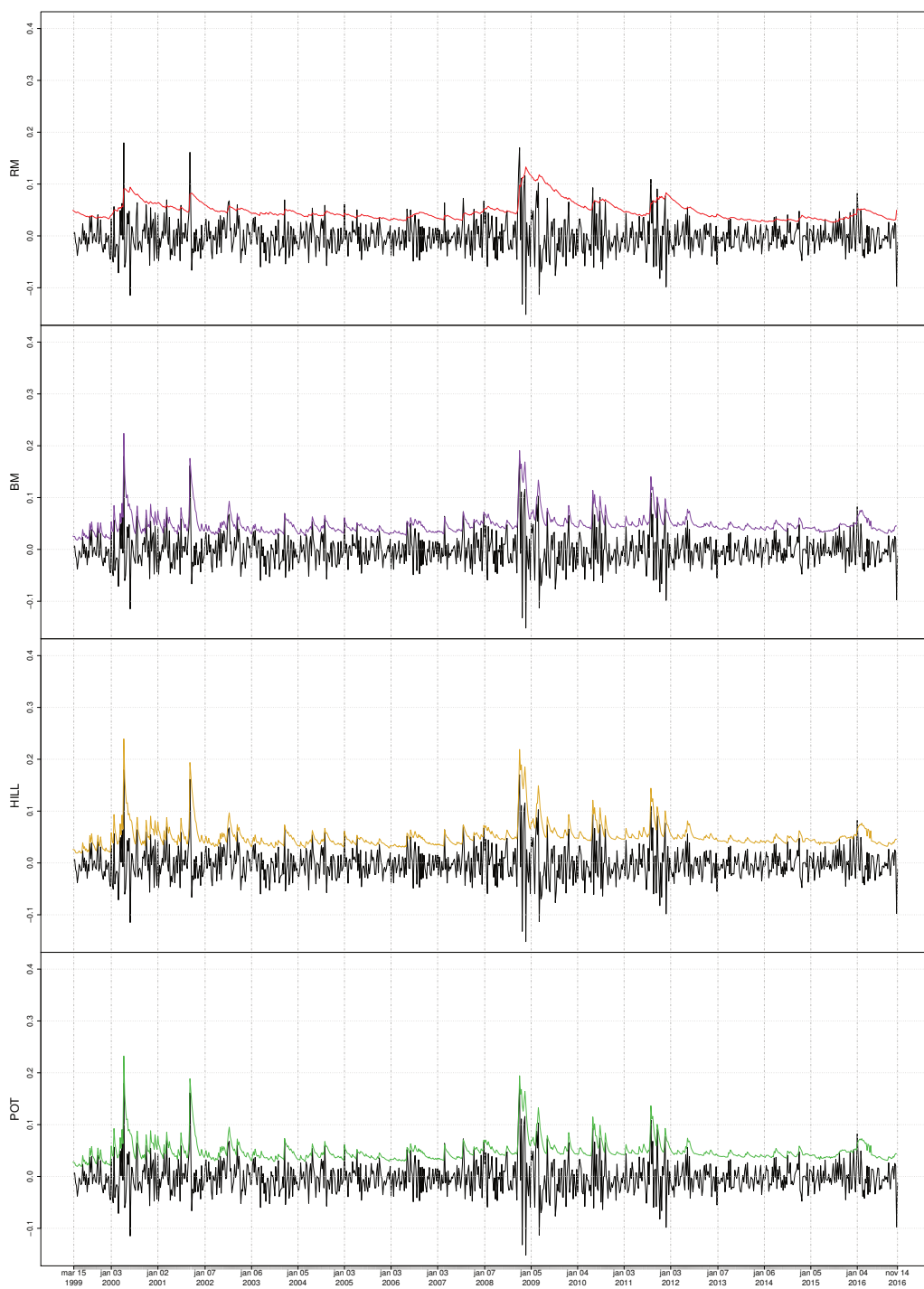
Obr. 6.18: Denné 97,5% odhady VaR JPMorgan v porovnání so skutočnými log výnosmi.



