

UNIVERZITA KOMENSKÉHO V BRATISLAVE

FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

**ZOVŠEOBECNENÉ LAMBDA ROZDELENIE
PRI MODELOVANÍ FINANČNÝCH AKTÍV**

Diplomová práca

UNIVERZITA KOMENSKÉHO V BRATISLAVE

FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



**ZOVŠEOBECNENÉ LAMBDA ROZDELENIE
PRI MODELOVANÍ FINANČNÝCH AKTÍV**

Diplomová práca

Študijný program: Ekonomická a finančná matematika

Študijný odbor: 9.1.9 Aplikovaná matematika

Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky

Vedúci diplomovej práce: Mgr. Ing. Urban Kováč, PhD

Bratislava 2013

Bc. Slavomíra Gregušová



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Slavomíra Gregušová
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský

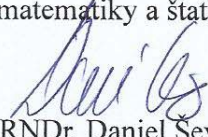
Názov: Zovšeobecnené lambda rozdelenie pri modelovaní finančných aktív

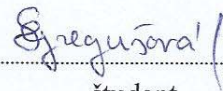
Cieľ: Hlavnou náplňou diplomovky je aplikácia zovšeobecneného lambda rozdelenia na popis distribúcie výnosov jednotlivých aktív na trhu. Zameriame sa na generovanie aktív v portfóliu pomocou viacrozmerného zovšeobecneného lambda rozdelenia, kde zohľadníme aj koreláciu medzi aktívami. Výsledky použijeme v rôznych finančných modeloch na predikciu alebo optimalizáciu výnosov portfólia.

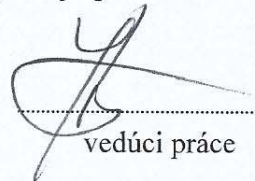
Vedúci: Mgr. Urban Kováč
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky

Dátum zadania: 25.01.2012

Dátum schválenia: 26.01.2012


prof. RNDr. Daniel Ševčovič, CSc.
garant študijného programu


študent


vedúci práce

Prehlásenie:

Čestne prehlasujem, že túto prácu som vypracovala samostatne pod vedením Mgr. Ing. Urbana Kováča, PhD a uviedla v zozname literatúry všetky použité zdroje.

V Bratislave, 22.4.2013

.....

Podpis autora

Pod'akovanie:

Chcela by som pod'akovať vedúcemu svojej diplomovej práce Mgr. Ing. Urbanovi Kováčovi, PhD za pomoc pri vypracovávaní, za užitočné rady a pripomienky. Ďalej by som rada pod'akovala priateľovi, rodine a kamarátom za trpezlivosť a podporu pri písaní diplomovej práce.

Abstrakt

Gregušová, Slavomíra: Zovšeobecnené lambda rozdelenie pri modelovaní finančných aktív [Diplomová práca].

Univerzita Komenského v Bratislave, Mlynská dolina 84248, Bratislava, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky.

Vedúci práce: Mgr. Ing. Urban Kováč, PhD

Bratislava, 2013. /55 s./

Hlavnou náplňou diplomovej práce bude aplikácia zovšeobecneného lambda rozdelenia na popis distribúcie výnosov jednotlivých aktív na trhu. Zameriame sa na generovanie aktív v portfóliu pomocou viacrozmerného zovšeobecneného lambda rozdelenia, pričom zohľadníme korelácie medzi jednotlivými aktívami. Výsledky použijeme vo finančných modeloch na predikciu miery rizika (VaR) v portfóliu.

Kľúčové slová: Zovšeobecnené lambda rozdelenie, Generalized Lambda Distribution, VaR, GARCH, Monte Carlo

Abstract

Gregušová, Slavomíra: Generalized lambda distribution for modeling of financial assets [Master's thesis].

Comenius University Bratislava, Mlynská dolina 84248, Bratislava, Faculty of Mathematics, Physics, and Informatics, Department of Applied Mathematics and Statistics.

Supervisor: Mgr. Ing. Urban Kováč, PhD

Bratislava, 2013. /55 p./

The main content of this thesis is the application of the Generalized Lambda Distribution to describe the distribution of the asset returns in the market. We focus on generating assets in the portfolio using a Multidimensional Generalized Lambda Distribution, taking into account the correlation between assets. The results are used in the financial model to predict the Value-at-Risk measure (VaR) for the given portfolio.

Keywords: Generalized Lambda Distribution, VaR, GARCH, Monte Carlo

Obsah

1	Základné pojmy	3
1.1	Distribučná funkcia	3
1.2	Hustota rozdelenia pravdepodobnosti	3
1.3	Kvantilová funkcia	4
2	Zovšeobecnené lambda rozdelenie (GLD)	6
2.1	Definícia zovšeobecneného lambda rozdelenia [1]	6
2.2	Priestor parametrov GLD [1]	8
2.3	Tvary funkcie hustoty GLD [1]	10
3	Aplikácia GLD na popis rozdelenia dát použitím metódy momentov	14
3.1	Momenty zovšeobecneného lambda rozdelenia [1]	14
3.2	Priestor (α_3^2, α_4) [1]	16
3.3	Odhad GLD parametrov pomocou metódy momentov [1]	17
3.4	Použitie GLD na aproximáciu normálneho rozdelenia [1]	18
3.4.1	Normálne rozdelenie	19
3.5	Testy dobrej zhody [1]	20
3.5.1	Pearson-ov χ^2 test	20
3.6	Odhad parametrov GLD pre reálne dáta [16]	21
3.7	Riešenie nelineárneho systému rovníc [4],[10], [13]	26
3.7.1	Newtonova metóda: [13]	27
3.7.2	Broyden metóda:	28
4	Rozšírený GLD systém - EGLD	29
4.1	Beta rozdelenie a jej momenty [1]	29
4.1.1	Vlastnosti	29
4.2	GBD a jej momenty [1]	30
4.3	Momenty GBD [1]	31
4.4	Odhad parametrov GBD [1]	31
5	Odhad Value at Risk (VaR) pomocou GLD	32
5.1	Value-at-Risk	32
5.2	Výpočet Value-at-Risk	32

5.2.1	Metóda historickej simulácie	34
5.2.2	Metóda Monte Carlo	35
5.2.3	GARCH modely [3], [8], [12]	40
5.3	Spätné testovanie VaR - Backtesting	43
6	Praktická časť výpočtu VaR	45
6.1	Výpočet VaR pomocou metódy Monte Carlo	45
6.1.1	Výpočet VaR pre jedno aktívum	45
6.1.2	Výpočet VaR pre portfólio akcií	48
6.2	VaR pomocou GARCH modelov	50

Úvod

Skúmanie správania finančných časových radov je aktuálnou témou vzhľadom na neočakávané extrémne udalosti pohybov cien na trhu. Nedávna celosvetová finančná kríza v roku 2008, a takisto pád trhu v októbri 1987 a mnoho ďalších sú vecné spúšťače k ďalšiemu výskumu a hľadaniu účinnejších predikčných nástrojov. Všetky tieto krízy na trhu boli charakteristické zvýšením volatility. Volatilita v oblasti financií charakterizuje turbulencie výnosov v čase. Teda záujmom je veľmi presný odhad tejto volatility a jej následné predikcie do budúcnosti.

Volatilita je stredobodom pozornosti v oblasti riadenia rizík a následných opatrení pre vyhodnotenie rizík. Najrozšírenejším a najpopulárnejším nástrojom v oblasti riadenia rizík je Value at Risk (VaR), teda miera rizika. VaR vyjadruje maximálnu možnú stratu portfólia s istou pravdepodobnosťou v priebehu istého časového okamihu. Pre výpočet VaR je dôležité poznať pravdepodobnostné rozdelenie finančného nástroja. V prvej časti sa preto zameriame na modelovanie distribúcie výnosov finančných aktív. Rozdelenie výnosov finančných aktív na trhu sa vyznačuje tým, že má ťažké chvosty, resp. pravdepodobnosť väčšieho výnosu/straty nie je nulová, ako napríklad predpokladá normálne rozdelenie - teda rozdelenie je charakteristické väčšou hodnotou špicatosti. Takisto rozdelenie výnosov nie je symetrické ako predpokladá normálne rozdelenie, ale je charakteristické istou šikmosťou rôznou od nuly. V našej diplomovej práci budeme na popis distribúcie používať rozdelenie, ktoré dokáže zachytiť aj takéto vlastnosti, a to *zovšeobecnené lambda rozdelenie* (*Generalized Lambda Distribution - GLD*). Je to štvorparametrové zovšeobecnenie Tukey-ho jednoparametrového lambda rozdelenia, ktoré v sebe zahŕňa širokú škálu rôznych tvarov hustôt a dokáže veľmi dobre aproximovať normálne rozdelenie a ďalšie iné známe rozdelenia.

Diplomová práca je rozdelená do šiestich kapitol. V prvej kapitole spomenieme základné pojmy zo štatistiky, ktoré budú potrebné pre ďalšiu prácu. V kapitole druhej vyslovíme definíciu GLD, predstavíme jeho štyri parametre a ich vlastnosti a ukážeme rôzne tvary, aké môže toto rozdelenie nadobúdať. Aplikácia GLD rozdelenia na popis distribúcie dát použitím metódy momentov bude predmetom tretej kapitoly. V tejto kapitole ukážeme, aký tvar majú predpisy pre momenty GLD rozdelenia a ako pomocou momentovej metódy odhadnúť jednotlivé parametre rozdelenia. Ďalším bodom tejto

časti bude aproximácia normálneho rozdelenia pomocou GLD rozdelenia, následne testy dobrej zhody a aplikácia odhadu parametrov GLD pre reálne výnosy akcie Hewlett Packard (HPQ). Záver tejto kapitoly bude venovaný riešeniu systému nelineárnych rovníc, ktoré sme v tejto kapitole využili. Kapitola štvrtá bude venovaná rozšírenému GLD rozdeleniu tzv. EGLD, ktoré vznikne pridaním skupiny zovšeobecneného beta rozdelenia, pretože samotné GLD nemožno použiť vždy. V piatej časti sa zameriame na teóriu ohľadom spomínaného nástroja na meranie rizika - Value at Risk. Na výpočet VaR existuje niekoľko možností, v našej práci sme sa zamerali na dve rôzne metódy - Metóda Monte Carlo a použitie GARCH modelov. Pre porovnanie "GLD metód" budeme používať normálne rozdelenie. V poslednej šiestej kapitole predstavíme praktické výsledky výpočtu VaR popísaného v kapitole piatej.

1 Základné pojmy

Na začiatok, pred definovaním samotného zovšeobecneného lambda rozdelenia, vysvetlíme základné pojmy zo štatistiky, s ktorými sa budeme v nasledujúcich častiach stretávať. V tejto kapitole budeme vychádzať z [1], [5].

Vo viacerých oblastiach ako sú veda, technika, medicína, inžinierstvo, manažment, atď. sa stretávame s problémom skonštruovať štatistický model na popis základných premenných danej oblasti. Najzákladnejší a najviac používaný model, tzv. rozdelenie pravdepodobnosti, dáva do súvislosti základné premenné a pravdepodobnosť ich výskytu.

Dôležitým pojmom v teórii pravdepodobnosti a štatistiky je pojem *náhodná premenná*. Náhodná premenná je premenná, ktorej hodnota je jednoznačne určená výsledkom náhodného pokusu. Za náhodný pokus považujeme takú činnosť, ktorá sa niekoľkokrát opakuje a ktorej výsledok je neistý, závislý od náhody.

1.1 Distribučná funkcia

Poznať náhodnú premennú znamená vedieť jej pravdepodobnostné rozdelenie, t.j. vedieť aké hodnoty a s akými pravdepodobnosťami nadobúda. Jednou z možností ako charakterizovať pravdepodobné správanie sa náhodnej premennej je distribučná funkcia (Distribution Function).

Definícia 1.1. *Distribučnou funkciou náhodnej premennej X nazývame reálnu funkciu $F : \mathbb{R} \rightarrow \langle 0, 1 \rangle$ definovanú pre každé $x \in \mathbb{R}$ vzťahom*

$$F(x) = P(X < x)$$

(T.j. pravdepodobnosť, že náhodná premenná X nadobudne hodnotu menšiu, nanajvýš rovnú ako reálne číslo x).

1.2 Hustota rozdelenia pravdepodobnosti

Druhým spôsobom, ako špecifikovať šance výskytu rôznych hodnôt náhodnej premennej X je hustota rozdelenia pravdepodobnosti (Probability Density Function - p.d.f.).

Funkcia hustoty (p.d.f.) je funkcia $f_X(x)$, pre ktorú platí:

$$F(x) = \int_{-\infty}^x f(t)dt \text{ a zároveň } \int_{-\infty}^{\infty} f(t)dt = 1,$$

Funkcia hustoty náhodnej veličiny vyjadruje pravdepodobnosť, že premenná bude mať hodnotu v nejakom intervale. Teda pravdepodobnosť, že spojitá náhodná premenná nadobudne hodnotu z intervalu $[a,b]$, je daná obsahom plochy pod krivkou funkcie hustoty na intervale $[a,b]$.