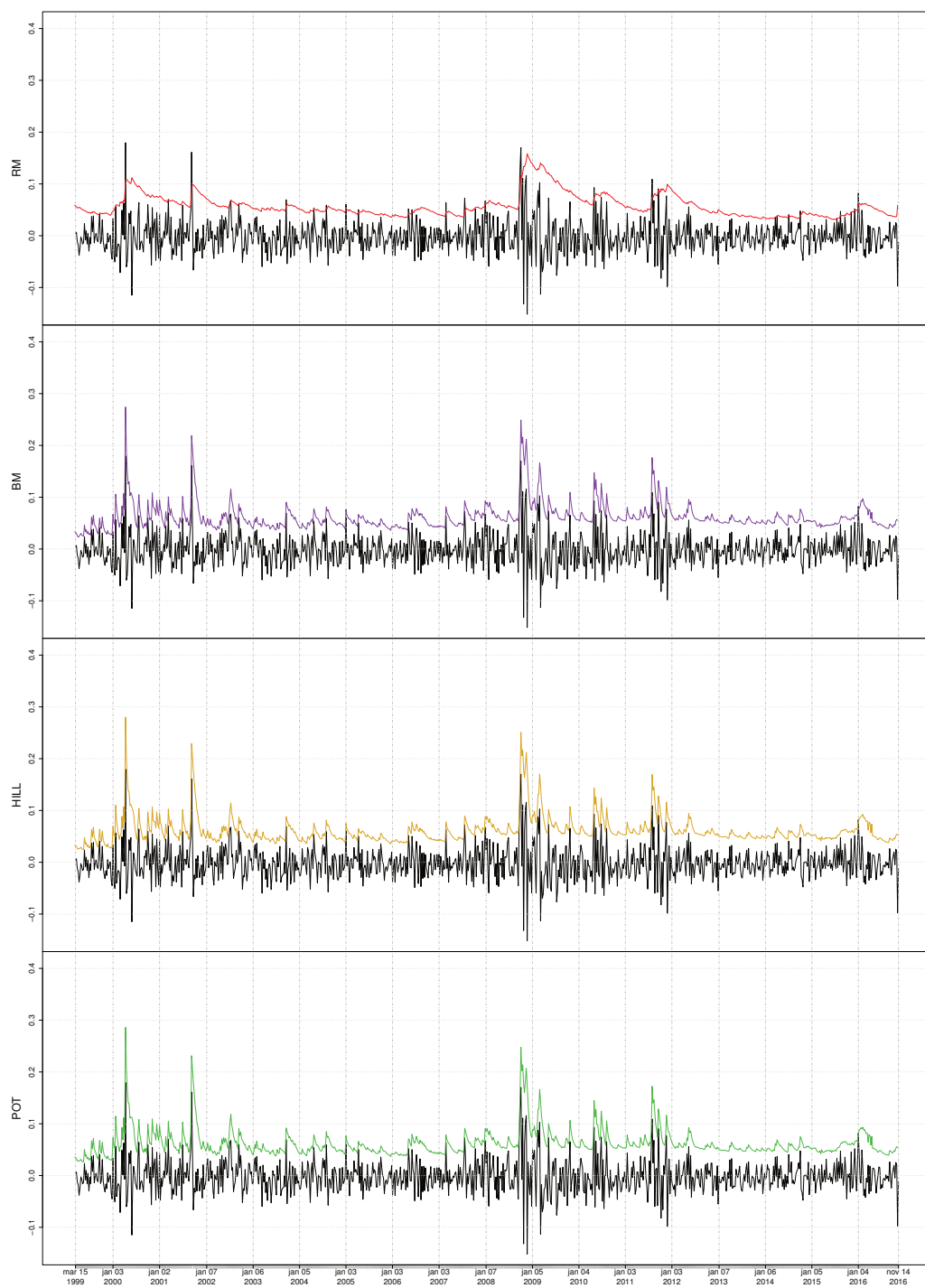
Obr. 6.19: Denné 99% odhady VaR JPMorgan v porovnání so skutočnými log výnosmi.