1.3 Kvantilová funkcia

Treťou možnosťou ako určiť šance výskytu rôznych hodnôt náhodnej premennej X je kvantilová funkcia (Quantile Function/Percentile function). Kvantilová funkcia je inverzná funkcia k distribučnej funkcii:

$$Q_X(y) = \{ \text{hodnota takého } x : F_X(x) = y \}, \text{ pre všetky } y \in (0, 1)$$

Hodnoty kvantilovej funkcie $Q(y)$ sa nazývajú kvantilmi. Kvantilová funkcia $x = F^{-1}(y)$ teda udáva, pre aké x bude výsledok náhodného pokusu s požadovanou pravdepodobnosťou y menší alebo rovný x .

Vidíme teda, že existujú tri prístupy ako špecifikovať šance výskytu náhodnej premennej a to: kumulatívna distribučná funkcia (c.d.f.), funkcia hustoty (p.d.f.) a percentilová funkcia (p.f.).

Príklad: Normálne rozdelenie

Normálne rozdelenie je jedno z najviac používaných rozdelení na popis spojitých náhodných premenných. Je charakterizované tým, že je symetrické a namerané hodnoty sú koncentrované skôr v strede ako na chvostoch (nemá ťažké chvosty). Tvar normálneho rozdelenia je špecifikovaný iba pomocou dvoch parametrov: strednej hodnoty μ a štandardnej odchýlky σ .

Funkcia hustoty normálneho rozdelenia so strednou hodnotou μ a disperziou σ , $N(\mu, \sigma)$ má tvar:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Percentilovú funkciu možno získať nasledovne:

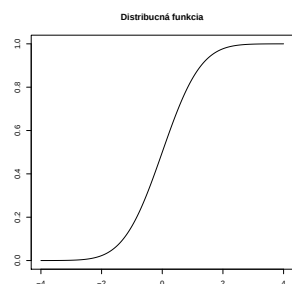
$$F_X(x) = P(X \leq x) = y, \text{ kde } F_X(.) \text{ je distribučná funkcia}$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = y = P\left(Z \leq \frac{x - \mu}{\sigma}\right)$$

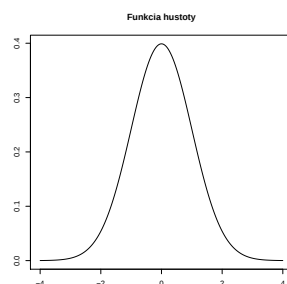
$$Q_Z(y) = \frac{x - \mu}{\sigma}, \text{ odtiaľ } x = \mu + \sigma Q_Z(y)$$

Potom: $Q_X(y) = x$, teda hodnota x , pre ktoré $F_X(x) = y$ sa rovná $\mu + \sigma Q_Z(y)$.

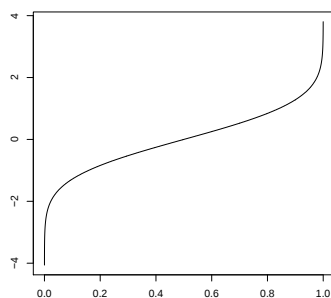
Na nasledujúcich obrázkoch sú grafy spomínaných funkcií - distribučnej funkcie, funkcie hustoty a kvantilovej funkcie normálneho rozdelenia so strednou hodnotou 0 a disperziou 1 (značíme $N(0,1)$).



Obr. 1.3.1: Distribučná funkcia



Obr. 1.3.2: Funkcia hustoty



Obr. 1.3.3: Kvantilová/percentilová funkcia

2 Zovšeobecnené lambda rozdelenie (GLD)

Hlavnou časťou diplomovej práce bude otázka, ako modelovať spojité rozdelenie pravdepodobnosti pre danú vzorku dát. V nasledujúcich troch kapitolách sa budeme venovať práve tomuto problému pomocou zovšeobecneného lambda rozdelenia. V tejto kapitole začneme tým, že uvedieme definíciu spomínaného rozdelenia, predstavíme si jeho štyri parametre a ich vlastnosti. Na záver kapitoly ukážeme rôzne tvary, aké môže toto rozdelenie nadobúdať. Budeme vychádzať z [1], [11].

Modelovaniu distribúcie dát bola venovaná v poslednej dobe veľká pozornosť a tento obor zaznamenal vysoký pokrok. Často najnovšie poznatky vyriešili predchádzajúce problémy s procesom aproximácie. Jedným z mnoho nástrojov na modelovanie rozdelenia dát je zovšeobecnené lambda rozdelenie. Zovšeobecnené lambda rozdelenie (Generalized lambda distribution - GLD), pôvodne navrhnuté Rambergom a Schmeisnerom (1974) je štvor-parametrové zovšeobecnenie Tukey-ho jednoparametrového lambda rozdelenia predstaveného Hastings-om v 1947. Vďaka flexibilitě GLD, berúc do úvahy jeho širokú škálu tvarov, sa začalo značne používať na aproximáciu a modelovanie spojitých pravdepodobnostných rozdelení širokého spektra rôznych javov, od aplikovania v meteorológii a modelovania vo finančnom sektore, až po štúdium Monte Carlo simulácií.

Ako bolo spomenuté, GLD vzniklo zovšeobecnením jednorozmerného Tukey-ho lambda rozdelenia. Tukey-ho Lambda rozdelenie je definované pomocou kvantilovej funkcie $Q(y)$, navrhnuté od Johna Tukeyho (1960).

$$Q(y) = \begin{cases} \frac{y^\lambda - (1-y)^\lambda}{\lambda}, & \lambda \neq 0 \\ \frac{\log(y)}{1-y}, & \lambda = 0, y \neq 1 \end{cases}$$

Tukey-ho rozdelenie bolo zovšeobecnené na štvor-parametrové, tzv. GLD (Generalized Lambda Distribution) za účelom generovania náhodných premenných pre Monte Carlo štúdie. Zovšeobecnené bolo Rambergom a Schmeiserom 1971-1974 a následne Mykytkom (1979).

2.1 Definícia zovšeobecneného lambda rozdelenia [1]

Definícia 2.1. *Zovšeobecnené lambda rozdelenie (GLD - Generalized Lambda Distribution) s parametrami $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, je najčastejšie definovaná po-*

mocou jej kvantilovej funkcie:

$$Q(y) = Q(y, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2}, \text{ kde } 0 \leq y \leq 1. \quad (2.1.1)$$

Parametre λ_1, λ_2 predstavujú parameter polohy a rozptyl a λ_3, λ_4 určujú šikmosť a špicatosť zovšeobecneného lambda rozdelenia. Je relatívne jednoduché odvodiť funkciu hustoty pre GLD. O tom nám hovorí nasledovná veta podľa [1].

Veta 2.1.1. *Pre $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ funkcia hustoty je:*

$$f(x) = \frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}}, \text{ pre } x = Q(y) \quad (2.1.2)$$

Dôkaz. Keď $x = Q(y)$, tak $y = F(x)$. Derivovaním y podľa x dostaneme funkciu hustoty:

$$f(x) = \frac{dy}{dx} = \frac{dy}{d(Q(y))} = \frac{1}{\frac{d(Q(y))}{dy}}. \quad (2.1.3)$$

Keďže poznáme tvar funkcie $Q(y)$, potom dostávame z 2.1.1 priamo derivovaním:

$$f(x) = \frac{1}{\frac{d(Q(y))}{dy}} = \frac{1}{\frac{d}{dy}(\lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2})} = \frac{1}{\frac{\lambda_3 y^{\lambda_3-1} - \lambda_4 (1-y)^{\lambda_4-1}}{\lambda_2}} = \frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}}$$

□

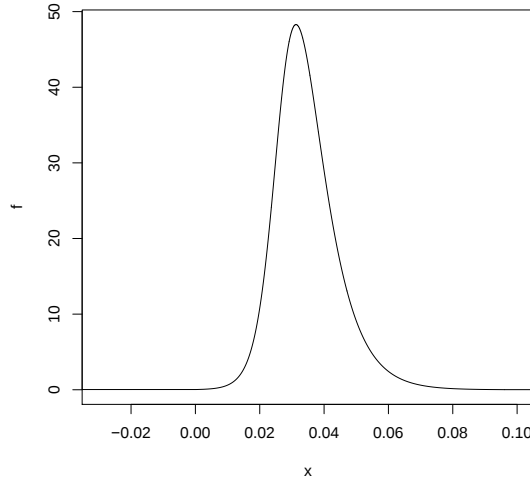
Pri vykresľovaní funkcie hustoty napríklad normálneho rozdelenia, kde $f(x)$ je dané ako funkcia od x , vypočítame $f(x)$ pre nejaké hodnoty x a vykreslíme dvojice $(x, f(x))$ a spojíme spojitou krivkou. Pri vykreslení funkcie hustoty GLD sa postupuje trochu inak, pretože predpis pre funkciu hustoty GLD nám dáva hodnotu $f(x)$ pre $x = Q(y)$ vid'. (2.1.2). Najskôr pre $y \in (0, 1)$ nájdeme x tak, že $x = Q(y)$. Potom nájdeme $f(x)$ podľa (2.1.2) a nakoniec vykreslíme dvojice $(x, f(x))$ a spojíme spojitou krivkou.

Príklad:[1] Chceli by sme nakresliť graf funkcie hustoty pre $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ s parametrami: $\lambda_1 = 0.0305, \lambda_2 = 1.3673, \lambda_3 = 0.004581, \lambda_4 = 0.01020$.

Kvantilová funkcia pre takéto GLD má tvar:

$$Q(y) = 0.0305 + \frac{y^{0.004581} - (1-y)^{0.01020}}{1.3673} \quad (2.1.4)$$

Vezmime si $y = 0.25$, potom $Q(0.25)=0.0280$ z (2.1.4). Potom $f(0.028) = 43.0399612$ podľa (2.1.2). Teda dvojica $(0.028, 43.04)$ bude jeden z bodov na grafe funkcie hustoty pre GLD s danými parametrami. Opakovaním tohto postupu pre $y=0.01, 0.02, \dots, 0.99$ dostaneme graf funkcie hustoty $GLD(0.0305, 1.3673, 0.004581, 0.01020)$.



Obr. 2.1.1: Funkcia hustoty pre $GLD(0.0305, 1.3673, 0.004581, 0.01020)$

2.2 Priestor parametrov GLD [1]

Zovšeobecnené lambda rozdelenie, známe aj ako nesymetrické lambda, alebo Ramberg-Schmeiser "RS rozdelenie", je rozdelenie so širokým spektrom tvarov. Keď si vezmeme vzťah pre výpočet kvantilovej funkcie (2.1.1), tak nie vždy špecifikuje platné rozdelenie. Dôvod je ten, že nemožno napísať ľubovoľnú formulu a predpokladať, že určuje platné rozdelenie bez toho, aby sme overili potrebné podmienky. V tejto časti sa zameriame na určenie hodnôt parametrov GLD tak, aby určovali platné rozdelenie. Z vlastností pre hustotu rozdelenia vieme, že funkcia $f(x)$ je hustota rozdelenia pravdepodobnosti vtedy a len vtedy, keď spĺňa nasledovné podmienky:

$$f(x) \geq 0 \quad \wedge \quad \int_{-\infty}^{\infty} f(x) dx = 1.$$

T.j. funkcia hustoty je vždy nezáporná a obsah plochy pod funkciou hustoty je rovná jednej. Podľa (2.1.2) a $x = Q(y)$ možno podmienky rozpísať nasledovne:

$$\frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}} \geq 0 \quad \wedge \quad \int_{-\infty}^{\infty} f(Q(y)) dQ(y) = 1 \quad (2.2.1)$$

Z (2.1.3) vieme, že $f(Q(y)) dQ(y) = dy$, a $y \in \langle 0, 1 \rangle$, teda druhá podmienka je splnená. Teda pre ľubovoľné $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ bude $\int_{-\infty}^{\infty} f(x) dx = 1$. Zostáva teda určiť, kedy platí 1. podmienka v 2.2.1. Pretože λ_1 v podmienke nevystupuje, tak tento parameter nemá žiadne obmedzenia. Nasledujúca veta (podľa [1]) špecifikuje význam λ_1 ako parametra polohy.

Veta 2.2.1. Ak náhodná premenná X je $GLD(0, \lambda_2, \lambda_3, \lambda_4)$, tak potom náhodná premenná $X + \lambda_1$ je $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

Veta 2.2.2. $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ určuje platné rozdelenie práve vtedy keď:

$$\frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}} \geq 0 \quad (2.2.2)$$

pre všetky $y \in \langle 0, 1 \rangle$.

Dôsledok 2.2.2.1. $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ je platným rozdelením práve vtedy keď

$$\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}$$

má rovnaké znamienko pre všetky $y \in [0, 1]$ ako parameter λ_2 . $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ špecifikuje platné rozdelenie najmä ak $\lambda_2, \lambda_3, \lambda_4$ majú rovnaké znamienko.

Okrem toho Ramberg v roku 1979 poznamenal, že existujú určité kombinácie λ_3, λ_4 , pre ktoré rozdelenie dané (2.1.1) je neplatné rozdelenie. Táto oblasť je daná nerovnosťou: $1 + \lambda_3^2 < \lambda_4 < 1.8(\lambda_3^2 + 1)$. Podľa Ramberga a Schmeisera existuje šesť oblastí v ktorých GLD parametre môžu ležať a v tieto oblasti sú charakteristické tým, že funkcie hustoty GLD majú podobné tvary. Keď chceme vymedziť dvojice (λ_3, λ_4) také, ktoré vedú k platnému $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, rozdelíme priestor (λ_3, λ_4) do niekoľkých oblastí.

$$Region1 = \{(\lambda_3, \lambda_4) | \lambda_3 \leq -1, \lambda_4 \geq 1\} \quad (2.2.3)$$

$$Region2 = \{(\lambda_3, \lambda_4) | \lambda_3 \geq 1, \lambda_4 \leq -1\} \quad (2.2.4)$$

$$Region3 = \{(\lambda_3, \lambda_4) | \lambda_3 \geq 0, \lambda_4 \geq 0\} \quad (2.2.5)$$

$$Region4 = \{(\lambda_3, \lambda_4) | \lambda_3 \leq 0, \lambda_4 \leq 0\} \quad (2.2.6)$$

$$V1 = \{(\lambda_3, \lambda_4) | \lambda_3 < 0, 0 < \lambda_4 < 1\} \quad (2.2.7)$$

$$V2 = \{(\lambda_3, \lambda_4) | 0 < \lambda_3 < 1, \lambda_4 < 0\} \quad (2.2.8)$$

$$V3 = \{(\lambda_3, \lambda_4) | -1 < \lambda_3 < 0, \lambda_4 > 1\} \quad (2.2.9)$$

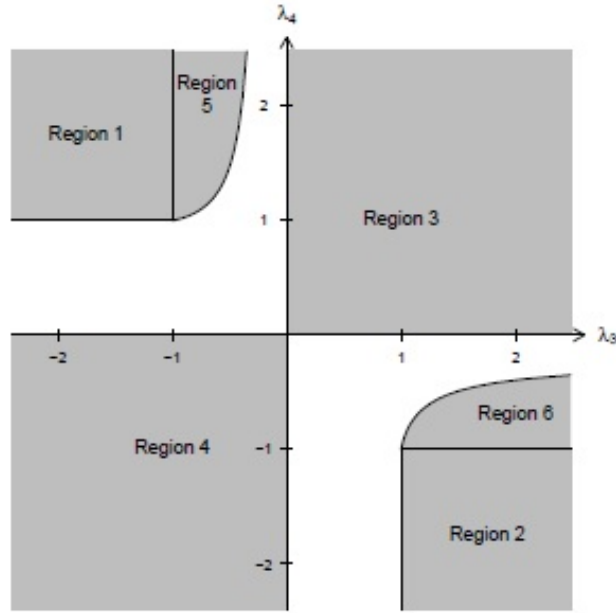
$$V4 = \{(\lambda_3, \lambda_4) | \lambda_3 > 1, -1 < \lambda_4 < 0\} \quad (2.2.10)$$

Platné GLD rozdelenie určujú dvojice (λ_3, λ_4) v oblastiach: Region 1,2,3,4, V_1, V_2 a niektoré dvojice z oblastí V_3, V_4 . Nasledovná veta, podľa [1], nám presnejšie definuje oblasti, kde dvojice (λ_3, λ_4) určujú platné $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ rozdelenie.

Veta 2.2.3. *Nech máme vhodné λ_2 . $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ je platné pre (λ_3, λ_4) vtedy a len vtedy, keď (λ_3, λ_4) je v jednom z vyfarbených oblastí na obrázku 2.2.1, pričom krivky hranice sú dané nasledovne:*

$$\frac{(1 - \lambda_3)^{1-\lambda_3}}{(\lambda_4 - \lambda_3)^{\lambda_4-\lambda_3}}(\lambda_4 - 1)^{\lambda_4-1} = \frac{-\lambda_3}{\lambda_4} \quad (\text{Region 5 v druhom kvadrante})$$

$$\frac{(1 - \lambda_4)^{1-\lambda_4}}{(\lambda_3 - \lambda_4)^{\lambda_3-\lambda_4}}(\lambda_3 - 1)^{\lambda_3-1} = \frac{-\lambda_4}{\lambda_3} \quad (\text{Region 6 v štvrtom kvadrante})$$



Obr. 2.2.1: Oblasti platného rozdelenia [11]

Pre hlbšiu štúdiu a dokazy viet pozri [1].

2.3 Tvary funkcie hustoty GLD [1]

V tejto časti budeme vyšetrovať možné tvary funkcie hustoty $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ a to pomocou rozdelenia priestoru (λ_3, λ_4) na oblasti, kde funkcia hustoty má jeden, dva alebo žiadny lokálny extrém.

Lema 2.3.1. *Funkcia hustoty $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ má lokálne extrém v bode y , ktoré spĺňa:*

$$g(y) = \frac{y^{\lambda_3-2}}{(1-y)^{\lambda_4-2}} = \frac{\lambda_4(\lambda_4-1)}{\lambda_3(\lambda_3-1)} \quad (2.3.1)$$

Dôkaz. Dôkaz tejto lemy dostaneme, ak deriváciu funkcie hustoty položíme rovnú 0:

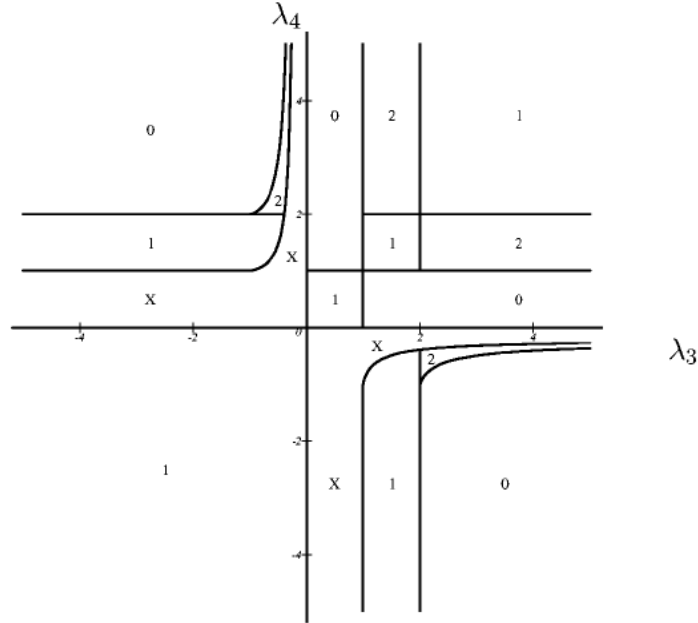
$$f(x) = \frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}}$$

Diferencovaním $f(x)$ podľa x dostaneme:

$$\begin{aligned} f'(x) &= \frac{d}{dy} \left(\frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4^{\lambda_4-1}} \right) \frac{dy}{dx} = \frac{d}{dy} \left(\frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4^{\lambda_4-1}} \right) f(x) = \\ &= -\lambda_2 \frac{(\lambda_3(\lambda_3 - 1)y^{\lambda_3-2} - \lambda_4(\lambda_4 - 1)(1 - y)^{\lambda_4-2})}{(\lambda_3 y^{\lambda_3-1} + \lambda_4^{\lambda_4-1})^2} f(x) = 0 \end{aligned} \quad (2.3.2)$$

Odtiaľ už ľahkou úpravou dostaneme 2.3.1. $\square$

Na nasledujúcom obrázku je znázornené pre každú oblasť príslušný počet lokálnych extrémov funkcie hustoty. Písmenom x sú oblasti, kde neexistuje platné GLD.



Obr. 2.3.1: Počet kritických bodov funkcie hustoty GLD rozdelení [1]

Symetria a prítomnosť chvostov sú dodatočné detaily tvaru funkcie hustoty GLD. Funkciu hustoty $f(x)$ možno vnímať ako funkciu k od $y, \lambda_1, \lambda_2, \lambda_3, \lambda_4$ a potom platí, že:

$$k(y, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = k(1 - y, \lambda_1, \lambda_2, \lambda_4, \lambda_3)$$

Teda ak $x = Q(y)$, dostávame:

$$x = Q(y, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1 - y)^{\lambda_4}}{\lambda_2} = \lambda_1 + A$$

A x hodnota ktorá korešponduje k $(1 - y)$ keď navzájom vymeníme λ_3, λ_4 bude:

$$x = Q(y, \lambda_1, \lambda_2, \lambda_4, \lambda_3) = \lambda_1 + \frac{(1 - y)^{\lambda_4} - y^{\lambda_3}}{\lambda_2} = \lambda_1 - A$$

Preto funkcia hustoty $\text{GLD}(\lambda_1, \lambda_2, \lambda_4, \lambda_3)$ je symetrický obraz funkcie hustoty $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ podľa priamky $x = \lambda_1$.

Veta 2.3.1. Funkcie hustoty $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ a $GLD(\lambda_1, \lambda_2, \lambda_4, \lambda_3)$ sú navzájom symetrické pričom os symetrie je priamka $x = \lambda_1$.

Poznámka: Pre $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ platí, že ak $\lambda_3 > \lambda_4$, tak rozdelenie má pozitívnu hodnotu skewness, t.j. hustota je od normálneho rozdelenia vychýlená na vpravo. A naopak, ak $\lambda_3 < \lambda_4$, tak rozdelenie má zápornú hodnotu skewness, t.j. hustota je od normálneho rozdelenia vychýlená vľavo.

Podľa toho, aké sú parametre λ_3, λ_4 , tak $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ môže mať všetky možné kombinácie konečných a nekonečných pravých a ľavých chvostov. Funkcia hustoty GLD má nekonečný pravý chvost práve vtedy, keď $\lambda_4 < 0$ a ľavý nekonečný chvost iba ak $\lambda_3 < 0$.

Na záver ešte spomenieme čo je na rozdelení GLD výnimočné a výhodné v modelovaní rozdelenia pravdepodobností. GLD je veľmi bohaté, čo sa týka tvarov funkcie hustoty a chvostov. Zahŕňa unimodálne, tzv. U-tvar, J-tvar, monotónne. Môže byť symetrické aj asymetrické, chvosty môžu byť hladké, abrupt, truncated, long, medium, short. Zahŕňa rovnomerné, exponenciálne, logistické rozdelenie..., dokonca môže veľmi dobre aproximovať aj normálne rozdelenie a ďalšie známe rozdelenia.

V oblasti financií sa často krát rozdelenie dát aproximuje pomocou normálneho rozdelenia. Avšak takéto dáta majú zväčša ťažšie chvosty - vyššia hodnota kurtosis (špicatosť), a hodnotu skewness (šikmosť) majú nenulové. Aproximáciou pomocou GLD možno zachytiť aj takéto vlastnosti, vďaka jeho štyrom parametrom. Pokým normálne rozdelenie má fixne danú hodnotu skewness(0) a kurtosis(3), tak GLD môže nadobúdať širokú škálu týchto parametrov, o čom sa možno presvedčiť pri nasledujúcich príkladoch rôznych hustôt rozdelenia GLD.

Obrázok (2.3.2): Hustota $GLD(0, 1, 2.5, \lambda_4)$, $\lambda_4 = 0.5, 0.75, 1$

Obrázok (2.3.3): Hustota $GLD(0, 1, 2.5, \lambda_4)$, $\lambda_4 = 1.25, 1.5, 1.75$

Obrázok (2.3.4): Hustota $GLD(0, 1, 2.5, \lambda_4)$, $\lambda_4 = 2, 2.25, 2.5$

Obrázok (2.3.5): Hustota $GLD(0, 1, 1.5, \lambda_4)$, $\lambda_4 = 2, 2.25, 2.5$

Obrázok (2.3.6): Hustota $GLD(0, 1, 0.5, \lambda_4)$, $\lambda_4 = 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5$

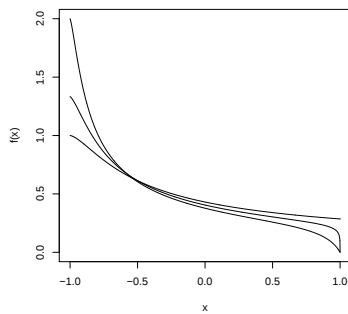
Obrázok (2.3.7): Hustota $GLD(0, -1, -0.2, \lambda_4)$, $\lambda_4 = 27, 30, 35, 50$

Obrázok (2.3.8): Hustota $GLD(0, -1, -0.5, \lambda_4)$, $\lambda_4 = 2.91, 2.92, 2.93, \dots, 2.99$

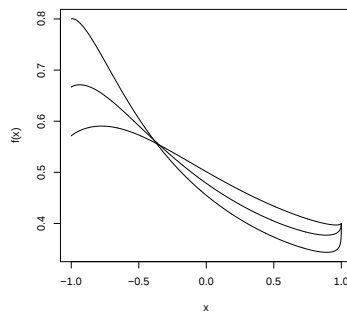
Obrázok (2.3.9): Hustota $GLD(0, -1, -3, \lambda_4)$, $\lambda_4 = 1, 1.25, 1.5, 1.75$

Obrázok (2.3.10): Hustota $GLD(0, -1, -0.25, \lambda_4)$, $\lambda_4 = -0.2, -0.3, -0.35, -0.5$

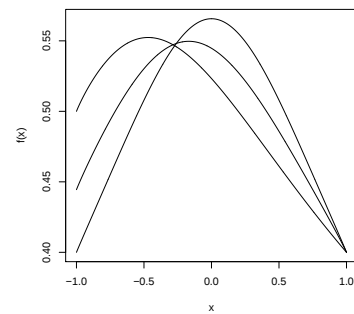
Pre viac informácií o vlastnostiach GLD rozdelenia nájdete v [1].