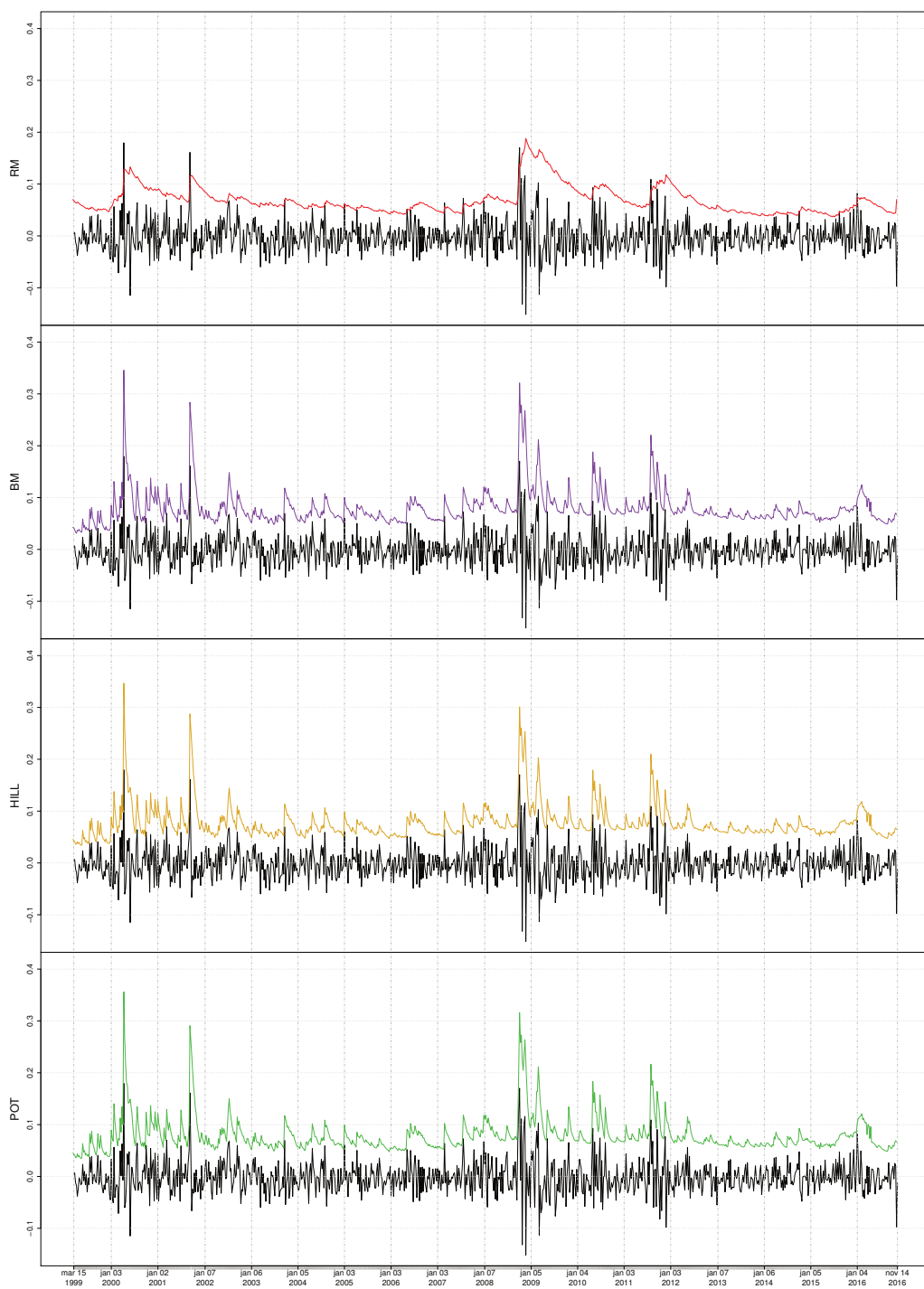
Obr. 6.20: Denné 99,5% odhady VaR JPMorgan v porovnání so skutočnými log výnosmi.



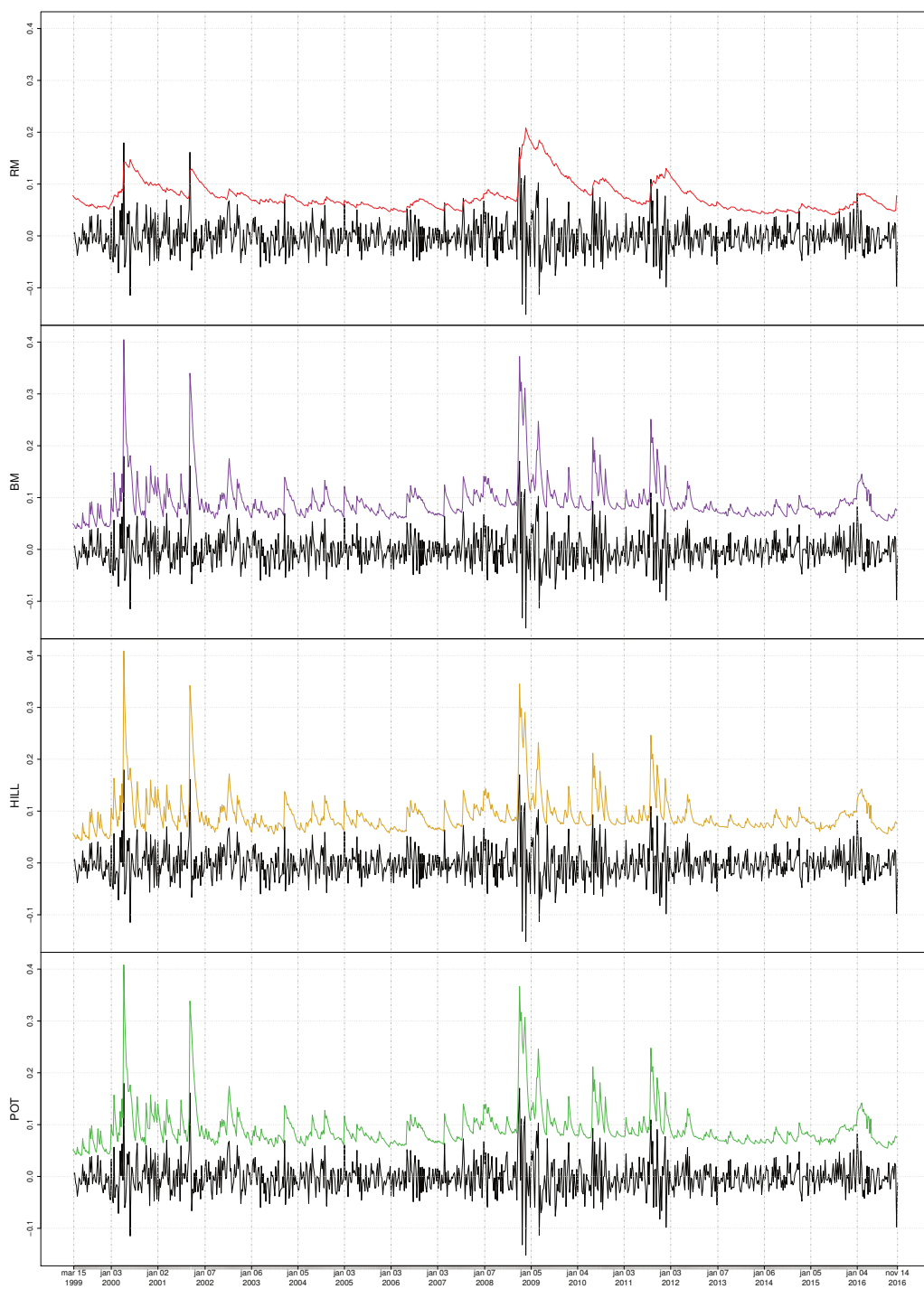
Obr. 6.21: 95% odhady VaR indexu Russell 2000 v porovnání so skutočnými log výnosmi.



Obr. 6.22: 97,5% odhady VaR indexu Russell 2000 v porovnaní so skutočnými log výnosmi.