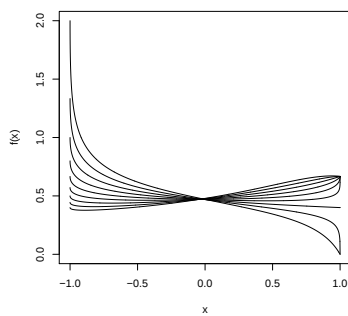
Obr. 2.3.2:



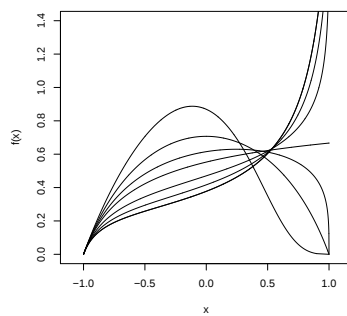
Obr. 2.3.3:



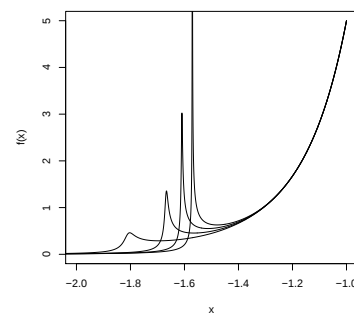
Obr. 2.3.4:



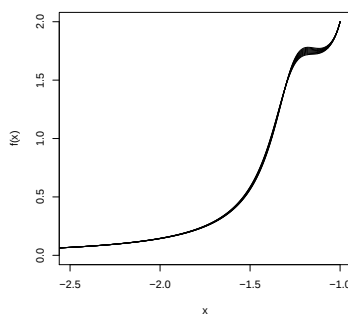
Obr. 2.3.5:



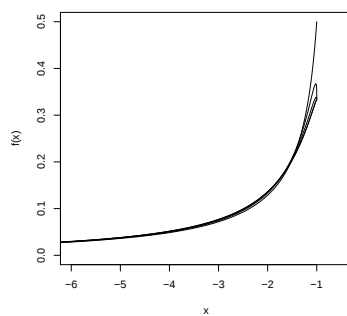
Obr. 2.3.6:



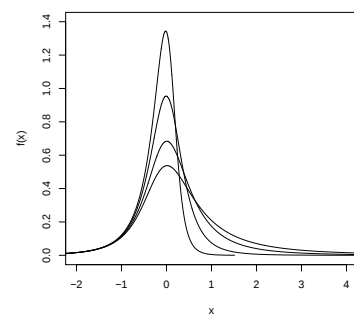
Obr. 2.3.7:



Obr. 2.3.8:



Obr. 2.3.9:



Obr. 2.3.10:

3 Aplikácia GLD na popis rozdelenia dát použitím metódy momentov

V tejto časti budeme pokračovať v predstavovaní GLD rozdelenia. Ukážeme, ako odhadnúť parametre rozdelenia GLD pre daný set dát. Ďalším bodom tejto časti bude aproximácia normálneho rozdelenia pomocou GLD rozdelenia, následne testy dobrej zhody a aplikácia odhadu parametrov GLD pre reálne výnosy akcie Hewlett Packard (HPQ). Záver kapitoly bude venovaný riešeniu systému nelineárnych rovníc, ktoré sme v tejto kapitole využili. V tejto kapitole vychádzame z [1], [4],[9], [10], [13],[16].

Vo väčšine prípadov, keď modelujeme štatistický model, nepoznáme vhodné pravdepodobnostné rozdelenie. Aj keď ho podľa nejakých rozumných predpokladov poznáme, tak často nepoznáme parametre tohto rozdelenia. Hlavná myšlienka odhadu parametrov GLD bude nasledovná:

- Pre daný set dát $X_1, X_2, \dots, X_n$ určíme prvé štyri momenty (mean, variance, skewness, kurtosis)
- Tieto momenty položíme rovné ich odhadom pomocou GLD rozdelenia
- Vyriešime príslušné rovnice pre $\lambda_1, \lambda_2, \lambda_3, \lambda_4$.

Postup popísaný vyššie sa nazýva metóda momentov.

3.1 Momenty zovšeobecneného lambda rozdelenia [1]

V tejto časti nájdeme vzťahy pre prvé štyri momenty GLD rozdelenia. Pre zjednodušenie, najskôr položíme $\lambda_1 = 0$. Získame tak necentrálne momenty $GLD(0, \lambda_2, \lambda_3, \lambda_4)$ a nakoniec odvodíme centrálné $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ momenty.

Veta 3.1.1. *Ak X je $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ náhodná premenná, potom $Z = X - \lambda_1$ je $GLD(0, \lambda_2, \lambda_3, \lambda_4)$*

Keď sme si už zaviedli λ_1 ako parameter polohy, môžeme teraz pomocou nasledujúcej vety (podľa [1]) zdefinovať necentrálne momenty $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ (ak existujú).

Veta 3.1.2. Ak Z je $GLD(0, \lambda_2, \lambda_3, \lambda_4)$, potom $E(Z^k)$, teda očakávaná hodnota Z^k , je daná:

$$E(Z^k) = \frac{1}{\lambda_2^k} \sum_{i=0}^k \left[\binom{k}{i} (-1)^i \beta(\lambda_3(k-i) + 1, \lambda_4 i + 1) \right], \quad (3.1.1)$$

kde $\beta(a, b)$ je beta funkcia definovaná vzťahom:

$$\beta(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx. \quad (3.1.2)$$

Predtým ako budeme pokračovať v hľadaní momentov $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, treba spomenúť tri vlastnosti beta funkcie, ktoré budú dôležité pri ďalšej práci.

Vlastnosti beta funkcie:

1. Integrál (4.1.1), ktorý definuje beta funkciu bude konvergovať práve vtedy keď a, b sú kladné.
2. Ak a, b sú kladné, tak $\beta(a, b) = \beta(b, a)$
3. Pre $u > -1$,

$$\beta(u+1, 1) = \beta(1, u+1) = \frac{1}{u+1} \quad (3.1.3)$$

Prvá spomenutá vlastnosť beta funkcie a vzťah (3.1.1) nám pomáha určiť podmienky, na základe ktorých momenty $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ existujú.

Dôsledok 3.1.2.1. k -ty moment $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ existuje vtedy a len vtedy, keď $\lambda_3 > -\frac{1}{k}$ a $\lambda_4 > -\frac{1}{k}$.

Keďže nás zaujímajú prvé štyri momenty, tak podľa predchádzajúceho dôsledku máme podmienku $\lambda_3 > -\frac{1}{4}$ a $\lambda_4 > -\frac{1}{4}$.

Nasledujúca veta podľa [1] dáva explicitné vyjadrenie pre prvé štyri centralizované $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ momenty.

Veta 3.1.3. Ak X je $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ s $\lambda_3 > -\frac{1}{4}$ a $\lambda_4 > -\frac{1}{4}$, potom jej prvé štyri momenty $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ (stredná hodnota, variancia, šikmosť a špicatosť), sú dané:

$$\alpha_1 = \mu = E(X) = \lambda_1 + \frac{A}{\lambda_2}, \quad (3.1.4)$$

$$\alpha_2 = \sigma^2 = E((X - \mu)^2) = \frac{B - A^2}{\lambda_2^2}, \quad (3.1.5)$$

$$\alpha_3 = \frac{E(X - E(X))^3}{\sigma^3} = \frac{C - 3AB + 2A^3}{\lambda_2^3 \sigma^3}, \quad (3.1.6)$$

$$\alpha_4 = \frac{E(X - E(X))^4}{\sigma^4} = \frac{D - 4AC + 6A^2B - 3A^4}{\lambda_2^4 \sigma^4}, \quad (3.1.7)$$

kde

$$A = \frac{1}{1 + \lambda_3} - \frac{1}{1 + \lambda_4} \quad (3.1.8)$$

$$B = \frac{1}{1 + 2\lambda_3} + \frac{1}{1 + 2\lambda_4} - 2\beta(1 + \lambda_3, 1 + \lambda_4) \quad (3.1.9)$$

$$C = \frac{1}{1 + 3\lambda_3} - \frac{1}{1 + 3\lambda_4} - 3\beta(1 + 2\lambda_3, 1 + \lambda_4) + 3\beta(1 + \lambda_3, 1 + 2\lambda_4) \quad (3.1.10)$$

$$D = \frac{1}{1 + 4\lambda_3} + \frac{1}{1 + 4\lambda_4} - 4\beta(1 + 3\lambda_3, 1 + \lambda_4) + 6\beta(1 + 2\lambda_3, 1 + 2\lambda_4) - 4\beta(1 + \lambda_3, 1 + 3\lambda_4) \quad (3.1.11)$$

Dôsledok 3.1.3.1. Ak $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ sú prvé štyri momenty $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, potom prvé štyri momenty $GLD(\lambda_1, \lambda_2, \lambda_4, \lambda_3)$ budú:

$$\alpha_1 - \frac{2A}{\lambda_2}, \quad \alpha_2, \quad -\alpha_3, \quad \alpha_4.$$

3.2 Priestor (α_3^2, α_4) [1]

V tejto časti sa zameriame na určenie hodnôt, ktoré momenty GLD môžu nadobúdať. Ak náhodná premenná Y má iné rozdelenie ako GLD, môžeme ho skúsiť aproximovať pomocou náhodnej premennej X , ktorá už bude mať GLD rozdelenie pre nejaké $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. Predpokladajme, že prvé štyri momenty sú $\alpha_1 = \mu, \alpha_2 = \sigma^2, \alpha_3$ a α_4 . Ak možno vybrať λ_3, λ_4 tak, že $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ má tretí a štvrtý moment α_3 a α_4 , potom λ_1, λ_2 sú dané ako:

$$\mu = \lambda_1 + \frac{A}{\lambda_2}$$

$$\sigma^2 = \frac{B - A^2}{\lambda_2^2}$$

Pretože A, B, C, D závisia iba od λ_3, λ_4 , tak dostávame nasledovný dôsledok vety 3.1.3.

Dôsledok 3.2.0.2. Pomocou $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ možno určiť ľubovoľne prvé dva momenty μ a σ^2 a nejaký tretí a štvrtý moment α_3, α_4 .

Čím väčšia je množina (α_3, α_4) , ktorú GLD môže generovať, tým užitočnejšie bude v aproximácii širokého rozsahu údajov a aproximovaní množstva ďalších náhodných premenných. Uvažujme teda rozsah hodnôt, ktoré môžu α_3, α_4 nadobúdať. Z 3.1.3.1 vieme, že ak (α_3, α_4) sa dá dosiahnuť, tak aj $(-\alpha_3, \alpha_4)$ sa dá výmenou λ_3, λ_4 dosiahnuť. Tento poznatok nám dovoľuje uvažovať priestor (α_3^2, α_4) priradený k $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

Veta 3.2.1. *Pre ľubovoľnú náhodnú premennú X pre ktorú existujú momenty α_3, α_4 platí: $\alpha_4 > 1 + \alpha_3^2$.*

Dôkazy k jednotlivým vetám a ďalšie poznatky možno nájsť v [1].

3.3 Odhad GLD parametrov pomocou metódy momentov [1]

Naším cieľom je odhadnúť parametre GLD pre dané dáta. Najznámejší a najpopulárnejší prístup na odhad parametrov rozdelenia je už spomenutá metóda momentov. Táto metóda je založená na porovnaní teoretických momentov $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ s empirickými (momenty určené z dát) a následne vyriešením rovníc pre $\lambda_1, \lambda_2, \lambda_3, \lambda_4$.

Momenty získané z dát $X_1, \dots, X_n$ označíme ako $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4$ a sú definované nasledovne:

$$\begin{aligned}\hat{\alpha}_1 = \bar{X} &= \sum_{i=1}^n \frac{x_i}{n} \\ \hat{\alpha}_2 = \hat{\sigma}^2 &= \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n} \\ \hat{\alpha}_3 &= \sum_{i=1}^n \frac{(X_i - \bar{X})^3}{(n\hat{\sigma}^3)} \\ \hat{\alpha}_4 &= \sum_{i=1}^n \frac{(X_i - \bar{X})^4}{(n\hat{\sigma}^4)}\end{aligned}\tag{3.3.1}$$

Toto nie sú maximum likelihood odhady (to by bolo ak by niektoré n boli nahradené $n-1$), ale korešponduje to s odhadmi pomocou metódy momentov.