Obr. 6.23: 99% odhady VaR indexu Russell 2000 v porovnání so skutečnými log výnosmi.



Obr. 6.24: 99,5% odhady VaR indexu Russell 2000 v porovnání so skutočnými log výnosmi.

Záver

Táto diplomová práca sa zaoberala teóriou extrémnych hodnôt (EVT) a jej potenciálom v rizikovom manažmente, špeciálne jej uplatnením pri výpočte známej rizikovej miery VaR. V nasledujúcej časti zhrnieme hlavné zistenia a výsledky, ktoré sme dosiahli.

Fundamentálna časť teoretickej časti EVT je tvorená tzv. Fisherovou-Tippetovou-Gnedenkovou vetou, ktorá opisuje limitné správanie vhodne normalizovaného maxima náhodného výberu. Za podmienky nezávislých a rovnako rozdelených dát, vhodne normalizované maximum konverguje v distribúcii k zobecnenému rozdeleniu extrémnych hodnôt (GEVD), ktoré je tvorené Fréchetovou, Gumbelovou a Weibullovou triedou štandardných rozdelení extrémnych hodnôt. Na týchto teoretických poznatkoch sme stavali teóriu blokových maxím (BM), pri ktorej je maximum priamo modelované GEV rozdelením. Následne sme ukázali neparametrický prístup odhadu VaR pre dáta pochádzajúce z Fréchetovej sféry príťažlivosti označovaný ako Hillov odhad. Posledný prístup bol založený na presahoch cez medzu (POT), ktorý využíva zobecnené Paretovo rozdelenie (GPD) na modelovanie distribučnej funkcie presahu cez medzu.

Aj keď predpoklad iid dát môže byť opodstatnený napr. pri analýze povodní alebo pri iných vedných disciplínach, ktoré využívajú EVT, nezávislosť dát vo finančných aplikáciách, napr. pri kvantifikácii rizika, kde sa extrémne udalosti zvyknú zhľukovať, väčšinou nie je možné predpokladať, keďže zhľukovanie svedčí o dočasnej závislosti v dátach. Praktická časť tejto práce ukázala množstvo prístupov ako sa vysporiadať s týmto porušeným predpokladom nezávislosti dát. Prístup, ktorý sme v praktickej časti aplikovali spočíval v prvotnom očistení dát pred samotnou aplikáciou EVT. Na očistenie dát sme použili jeden z triedy AR-GARCH modelov, presnejšie AR(1)-GJR-GARCH(1,1) model, následne sme EVT aplikovali na štandarizované reziduá a ukázali postup, ako dostať VaR pre pôvodný rad finančných výnosov.

Ukázali sme, že tento prístup predčil tradičnú metodológiu RiskMetrics, hlavne na vyšších konfidenčných hladinách, kde tradičná metóda úplne zlyhala. Metódy sme porovnávali na základe testov podmieneného testu pokrytia, testu nezávislosti a ich združenej verzii. Ukázalo sa, že pri troch skúmaných dátových súboroch, ktoré tvorili denné negatívne logaritmické výnosy akciového indexu Dow Jones Industrial Average, banky JPMorgan a týždenné výnosy akciového indexu Russell 2000 s rôznymi hodnotami vstupných parametrov si metódy vo všeobecnosti počínali skutočne dobre a rozhodne ich odporúčame aplikovať pre výpočet VaR. Každý by však mal pred aplikáciou zvážiť komplexitu metód nad ich výpočetnou presnosťou.

Zoznam použitej literatúry

- ALEXANDER, C. (1996). Evaluating the Use of RiskMetrics as a Risk Measurement Tool for Your Operation: What are Its Advantages and Limitations? *Derivatives: Use Trading and Regulation*, **2**(3), 277–285.
- ARTZNER, P., DELBAEN, F., EBER, J.-M. a HEATH, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, **9**(3), 203–228.
- BALKEMA, A. A. a DE HAAN, L. (1974). Residual Life Time at Great Age. *The Annals of Probability*, **2**(5), 792–804.
- BCBS: FUNDAMENTAL REVIEW OF THE TRADING BOOK: A REVISED MARKET RISK FRAMEWORK (2013). <http://www.bis.org/publ/bcbs265.pdf>, navštívené 24.03.2017.
- BCBS: MINIMUM CAPITAL REQUIREMENTS FOR MARKET RISK (2016). <https://www.bis.org/bcbs/publ/d352.pdf>, navštívené 24.03.2017.
- BLACK, F. (1976). *Studies of Stock Price Volatility Changes*. Proceedings of the 1976 Meeting of Business and Economic Statistics Section. American Statistical Association, Washington DC.
- CHRISTOFFERSEN, P. F. (1998). Evaluating Interval Forecasts. *International Economic Review*, **39**(4), 841–862.
- CIPRA, T. (2008). *Finanční ekonometrie*. 1. vydanie, Ekopress, Praha.
- CIPRA, T. (2015). *Riziko ve financích a pojišťovnictví: Basel III a Solvency II*. Ekopress, Praha.
- CONT, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, **1**(2), 223–236.
- DANIELSSON, J. a DE VRIES, C. G. (1997). *Value at Risk and Extreme Returns*. Financial Markets Group Discussion Paper, no. 273. London School of Economics, Londýn.
- DANIELSSON, J. a DE VRIES, C. G. (2000). Value-at-Risk and Extreme Returns. *Annales d'Economie et de Statistique*, **60**, 239–270.
- DÁTOVÝ SÚBOR DENNÝCH UZATVÁRACÍCH CIEN AKCIÍ JPMORGAN OD 1.1.1997 DO 11.11.2016 (2016). <http://chart.finance.yahoo.com/table.csv?s=JPM&a=0&b=1&c=1997&d=10&e=11&f=2016&g=d&ignore=.csv>, navštívené 24.03.2017.
- DÁTOVÝ SÚBOR DENNÝCH UZATVÁRACÍCH CIEN INDEXU DJI OD 1.1.1997 DO 11.11.2016 (2016). <http://chart.finance.yahoo.com/table.csv?s=~DJI&a=0&b=1&c=1997&d=10&e=11&f=2016&g=d&ignore=.csv>, navštívené 24.03.2017.

- DÁTOVÝ SÚBOR TÝŽDENNÝCH UZATVÁRACÍCH CIEN INDEXU RUT OD 10.9.1987 DO 14.11.2016 (2016). <http://chart.finance.yahoo.com/table.csv?s=RUT&a=8&b=10&c=1987&d=10&e=14&f=2016&g=w&ignore=.csv>, navštívené 24.03.2017.
- EMBRECHTS, P., KLÜPPELBERG, C. a MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. 4. vydanie, Springer, Berlín.
- ENGLE, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, **50**(4), 987–1008.
- FAMA, E. F. (1965). The Behavior of Stock-Market Prices. *Journal of Business*, **38**(1), 34–105.
- FERRO, C. A. T. a SEGERS, J. (2003). Inference for clusters of extreme values. *J.R. Statist. Soc. B*, **65**(2), 545–556.
- FISHER, R. A. a TIPPETT, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge philosophical society*, **24**(2), 180–190.
- FRANKE, J., HÄRDLE, W. a HAFNER, C. M. (2004). *Statistics of Financial Markets: An Introduction*. Springer, Berlín.
- GHALANOS, A. (2015). *rugarch: Univariate GARCH models*. R package version 1.3-6.
- GNEDENKO, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, **44**(3), 423–453.
- HANSEN, P. R. a LUNDE, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics*, **20**(7), 873–889.
- HILL, B. M. (1975). A Simple General Approach to Inference About the Tail of a Distribution. *Annals of Statistics*, **3**(5), 1163–1174.
- JARQUE, C. M. a BERA, A. K. (1980). Efficient test for normality, homoscedasticity and serial independence of regression residuals. *Economic Letters*, **6**(3), 255–259.
- JENKINSON, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, **81**(348), 158–171.
- KUESTER, K., MITTNIK, S. a PAOLELLA, M. S. (2006). Value-at-Risk Prediction: A Comparison of Alternative Strategies. *Journal of Financial Econometrics*, **4**(1), 53–89.
- LAURIDSEN, S. (2000). Estimation of Value at Risk by Extreme Value Methods. *Extremes*, **3**(2), 107–144.