Teoretické momenty z (3.1.4) až (3.1.7) označíme ako $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. Takže potrebujeme riešiť systém:

$$\alpha_i = \hat{\alpha}_i \text{ pre } i = 1, 2, 3, 4$$

Systém možno ľahko zjednodušiť na riešenie podsystému:

$$\alpha_3 = \hat{\alpha}_3, \quad \alpha_4 = \hat{\alpha}_4,$$

ako dve rovnice o dvoch neznámych λ_3, λ_4 , potom použiť (3.1.4), (3.1.5) a dostaneme postupne λ_2 a λ_1 . Po úpravách dostaneme nasledovný systém rovníc:

$$\hat{\alpha}_3 = \frac{C - 3AB + 2A^3}{(B - A^2)^{\frac{2}{3}}},\tag{3.3.2}$$

$$\hat{\alpha}_4 = \frac{D - 4AC + 6A^2B - 3A^4}{(B - A^2)^2},\tag{3.3.3}$$

kde A, B, C, D sú definované v (3.1.8) až (3.1.11) a závisia iba od λ_3, λ_4 . Teda $\hat{\alpha}_3, \hat{\alpha}_4$ poznáme, resp. máme odhadnuté z dát a hľadáme λ_3, λ_4 . Presné riešenia rovníc tohto typu (nelineárnych rovníc) sa hľadajú veľmi ťažko. Preto musíme hľadať približné

riešenie pomocou numerických metód. Metódy aproximácie sú založené na tom, že začneme z istých počiatočných hodnôt (v našom prípade $\lambda_3 = \lambda_3^*$, $\lambda_4 = \lambda_4^*$) a zakaždým skontrolujeme, či je splnená podmienka

$$\max \{|\alpha_3 - \alpha_3^*|, |\alpha_4 - \alpha_4^*|\} < \epsilon, \quad (3.3.4)$$

pre vopred zvolené epsilon. Ak táto podmienka neplatí, tak algoritmus hľadá lepšiu dvojicu λ_3^*, λ_4^* , pokiaľ nie je splnená podmienka (3.7.3). Nevýhodou je, že takéto algoritmy často končia neúspešne po istom počte iterácií. V našej práci sme na riešenie tohto systému rovníc využili balík v programovacom jazyku R - nleqslv (non-linear equation solver), o ktorom si povieme viac na konci tejto kapitoly.

3.4 Použitie GLD na aproximáciu normálneho rozdelenia [1]

V druhej kapitole sme videli, že funkcia hustoty GLD môže nadobúdať širokú škálu rôznych tvarov. Vďaka tomu by GLD malo byť schopné produkovať dobré výsledky pri rôznych rozdeleniach, z ktorých môžu dáta pochádzať. V tejto časti sa presvedčíme, že GLD veľmi dobre fituje jedno z najpoužívanějších rozdelení - normálne rozdelenie. Na to, aby sme overili vhodnosť rozdelenia, tak aplikujeme dva testy. Prvý test uvažuje blízkosť $\hat{f}(x)$, teda aproximovanú funkciu hustoty a $f(x)$, teda skutočnú funkciu hustoty rozdelenia, ktoré aproximujeme. Blízkosť $\hat{f}(x)$ a $f(x)$ je určená pomocou $\max |\hat{f}(x) - f(x)|$. Funkciu hustoty rozdelenia, ktorú aproximujeme, $f(x)$ nie je ťažké získať. Pri určení $\hat{f}(x)$ budeme postupovať nasledovne:

- Vezmeme si 249 bodov $y_i = i/250$ pre $i = 1, 2, \dots, 249$ z intervalu $(0,1)$.
- Použijeme 2.1.1 na $x_i = Q(y_i)$ pre $i = 1, 2, \dots, 249$
- Vypočítame $\hat{f}(x_i)$ podľa 2.1.2 pre $i = 1, 2, \dots, 249$

Potom použijeme aproximáciu

$$\max |\hat{f}(x) - f(x)| = \max_{1 \leq i \leq 249} |\hat{f}(x_i) - f(x_i)|$$

Pri druhom teste budeme porovnávať distribučné funkcie $\hat{F}(x)$ a $F(x)$ a použijeme tú istú formulu ako v prvom teste: $\max |\hat{F}(x) - F(x)| = \max_{1 \leq i \leq 249} |\hat{F}(x_i) - F(x_i)|$.

3.4.1 Normálne rozdelenie

Postup preberaný v predchádzajúcich častiach aplikujeme na odhad parametrov pre normálne rozdelenie. Funkcia hustoty normálneho rozdelenia so strednou hodnotou μ a disperziou σ má tvar:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

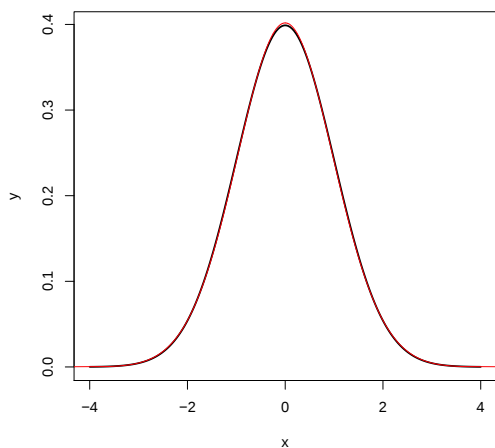
Každé normálne rozdelenie sa dá upraviť zo štandardizovaného normálneho rozdelenia, teda $N(0, 1)$. Stačí nám teda aproximácia $N(0, 1)$. Momenty pre $N(0, 1)$ sú všeobecne známe:

$$\alpha_1 = 0, \quad \alpha_2 = 1, \quad \alpha_3 = 0, \quad \alpha_4 = 3$$

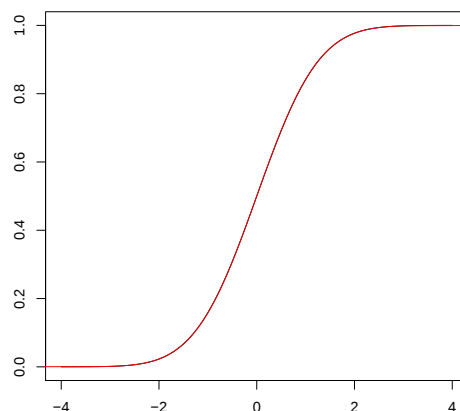
Teraz aplikujeme algoritmus na výpočet parametrov GLD, so štartovacím bodom $(\lambda_3^*, \lambda_4^*) = (0.15, 0.15)$ a dostaneme takéto parametre:

$$\lambda_1 = 0, \quad \lambda_2 = 0.1975, \quad \lambda_3 = 0.1349, \quad \lambda_4 = 0.1349$$

Vykreslíme si funkcie hustoty normálneho rozdelenia (znázornené čiernou) a do toho



Obr. 3.4.1: Funkcia hustoty



Obr. 3.4.2: Distribučná funkcia

istého obrázka aj odhad pomocou GLD (znázornené červenou), vid'. obr. 3.4.1. Vidíme, že sú takmer nerozoznateľné. To isté platí aj pre distribučné funkcie, vid'. obr. 3.4.2. Podľa obrázkov sa teda zdá, že odhad parametrov GLD pre normálne rozdelenie je veľmi dobrý. Aplikujme dva spomínané testy pre určenie ako presný je tento odhad. Aplikovaním prvého spomínaného testu (maximálny rozdiel $|\hat{f}(x) - f(x)|$) dostaneme:

$$\max |\hat{f}(x) - f(x)| = 0.002812276$$

Porovnaním distribučných funkcií dostaneme:

$$\max |\hat{F}(x) - F(x)| = 0.001085287$$

Druhý test sa dá veľmi ľahko interpretovať. Znamená to, že $P(X < x)$ sa odlišuje od jej aproximácie o nie viac ako 0.0010853 pre každé x . Aproximácie ďalších známych rozdelení pomocou GLD sú detailnejšie popísané v [1].

3.5 Testy dobrej zhody [1]

V tejto časti sa budeme zaoberať tým, ako otestovať vhodnosť použitého rozdelenia na konkrétnu dátovú vzorku pomocou testu dobrej zhody.

Test dobrej zhody, alebo taktiež známy ako Goodness Of Fit Test overuje, či empirické rozdelenie je štatisticky zhodné s niektorým z teoretických rozdelení pravdepodobnosti. Čo nás hlavne bude zaujímať v tejto časti je to, že ak sme už na základe dátovej vzorky odhadli parametre GLD rozdelenia podľa predchádzajúcich častí, tak či toto rozdelenie s odhadnutými parametrami štatisticky správne aproximuje skutočné rozdelenie dát. Tzn., že budeme testovať, či sa dané empirické rozdelenie štatisticky zhoduje s teoretickým GLD rozdelením so získanými parametrami. Existuje niekoľko rôznych štatistických testov pre test dobrej zhody, v našej práci budeme používať Pearson-ov Chi-kvadrát (χ^2) test.

3.5.1 Pearson-ov χ^2 test

Pearsonov Chi-kvadrát test dobrej zhody vychádza z frekvenčnej tabuľky a testuje nulovú hypotézu, že početnosti v jednotlivých kategóriách sa rovnajú očakávaným (teoretickým) početnostiam plynúcich z odhadnutého rozdelenia. Hlavnou myšlienkou tohto testu je rozdeliť získané hodnoty zo vzorky do niekoľkých kategórií, v našom prípade to budú intervaly a v každom intervale zistiť empirickú, teda skutočnú početnosť, koľko hodnôt padlo do daného intervalu a porovnať ju z teoretickou početnosťou, ktorú dostaneme z rozdelenia, ktoré chceme použiť na odhad. Teda testujeme hypotézu:

H_0 ...vzorka pochádza z predpokladaného rozdelenia

oproti alternatíve

H_1 ...vzorka nepochádza z predpokladaného rozdelenia

Ak hypotéza H_0 platí, znamená to, že odlišnosť skutočných početností v jednotlivých intervaloch a teoretických početností na základe predpokladaného rozdelenia, nie je štatisticky významná. Otázkou zostáva, akým spôsobom túto štatistickú významnosť testovať. Označme skutočné početnosti (**O**bserved) v jednotlivých intervaloch ako O_i , očakávané početnosti (**E**xpected) ako E_i , $i = 1, \dots, k$. Potom testovacia štatistika má tvar:

$$STAT = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

pričom k je počet intervalov, do ktorých sú dáta rozdelené. Pri platnosti H_0 má STAT chí kvadrát rozdelenie s $(k - p - 1)$ stupňami voľnosti, kde p je počet odhadovaných parametrov. Teda:

$$STAT = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{(k-p-1)}$$

Poznámka: Uvažujeme bežnú hladinu spoľahlivosti 95%.

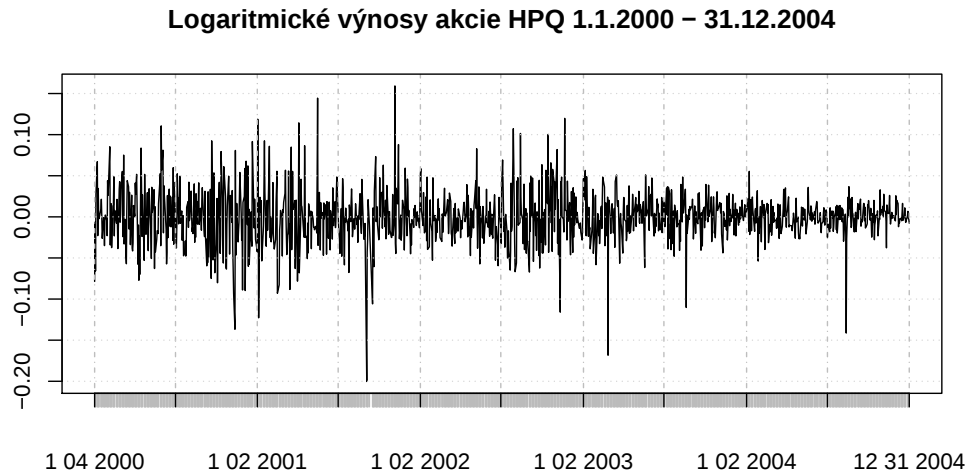
Teda v prípade GLD rozdelenia má testovacia štatistika za platnosti nulovej hypotézy chí kvadrát rozdelenie s $k - 4 - 1$ stupňami voľnosti, pretože pri aproximácii GLD odhadujeme štyri parametre. Teda získanú hodnotu STAT porovnávame s chí kvadrát rozdelením s $k - 5$ stupňami voľnosti. Ak je testovacia štatistika STAT väčšia ako hodnota chí kvadrát rozdelenia, tak nulovú hypotézu o správnom predpoklade rozdelenia zamietame, v opačnom prípade hypotézu H_0 prijímame a v tom prípade náš predpoklad rozdelenia môžeme považovať za vhodný.

Problémom tejto metódy je to, akým spôsobom zvoliť jednotlivé intervaly. To závisí hlavne od konkrétneho prípadu, avšak jedným z kritérií pre vhodné testovanie je, aby početnosti v jednotlivých intervaloch boli aspoň 5.