- LEADBETTER, M. R. (1983). Extremes and local dependence in stationary sequences. *Probability Theory and Related Fields*, **65**(2), 291–306.
- LJUNG, G. M. a BOX, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**(2), 297–303.
- LOPEZ, J. A. (1998). Testing your risk tests. *The Financial Survey*, **20**(3), 18–20.
- MARIMOUTOU, V., RAGGAD, B. a TRABELSI, A. (2009). Extreme Value Theory and Value at Risk: Application to oil market. *Energy Economics*, **31**(4), 519–530.
- MCNEIL, A. J. a FREY, R. (2000). Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach. *Journal of empirical finance*, **7**(3), 271–300.
- MCNEIL, A. J., FREY, R. a EMBRECHTS, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton.
- MORGAN, J. (1996). *RiskMetrics™–Technical Document*. 4. vydanie, Morgan Guaranty Trust Company of New York, New York.
- O'BRIEN, G. L. (1987). Extreme Values for Stationary and Markov Sequences. *Annals of Probability*, **15**(1), 281–291.
- RSTUDIO TEAM (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.
- SMITH, R. L. a WEISSMAN, I. (1994). Estimating the extremal index. *Journal of the Royal Statistical Society*, **56**(3), 515–528.
- SOLVENCY II: DIREKTÍVA OMNIBUS II (2015). <http://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=OJ:L:2015:012:FULL&from=CS>, navštívené 24.03.2017.
- TSAY, R. S. (2010). *Analysis of Financial Time Series*. 3. vydanie, Wiley, New York.
- VOIGT, T. (2015). Replacing VaR with Expected Shortfall in market risk. <https://www.finextra.com/blogs/fullblog.aspx?blogid=11678>, navštívené 24.03.2017.
- WOLFRAM RESEARCH, I. (2015). *Mathematica*. Wolfram Research, Inc., Champaign, Illinois, verzia 10. 1.

Zoznam obrázkov

1.1	95% VaR Gumbelovho rozdelenia.	8
2.1	Porovnanie distribučných funkcií leptokurtického štandarizovaného t -rozdelenia so 7 stupňami voľnosti a štandardného normálneho rozdelenia.	12
3.1	Distribučné funkcie a hustoty GEV rozdelení.	15
6.1	Uzatváracie ceny a negatívne logaritmičné výnosy akciového indexu DJI.	35
6.2	Histogram log výnosov okna indexu DJI v porovnaní s normálnym rozdelením so strednou hodnotou a rozptylom okna.	35
6.3	Korelogramy log výnosov okna a ich štvorcov spoločne s korelogram log výnosov celkového výberu a ich štvorcov v prípade indexu DJI.	36
6.4	Graf štandarizovaných reziduí z modelu 6.3 spolu s korelogramom reziduí, štvorcov reziduí a QQ grafmi.	39
6.5	Hillov graf okna v prípade indexu DJI.	40
6.6	Graf stredného presahu cez medzu η v prípade indexu DJI pre celé okno reziduí a detailnejší pohľad pre reziduí z intervalu $(0,2)$. . .	41
6.7	Graf distribučnej funkcie presahu cez medzu η v prípade indexu DJI spolu s vyrovnanými hodnotami, QQ graf empirických kvantilov vs. kvantilov exponenciálneho rozdelenia.	42
6.8	Vývoj koeficientov AR(1)-GJR-GARCH(1,1) modelov v čase v prípade indexu DJI.	44
6.9	Uzatváracie ceny očistené od dividend a negatívne logaritmičné výnosy akcií JPM.	47
6.10	Vývoj koeficientov AR(1)-GJR-GARCH(1,1) modelov v čase v prípade akcií JPM.	47
6.11	Uzatváracie ceny očistené od dividend a negatívne logaritmičné výnosy akciového indexu RUT.	50
6.12	Vývoj koeficientov AR(1)-GJR-GARCH(1,1) modelov v čase v prípade indexu RUT.	51
6.13	Denné 95% odhady VaR indexu DJI v porovnaní so skutočnými log výnosmi.	54
6.14	Denné 97,5% odhady VaR indexu DJI v porovnaní so skutočnými log výnosmi.	55
6.15	Denné 99% odhady VaR indexu DJI v porovnaní so skutočnými log výnosmi.	56
6.16	Denné 99,5% odhady VaR indexu DJI v porovnaní so skutočnými log výnosmi.	57
6.17	Denné 95% odhady VaR JPMorgan v porovnaní so skutočnými log výnosmi.	58
6.18	Denné 97,5% odhady VaR JPMorgan v porovnaní so skutočnými log výnosmi.	59