3.6 Odhad parametrov GLD pre reálne dáta [16]

Teóriu ohľadom aproximácie distribúcie nejakej vzorky pomocou GLD sme si vysvetlili v predchádzajúcich častiach. Teraz si prakticky ukážeme ako sa tieto poznatky využijú na reálnych dátach. Pre našu vzorku dát sme si vybrali logaritmické výnosy akcie Hewlett-Packard Company (HPQ) z obdobia od 1.1.2000 až po 31.12.2004. Môžeme sa pozrieť aj na ich priebeh počas tohto obdobia, viď. obrázok (3.6).

Ešte predtým ako sa pustíme do odhadovania GLD parametrov tak sa pozrieme na histogram týchto výnosov a otestujeme, či tieto dáta nepochádzajú z normálneho rozdelenia.



Obr. 3.6.1: Denné logaritmické výnosy akcie HPQ

Na obrázku (3.6.2) je znázornený histogram denných výnosov akcie HPQ a spojitou zelenou krivkou je vykreslená funkcia hustoty normálneho rozdelenia so strednou hodnotou a disperziou výnosov. Podľa obrázka sa zdá, že rozdelenie výnosov nie je z normálneho rozdelenia, hustota výnosov má ťažšie chvosty. Aby sme sa o tom presvedčili, tak sme vykonali chí-kvadrát test a p-hodnota pre tento test nám vyšla 0.

Teraz skúsime tieto dáta aproximovať pomocou prebraného GLD rozdelenia. Najskôr vypočítame momenty podľa 3.3.1 pre výnosy a dostaneme:

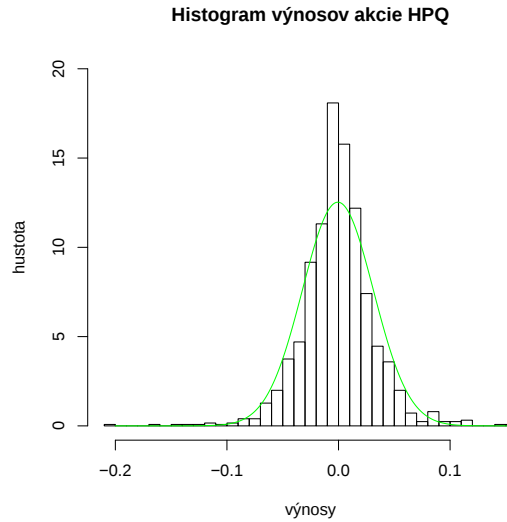
$$\hat{\alpha}_1 = -0.0005661917, \quad \hat{\alpha}_2 = 0.0010137762, \quad \hat{\alpha}_3 = -0.1219729671, \quad \hat{\alpha}_4 = 6.9224475571$$

Ďalej sa presvedčíme, že α_3, α_4 nespĺňajú podmienku: $1 + \lambda_3^2 < \lambda_4 < 1.8(\lambda_3^2 + 1)$, teda že nepatria do oblasti, kde α_3, α_4 neurčujú platné rozdelenie, spomínané v kapitole 2. Ďalším krokom bude riešenie dvoch nelineárnych rovníc s dvoma neznámymi:

$$\hat{\alpha}_3 = \frac{C - 3AB + 2A^3}{(B - A^2)^{\frac{2}{3}}}, \quad (3.6.1)$$

$$\hat{\alpha}_4 = \frac{D - 4AC + 6A^2B - 3A^4}{(B - A^2)^2}, \quad (3.6.2)$$

kde A,B,C,D závisia iba od λ_3, λ_4 a sú definované v 3.1.8 až 3.1.11. Ako bolo už spomenuté, tak riešenie takýchto rovníc sa hľadá numericky, čo si vyžaduje určiť štartovacie body pre λ_3, λ_4 . Tento problém vyriešime tak, že si vykreslíme krivky definované pomocou 3.6.1, 3.6.2, pozrieme sa na priesečníky týchto dvoch kriviek a definujeme štartovacie body na riešenie daného systému. Takýchto bodov je zväčša



Obr. 3.6.2: Histogram výnosov a aproximácia hustoty pomocou normálneho rozdelenia

niekoľko, väčšinou štyri. Pre finančné aktíva sú zaujímavé spravidla iba dvojice z tretieho kvadrantu (region 4, obr. (2.2.1)), teda také λ_3, λ_4 , ktoré sú obe záporné, avšak skúsime aj ostatné možnosti, aby sme sa o tom presvedčili. Keďže hodnota skewness ($\hat{\alpha}_3$) je záporná, tak hľadáme také parametre λ_3, λ_4 , že $\lambda_3 < \lambda_4$, podľa poznámky v kapitole 2. Pre tento prípad sme našli štyri priesečníky, teda štyri možné riešenia sústavy, t.j. štyri možné dvojice (λ_3, λ_4) . Tieto štyri dvojice budú štartovacie body pre algoritmus na hľadanie riešenia systému nelineárnych rovníc. V našom prípade bol zvolený už spomínaný balík v jazyku R a to package `nleqslv` (Non linear equation solver). Pomocou algoritmu sme dostali nasledujúce štvorice možných parametrov GLD:

$$GLD1 \quad (0.0002932782, -7.0626333938, -0.1047786071, -0.0998873030)$$

$$GLD2 \quad (-0.002503155, 8.504659630, 11.628442983, 14.945628036)$$

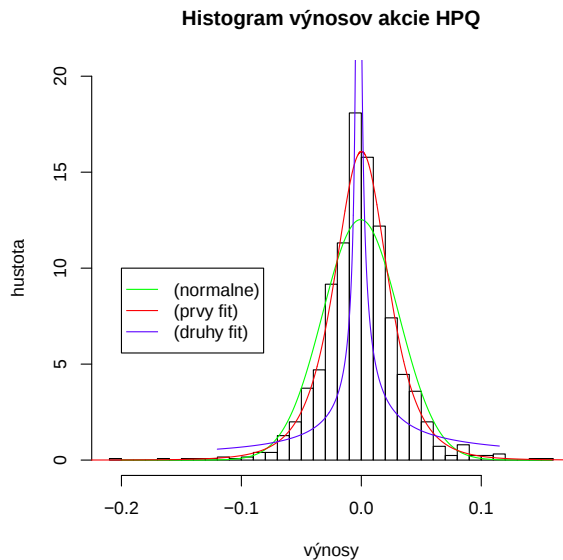
$$GLD3 \quad (0.1682218, -5.0494589, -0.1293255, 2.3755633)$$

$$GLD4 \quad (0.5287315, -2.0083450, -0.0704261, 77.4357622)$$

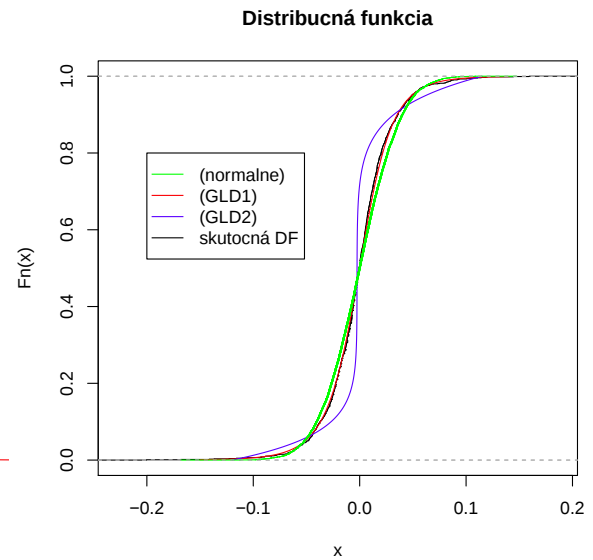
Z týchto štyroch možností sú prípustné len dve, pretože posledné dve nespĺňajú podmienku rozdelenia (λ_3, λ_4 majú opačné znamienka, čo podľa 2.2.2.1 nevedie k platnému GLD).

Takže získali sme dve dvojice parametrov GLD rozdelenia pre výnosy akcie HPQ (λ_3, λ_4). Aj keď len jedno riešenie spĺňa, že obe λ_3, λ_4 sú záporné, tak vyskúšame obe riešenia a presvedčíme sa, že rozumné výsledky dáva iba jedno z nich a konkrétne to so zápornými hodnotami λ_3, λ_4 . Nato, aby sme rozhodli, ktoré GLD rozdelenie je

pre naše dáta vhodnejšie, začneme tým, že si nakreslíme histogram dát a do neho dokreslíme hustoty oboch získaných GLD rozdelení. Na obrázku (3.6.3) teda máme histogram pre výnosy, zelenou krivkou hustotu normálneho rozdelenia, modrou GLD2 a červenou GLD1. Na obrázku (3.6.4) je kumulatívna distribučná funkcia dát (čiernou), aproximácia pomocou normálneho rozdelenia (zelenou), aproximácia pomocou GLD1 (červenou) a pomocou GLD2 (modrou).



Obr. 3.6.3: Funkcia hustoty

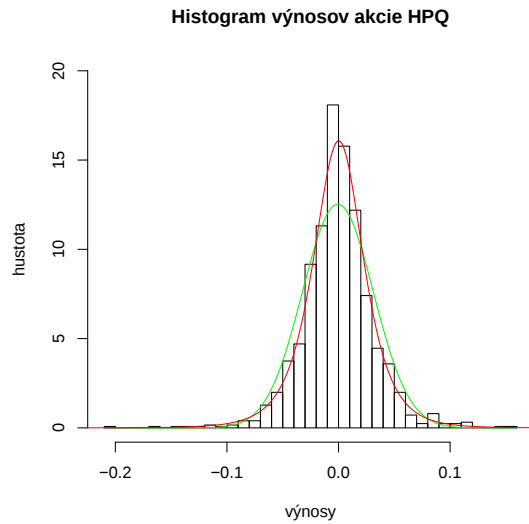


Obr. 3.6.4: Distribučná funkcia

Vidíme, že zrejme najlepšie popisuje hustotu dát červená krivka, teda GLD1. GLD2 síce spĺňa systém rovníc 3.6.1, 3.6.2, avšak hustotu rozdelenia dobre nepopisuje. Takže zdá sa, že najlepšiu aproximáciu nám dáva GLD1 rozdelenie (červené krivky) s parametrami:

$$\lambda_1 = 0.0002932782, \quad \lambda_2 = -7.0626333938, \quad \lambda_3 = -0.1047786071, \quad \lambda_4 = -0.0998873030$$

Na ďalšom obrázku je opäť histogram a pre porovnanie hustota normálneho rozdelenia (opäť zelenou) a hustota GLD1 najlepšie popisujúca naše dáta (červenou). Pozrime sa teraz na chí-kvadrát test pre toto najvhodnejšie GLD rozdelenie. V tabuľke na obr. (3.6.6) máme v prvom riadku rozdelenie do intervalov, v riadku druhom sú početnosti získané zo vzorky a v posledných dvoch riadkoch sú očakávané početnosti najskôr pre normálne rozdelenie a následne pre odhadnuté GLD rozdelenie. Vidíme, že rozdiely sú menšie pri GLD rozdelení, čo korešponduje aj s histogramom v predchádzajúcom obrázku. Hodnota chí-kvadrát rozdelenia pri hladine spoľahlivosti 95% so stupňami



Obr. 3.6.5: Histogram výnosov a aproximácia hustoty pomocou GLD rozdelenia

Interval	[-1;-0,08]	[-0,08;-0,05]	[-0,05;-0,03]	[-0,03;-0,02]	[-0,02;-0,015]	[-0,015;-0,01]	[-0,01;-0,005]	[-0,005;-0,003]	[-0,003;0]
Observed	15	46	106	115	63	79	102	47	59
Expected NORM	7,93	67,79	147,32	116,92	68,18	73,20	76,68	31,25	47,12
Expected GLD	15,06	47,69	107,93	106,68	70,28	82,83	92,87	37,65	58,99

Interval	[0;0,003]	[0,003;0,005]	[0,005;0,01]	[0,01;0,015]	[0,015;0,02]	[0,02;0,03]	[0,03;0,05]	[0,05;0,08]	[0,08;1]
Observed	49	45	104	83	70	93	101	37	22
Expected NORM	47,04	31,11	76,04	72,19	66,87	113,73	141,09	63,36	7,17
Expected GLD	58,99	38,91	92,87	82,83	71,54	105,42	104,17	43,93	12,55

Obr. 3.6.6: Tabuľka pre chí-kvadrát test

voľnosti $18-4-1=13$, teda χ^2_{13} sa rovná 22.36203. Pri normálnom rozdelení sa odhadujú len dva parametre, teda budeme porovnávať s $\chi^2_{15} = 24.99579$. Jednotlivé štatistiky pre obe rozdelenia sú nasledovné:

Normálne rozdelenie: $STAT = 120.22$

GLD rozdelenie: $STAT = 18.67$

Teda vidíme, že testovacia štatistika pre normálne rozdelenie je omnoho väčšia ako kritická hodnota. A tak isto môžeme vidieť, že testovacia štatistika pre GLD rozdelenie je menšia ako chí-kvadrát hodnota, teda hypotézu o správnom predpoklade rozdelenia pre GLD nezamietame. P-hodnota v tomto prípade vyšla $13.36\% > 5\%$.

3.7 Riešenie nelineárneho systému rovníc [4],[10], [13]

V tejto časti predstavíme balík programovacieho jazyka R - nleqslv určeného na riešenie nelineárneho systému rovníc, ktorý využívame pri odhadoch parametrov GLD rozdelenia.

Pri odhadovaní parametrov GLD v predchádzajúcej časti sme narazili na problém vyriešiť systém dvoch nelineárnych rovníc s dvoma neznámymi λ_3, λ_4 :

$$\hat{\alpha}_3 = \frac{C - 3AB + 2A^3}{(B - A^2)^{\frac{2}{3}}}, \quad (3.7.1)$$

$$\hat{\alpha}_4 = \frac{D - 4AC + 6A^2B - 3A^4}{(B - A^2)^2}, \quad (3.7.2)$$

kde A,B,C,D závisia iba od λ_3, λ_4 a sú definované v (3.1.8) až (3.1.11). Ako sme už spomenuli, takýto systém sa nedá vyriešiť analyticky, preto musíme hľadať približné riešenie pomocou numerických metód. Numerické metódy sú založené na tom, že začneme z nejakého štartovacieho bodu (v našom prípade $\lambda_3 = \lambda_3^*, \lambda_4 = \lambda_4^*$) a zakaždým skontrolujeme, či je splnená podmienka

$$\max \{|\alpha_3 - \alpha_3^*|, |\alpha_4 - \alpha_4^*|\} < \epsilon, \quad (3.7.3)$$

pre vopred zvolené epsilon, teda či sme požadovane blízko riešenia. Ak táto podmienka neplatí, tak algoritmus hľadá lepšiu dvojicu λ_3^*, λ_4^* pokým nie je splnené podmienka (3.7.3). Rozdiely v jednotlivých metódach algoritmov spočívajú v tom, ako tú lepšiu dvojicu hľadať. V balíku je možnosť využiť dva algoritmy na riešenie systému nelineárnych rovníc:

- Broyden Secant metódu
- Kompletnú Newton metódu

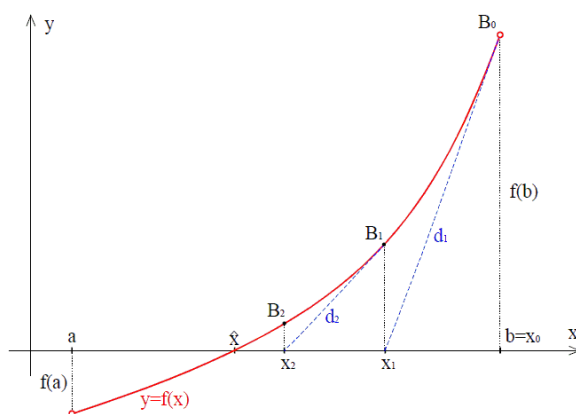
Obe metódy využívajú globálne stratégie ako linesearch alebo trust region, keď štandardná Newton/Broyden metóda nevedie bližšie k riešeniu systému. Linesearch môže byť kvadratická, alebo geometrická. Trust region metódy sú buď double dogleg, alebo Powel single dogleg. Jednotlivé globálne stratégie detailnejšie v [9]. Balík obsahuje možnosť použitia numerického alebo analytického výpočtu jakobiánu. Dokáže škálovať parametre buď automaticky, alebo fixne. Keďže jednotlivé metódy balíka budeme využívať, stručne popíšeme ich algoritmy.

3.7.1 Newtonova metóda: [13]

Newtonova metóda je jednou z často používaných iteračných metód a niekoľko ďalších metód je odvodených práve od Newton metódy. Táto metóda sa nazýva aj metódou dotyčníc, kvôli jej geometrickému princípu. Je založená na princípe postupnej linearizácie, teda nahradení ťažkého nelineárneho problému na postupnosť lineárnych problémov, ktorých riešenie konverguje k riešeniu nelineárneho problému. Pre jednoduchosť uvažujme najskôr jednorozmerný problém. Keď máme štartovací bod, z ktorého začína aproximácia, tak vytvárame postupnosť iterácií $\{x_p\}_{p=0}^{\infty}$ podľa vzťahu:

$$x_{p+1} = x_p - \frac{f(x_p)}{f'(x_p)}$$

Teda ak x_p je p -ta aproximácia, tak priesečník dotyčnice k funkcii f v bode x_p s osou x bude ďalšia aproximácia, teda x_{p+1} . Lepšie pochopenie docieli obrázok 3.7.1.



Obr. 3.7.1: Princíp Newton metódy [4]

Predpokladajme teraz systém nelineárnych rovníc:

$$F(x) = 0,$$

kde $F : \mathbb{R}^n \mapsto \mathbb{R}^n$ je nelineárna funkcia so spojitými parciálnymi deriváciami a $J(x) = F'(x)$ pre všetky $x \in \mathbb{R}^n$. Newtonova iterácia v prípade systému nelineárnych rovníc je daná:

$$x_{k+1} = x_k - J(x_k)^{-1}F(x_k)$$

kde $J(x_k)^{-1}$ je inverzná funkcia k jakobiánu $J(x_k)$. Výpočet inverznej funkcie $J(x_k)^{-1}$ sa dá nahradiť riešením rovnice:

$$J(x_k)s_k = J^{-1}(x_k)$$

Newtonova metóda má veľmi atraktívne teoretické a praktické vlastnosti. Napr. ak x^* je riešenie systému, kde jakobián $J(x^*)$ nie je singulárny a štartovací bod x_0 je blízko riešenia, tak x_k konverguje super lineárne k riešeniu x^* . Nevýhodou Newton metódy je to, že výpočet jakobiánu v každej iterácii je časovo náročný. Vylepšením sú potom quasi Newton metódy, ktoré jakobián aktualizujú nie po každej iterácii.

3.7.2 Broyden metóda:

Broyden metóda je zovšeobecnená metóda sečníc pre systém nelineárnych rovníc. Metóda sečníc (secant method) je založená na tom, že deriváciu nahrádza konečnou diferenciou.

$$f'(x_k) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

Teda prvým rozdielom oproti Newton metóde je to, že Broyden metóda používa aproximáciu derivácie. Ďalším rozdielom je, že jakobián nie je prepočítavaný v každej iterácii, ale po každej hlavnej iterácii využitím tzv. Broyden rank 1 update. Algoritmus začína s počítaným jakobiánom funkcie a aktualizuje ho po každej úspešnej iterácii použitím Broyden aktualizácie. Tento prístup často vykazuje super lineárnu konvergenciu k riešeniu. Ak funkcia nleqslv rozhodne, že nemožno pokračovať s aktuálnou maticou derivácií, počíta novú.

Riešenie nelineárneho systému rovníc nie je predmetom našej práce, preto sa ňou nebudeme viac zaoberať. Pre viac informácií pozri [9], [13].

4 Rozšírený GLD systém - EGLD

V tejto kapitole sa budeme venovať rozšíreniu GLD na rozšírené GLD (Extended GLD) pridaním skupiny zovšeobecneného beta rozdelenia (Generalized Beta Distribution). Aj keď v tejto diplomovej práci budeme využívať iba samotné GLD, pretože na popis distribúcie výnosov aktív na trhu bolo GLD v našej práci postačujúce, pre kompletnosť spomenieme aj zovšeobecnené beta rozdelenie. Vychádzať budeme opäť z [1].

Samotné GLD možno použiť len v prípade ak dvojica α_3, α_4 nespĺňa nasledovné:

$$1 + \alpha_3^2 < \alpha_4 \leq 1.8(\alpha_3^2 + 1)$$

Teda ak tretí a štvrtý moment spĺňajú predchádzajúcu podmienku, tak pomocou GLD nemožno rozdelenie vzorky aproximovať. Avšak túto oblasť dokáže pokryť už spomínané zovšeobecnené beta rozdelenie alebo Generalized Beta Distribution (ďalej už len GBD). Skôr ako sa dostaneme k samotnej definícii GBD, najskôr pre lepšie pochopenie definujeme beta rozdelenie a jej momenty. Ďalej rozvineme beta rozdelenie na zovšeobecnené beta rozdelenie - GBD a následne použitím metódy momentov odhadneme parametre GBD.

4.1 Beta rozdelenie a jej momenty [1]

Beta rozdelenie s parametrami $\beta_3 < -1, \beta_4 > -1$ je definované pomocou jej funkcie hustoty nasledovne:

$$f(x) = \begin{cases} \frac{x^{\beta_3}(1-x)^{\beta_4}}{\beta(\beta_3+1, \beta_4+1)}, & 0 \leq x \leq 1 \\ 0, & \text{inak} \end{cases}$$

pričom $\beta(a, b)$ je beta funkcia definovaná:

$$\beta(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx. \quad (4.1.1)$$

Predtým ako určíme momenty, spomenieme niekoľko užitočných vlastností beta a gamma funkcie. Pre $a > 0$, gamma funkcia je definovaná:

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$$

4.1.1 Vlastnosti

1. $\Gamma(a+1) = a\Gamma(a)$

$$2. \Gamma(a) = 2 \int_0^\infty u^{2a-1} e^{-u^2}$$

$$3. \text{ Pre } a, b \text{ kladné: } \beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Veta 4.1.1. Ak X je beta náhodná premenná s parametrami β_3, β_4 a ak $B_i = \beta_3 + \beta_4 + i$, kde $i=1, 2, \dots$, potom

$$E(X^k) = \frac{\beta(\beta_3 + k + 1, \beta_4 + 1)}{\beta(\beta_3 + 1, \beta_4 + 1)}$$

A pomocou rekurzie:

$$E(X^k) = \left(\frac{\beta_3 + k}{B_{k+1}} \right) E(X^{k-1}) \text{ pre } k = 1, 2, 3, \dots, E(X^0) = 1$$

Veta 4.1.2. Ak X je beta náhodná premenná s parametrami β_3, β_4 a ak $B_i = \beta_3 + \beta_4 + i$, kde $i=1, 2, \dots$, potom:

$$\begin{aligned} \alpha_1 &= E(X) = \mu = \frac{\beta_3 + 1}{B_2} \\ \alpha_2 &= E[(X - \mu)^2] = \sigma^2 = \frac{(\beta_3 + 1)(\beta_4 + 1)}{B_2^2 B_3} \\ \sigma^3 \alpha_3 &= E[(X - \mu)^3] = 2 \frac{(\beta_3 + 1)(\beta_4 + 1)(\beta_4 - \beta_3)}{B_2^3 B_3 B_4} \\ \sigma^4 \alpha_4 &= E[(X - \mu)^4] = 3 \frac{(\beta_3 + 1)(\beta_4 + 1)}{B_2^4 B_3 B_4 B_5} \times (\beta_3 \beta_4 B_2 + 3\beta_3^2 + 3\beta_4^2 + 5\beta_3 + 5\beta_4 + 4) \end{aligned}$$

4.2 GBD a jej momenty [1]

Na to, aby sme dokázali fitovať dáta pomocou beta rozdelenia, potrebujeme aby beta rozdelenie bolo viac flexibilné, teda nie len na intervale $(0, 1)$. Toto možno vyriešiť transformáciou beta rozdelenia na zovšeobecnené beta rozdelenie (generalized beta distribution - GBD).

Definícia 4.1. Ak X je beta náhodná premenná s parametrami β_3, β_4 , pre $\beta_2 > 0$ a pre ľubovoľné β_1 náhodná premenná $Y = \beta_1 + \beta_2 X$ má zovšeobecnené beta rozdelenie (ďalej už len GBD).

Veta 4.2.1. Funkcia hustoty pre náhodnú premennú s rozdelením $GBD(\beta_1, \beta_2, \beta_3, \beta_4)$ je:

$$f(x) = \begin{cases} \frac{(x - \beta_1)^{\beta_3} (\beta_1 + \beta_2 - x)^{\beta_4}}{\beta(\beta_3 + 1, \beta_4 + 1) \beta_2^{\beta_3 + \beta_4 + 1}} & \text{ak } \beta_1 \leq x \leq \beta_1 + \beta_2 \\ 0, & \text{inak} \end{cases}$$

4.3 Momenty GBD [1]

Ak chceme fitovať $GBD(\beta_1, \beta_2, \beta_3, \beta_4)$ na dáta použitím metódy momentov, potrebujeme jednak odhady momentov získané z dát, označme ich: $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4$. A ďalej potrebujeme predpis pre momenty $GBD(\beta_1, \beta_2, \beta_3, \beta_4)$ rozdelenia, tie označme $\alpha_1, \alpha_2, \alpha_3, \alpha_4$.