6.19	Denné 99% odhady VaR JPMorgan v porovnaní so skutočnými log výnosmi.	60
6.20	Denné 99,5% odhady VaR JPMorgan v porovnaní so skutočnými log výnosmi.	61
6.21	95% odhady VaR indexu Russell 2000 v porovnaní so skutočnými log výnosmi.	62
6.22	97,5% odhady VaR indexu Russell 2000 v porovnaní so skutočnými log výnosmi.	63
6.23	99% odhady VaR indexu Russell 2000 v porovnaní so skutočnými log výnosmi.	64
6.24	99,5% odhady VaR indexu Russell 2000 v porovnaní so skutočnými log výnosmi.	65

Zoznam tabuliek

6.1	Deskriptívne štatistiky okna a celého výberu log výnosov indexu DJI.	34
6.2	Hodnoty testových štatistík a príslušné p-hodnoty JB testu, LB testu pre pôvodné výbery a štvorcové výbery spolu s testom na dodatočný ARCH efekt pre okno a celkový výber v prípade indexu DJI.	36
6.3	Odhady koeficientov v modeli AR(1)-GJR-GARCH(1,1) v prípade okna indexu DJI.	37
6.4	Odhady jednodenných VaR v percentách pre 1501. obchodný deň metódami RiskMetrics, BM metódou, Hillovou metódou, POT metódou v prípade indexu DJI.	43
6.5	Výsledky spätného testovania v prípade akciového indexu DJI. . .	45
6.6	Deskriptívne štatistiky negatívnych logaritmických výnosov a cien akcií JPM.	46
6.7	Výsledky spätného testovania v prípade akcií JPM.	48
6.8	Deskriptívne štatistiky negatívnych logaritmických výnosov a cien akciového indexu RUT.	50
6.9	Výsledky spätného testovania v prípade akciového indexu RUT. .	51

Prílohy

1. *DJI.csv*: denné ceny akcií akciového indexu Dow Jones Industrial Average od 1.1.1997 do 11.11.2016. Jednotlivé čísla v tabuľke (obsahuje záhlavie) sú oddelené čiarkami a reprezentujú postupne dátum, otváraciu, najvyššiu, najnižšiu, uzatváraciu cenu, objem akcií a konečne, uzatváraciu cenu akcií očistenú od dividend, ktorá bola použitá na analýzu.
2. *JPM.csv*: denné ceny akcií banky JPMorgan od 1.1.1997 do 11.11.2016 s rovnakou štruktúrou ako dáta z *DJI.csv*.
3. *RUT.csv*: týždenné ceny akcií akciového indexu Russell 2000 od 10.9.1987 do 14.11.2016 s rovnakou štruktúrou ako dáta z *DJI.csv*.
4. *zdroj.R*: zdrojový kód usporiadaný na kompletnú analýzu ľubovoľného dátového súboru prevedenú v kapitole 6, vrátane grafických výstupov, vlastného nastavenia parametrov, hlavnej procedúry, ktorá prevádza zautomatizovaný odhad VaR každou zo skúmaných metód spoločne s grafmi a funkciou, ktorá prevádza spätné testovanie. Hodnoty parametrov sú v zdrojovom kóde nastavené primárne pre dátový súbor *DJI.csv*.