Veta 4.3.1. *Nech Y je náhodná premenná z rozdelenia $GBD(\beta_1, \beta_2, \beta_3, \beta_4)$ so strednou hodnotou μ_Y a varianciou σ_Y^2 . Potom momenty GBD majú nasledovné predpisy:*

$$\alpha_1 = E(X) = \mu = \beta_1 + \frac{\beta_2(\beta_3 + 1)}{B_2} \quad (4.3.1)$$

$$\alpha_2 = E[(X - \mu)^2] = \sigma^2 = \frac{\beta_2^2(\beta_3 + 1)(\beta_4 + 1)}{B_2^2 B_3} \quad (4.3.2)$$

$$\alpha_3 = E[(X - \mu)^3] = \frac{2(\beta_4 - \beta_3)\sqrt{B_3}}{B_4\sqrt{(\beta_3 + 1)(\beta_4 + 1)}} \quad (4.3.3)$$

$$\alpha_4 = E[(X - \mu)^4] = \frac{3B_3(\beta_3\beta_4B_2 + 3\beta_3^2 + 5\beta_3 + 3\beta_4^2 + 5\beta_4 + 4)}{B_4B_5(\beta_3 + 1)(\beta_4 + 1)} \quad (4.3.4)$$

4.4 Odhad parametrov GBD [1]

V tejto časti, budeme postupovať identicky ako pri odhade parametrov GLD. Použijeme opäť metódu momentov. Tak ako pri odhade GLD parametrov potrebujeme riešiť systém rovníc:

$$\alpha_i = \hat{\alpha}_i \quad \text{pre } i = 1, 2, 3, 4,$$

pre $\beta_1, \beta_2, \beta_3, \beta_4$, kde $\hat{\alpha}_i$ sú skutočné momenty získané zo vzorky dát. Aj v tomto prípade sa tieto štyri rovnice dajú zredukovať na dve, vďaka tomu, že rovnice (4.3.3), (4.3.4) závisia iba od β_3, β_4 . Teda aj pri odhade parametrov GBD nám stačí vyriešiť iba dve rovnice:

$$\alpha_3 = \hat{\alpha}_3 \quad \alpha_4 = \hat{\alpha}_4,$$

pre β_3, β_4 a následne dopočítať β_1, β_2 z rovníc (4.3.1), (4.3.2). Rovnako ako pri GLD sme sa dostali k problému riešiť systém nelineárnych rovníc, teda na riešenie potrebujeme numerické metódy. Viac informácií ohľadom teórie GBD nájdete v [1].

5 Odhad Value at Risk (VaR) pomocou GLD

V ďalších častiach diplomovej práce budeme odhadovať mieru rizika - Value at Risk (VaR) pomocou preberaného zovšeobecneného lambda rozdelenia. Konkrétne v tejto kapitole sa budeme venovať teórii výpočtu VaR. Spomenieme niekoľko metód, pričom sa hlavne zameriame na klasickú metódu Monte Carlo ako aj Monte Carlo metódu v kombinácii s GARCH modelmi. Pri každej metóde budeme porovnávať výsledky pomocou GLD rozdelenia s rozdelením normálnym. Teóriu ohľadom jednotlivých metód sme čerpali z nasledovných zdrojov [1],[2],[3],[6], [7],[8], [12], [14], [15].

5.1 Value-at-Risk

Vývoj na finančnom trhu je veľmi ťažko predvídateľný a snahou každého účastníka trhových aktivít je dostatočne sa zabezpečiť proti vysokým stratám. Avšak ešte predtým ako sa dostaneme k manažovaniu rizika, tak musíme byť schopný zmerať/odhadnúť, akému riziku a v akej podobe čelíme. Hlavnou otázkou je to, aká je maximálna možná strata na určitom časovom horizonte pri danej hladine spoľahlivosti ak investujeme do istej množiny aktív. Odpoveď na túto otázku nám dáva jeden z najrozšírenejších a najviac používaných nástrojov pre manažment v bankovom sektore, a tým je Value-at-Risk (VaR), teda miera rizika. Je to štandardný nástroj na meranie trhového rizika, teda rizika spojeného s prudkými zmenami cien jednotlivých aktív na finančnom trhu.

Pre lepšiu predstavu, čo VaR predstavuje, uvedieme príklad. Ak sa povie, že jednodňová VaR portfólia pri hladine spoľahlivosti 99% je 1000 eur, znamená to, že s pravdepodobnosťou 99% dosiahneme maximálnu možnú stratu portfólia 1000 eur za jeden deň.

5.2 Výpočet Value-at-Risk

Ako sme sa už zmienili, VaR vyjadruje maximálnu možnú stratu portfólia pri istej hladine spoľahlivosti za dané časové obdobie. Teda merať trhové riziko znamená priradiť portfóliu istú hodnotu, ktorá charakterizuje ako veľmi sa môže vychýliť od hodnoty v súčasnosti (resp. koľko môžeme stratiť). Stratou sa rozumie pokles hodnoty rizikového faktora - napríklad ceny akcie, či portfólia akcií.

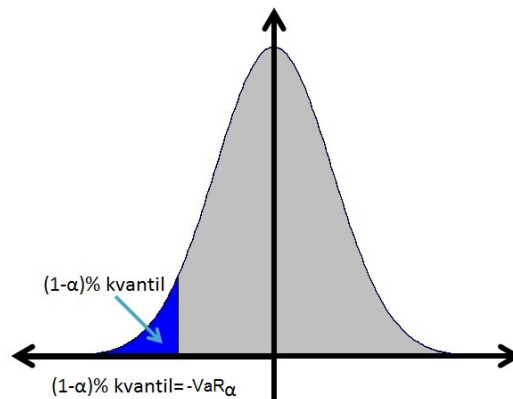
Označme L ako jednodňovú stratu portfólia ($L > 0$). Uvažujme hladinu spoľahlivosti

α . Potom jednoduchú VaR_α môžeme vyjadriť nasledovne:

$$VaR_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\}, \alpha \in (0, 1)$$

To znamená, že jednoduchá strata portfólia prekročí VaR s pravdepodobnosťou menšou, najvyššou rovnou ako $(1 - \alpha) \cdot 100\%$. Teda hľadáme takú hodnotu, že pravdepodobnosť jej prekročenia je rovná $(1 - \alpha) \cdot 100\%$. Hlavnou myšlienkou výpočtu VaR_α je odhad rozdelenia výnosov a určenie príslušného $(1 - \alpha) \cdot 100\%$ kvantilu rozdelenia, teda $-VaR_\alpha$.

Poznámka: VaR sa chápe ako strata a vyjadruje sa ako kladná hodnota.



Obrázok nám dáva grafickú predstavu výpočtu VaR. Keď už máme informáciu o tom, aké pravdepodobnostné rozdelenie výnosy sledujú a vieme ich parametre, tak výpočet VaR je jednoduchý, predstavuje $(1 - \alpha)\%$ kvantil (na obrázku je znázornený 5% kvantil normálneho rozdelenia). Aj keď podľa predchádzajúcej formulky má výpočet VaR jasnú štruktúru, určenie príslušného kvantilu je náročný štatistický problém. Existuje niekoľko rôznych prístupov určenia VaR, ktoré sa líšia najmä v prístupe odhadu rozdelenia výnosov portfólia.

Výpočet VaR sa rozdeľuje do troch hlavných skupín.

1. **Neparametrické (Historická simulácia, Metóda Monte Carlo)**
2. **Semiparametrické (Teória extrémnych hodnôt, CAViaR)**
3. **Parametrické (RiskMetrics, GARCH)**

V našej práci budeme na výpočet VaR metódu Monte Carlo a v ďalšej časti aj s kombináciou GARCH modelov, preto si tieto metódy popíšeme bližšie, no predtým pre lepšie pochopenie predstavíme aj myšlienku metódy historickej simulácie.

5.2.1 Metóda historickej simulácie

Príkladom neparametrickej metódy je metóda historickej simulácie. Historická simulácia je založená na empirickom pravdepodobnostnom rozdelení. Hlavným predpokladom je to, že výnosy v blízkej budúcnosti sa vyvíjajú podľa nedávnej minulosti. Problém tejto metódy je určenie toho, čo znamená nedávna minulosť. Ak vezmeme príliš veľa historických dát na výpočet VaR, tak sa môže stať, že náš odhad môže byť skreslený starými udalosťami vo vývoji výnosov. Taktiež naopak, ak vezmeme príliš málo historických údajov, tak nemusia zachytiť prípadné extrémne udalosti vo vývoji a v tom prípade bude VaR podhodnotená.

Ak už máme rozmyslené aké veľké "okno" na odhad VaR vyberieme, tak výpočet je veľmi jednoduchý a predstavuje $(1 - \alpha)\%$ kvantil empirického rozdelenia výnosov v danom "okne".

Existujú aj rozšírenia metódy historickej simulácie tzv. *vážené historické simulácie*. Pokým v klasickej historickej simulácii majú všetky pozorovania rovnakú váhu, pri váženej historickej simulácii môže váha závisieť od:

- časovej aktuálnosti
- od aktuálnej volatility
- od aktuálnej korelácie

Napríklad pri časovo váženej historickej simulácii majú najstaršie pozorovania najmenšiu váhu a tie najaktuálnejšie najväčšiu.

Výhody metódy historickej simulácie:

- Nie je potrebný predpoklad na štatistické rozdelenie rizikových faktorov.
- Intuitívna, ľahko interpretovateľná metóda, má jasnú a jednoduchú koncepciu výpočtu.
- Jednoduché rozšírenie pre portfólio.

Nevýhody metódy historickej simulácie:

- Predpoklad rovnakého rozdelenia výnosov v čase nie je správny.

- Dĺžka okna by mala spĺňať dve protichodné podmienky: okno má byť dostatočne dlhé kvôli dobrým štatistickým záverom a nesmie byť príliš dlhé aby sa nebrali do úvahy pozorovania veľmi vzdialené do minulosti.
- Extrémny výsledok v jeden deň môže negatívne ovplyvniť VaR počas celej doby v rámci daného časového okna.
- VaR je limitovaná do takej výšky hodnoty výnosu, ktorá skutočne nastala.
- Pomaly reaguje na turbulencie na trhu.

5.2.2 Metóda Monte Carlo

Výpočet VaR pomocou metódy Monte Carlo má veľmi podobný algoritmus ako samotná historická simulácia. Avšak namiesto priameho určenia kvantilu z empirického rozdelenia výnosov daného obdobia, v tejto metóde opakovane náhodne simulujeme výnosy a každá jedna simulácia nám dá potom možnú budúcu hodnotu portfólia. Ak simuláciu zopakujeme mnoho krát, tak rozdelenie nasimulovaných výnosov sa bude približovať skutočnému ale neznámemu rozdeleniu. Potom z neho odvodíme empirické rozdelenie a určíme opäť $(1 - \alpha)\%$ kvantil a teda získame VaR. Napríklad ak počítame 99% VaR, tak ak vykonáme 1000 simulácií výnosu, tak 10. najnižšia hodnota bude VaR, teda hodnota, ktorá oddeľuje 1% najnižších výnosov daného portfólia.

Najväčším problémom tejto metódy je nájsť vhodné štatistické rozdelenie rizikového faktora, alebo stochastický proces, ktorý by vhodne popisoval správanie rizikového faktora. Čo sa týka stochastických procesov, tak najčastejšie sa vo finančníctve využíva Brownov geometrický pohyb, či od neho odvodené procesy. Ako predpoklad štatistického rozdelenia rizikového faktora(výnosov) sa používa normálne rozdelenie, ktoré však nezodpovedá rozdeleniu skutočných výnosov najmä kvôli ťažkým chvostom a väčšej hodnote šikmosti skutočného rozdelenia výnosov. Niekedy sa namiesto normálneho rozdelenia používa studentovo rozdelenie. V našom prípade budeme predpokladať, že výnosy sledujú zovšeobecnené lambda rozdelenie, ktoré dokáže vhodnejšie zachytiť ťažšie chvosty a šikmosť v porovnaní s normálnym rozdelením.

Výhody Monte Carlo metódy

- Ľahké použitie a implementácia

- Možnosť modelovať akékoľvek štatistické rozdelenie výnosov aktív

Nevýhody Monte Carlo metódy:

- Náročné na technické vybavenie, hlavne pamäť a výpočtový čas
- Pomaly reaguje na turbulencie na trhu

Výhodou oproti klasickej historickej simulácii je to, že nepredpokladá rovnaké rozdelenie výnosov v jednotlivých časových oknách.

Pri aplikácii metódy Monte Carlo najskôr budeme predpokladať len jedno aktívum v portfóliu a následne aj portfólio obsahujúce niekoľko akcií, kde bude treba zohľadniť aj vzájomné väzby medzi jednotlivými aktívami.

Monte Carlo pre jedno aktívum:

Ako sme spomínali, tak najskôr sa zameriame len na výpočet VaR pre jednu akciu. Algoritmus Monte Carlo metódy v tomto prípade je nasledovný:

1. Nájsť vhodné pravdepodobnostné rozdelenie, prípadne stochastický proces pre rizikový faktor (v našom prípade výnosy modelujeme pomocou GLD rozdelenia, pre porovnanie aj pomocou normálneho rozdelenia)
2. Odhadnúť parametre predpokladaného rozdelenia výnosov z historických dát. To aké obdobie na odhad zvolíme je na posúdenie. Príliš málo dát môže VaR podhodnotiť, ale zase veľa dát môže zachytiť hektické obdobia a tie môžu VaR dlhodobo ovplyvňovať.
3. Generujeme veľký počet možných výnosov z predpokladaného rozdelenia (v našom prípade GLD a normálne rozdelenie) pre nasledujúci deň, pre ktorý chceme odhadovať VaR. Čím viac možných výnosov vygenerujeme, tým lepšie sa bude toto namodelované rozdelenie približovať skutočnému.
4. Zoradíme jednotlivé výnosy vzostupne a určíme príslušný $(1 - \alpha) * 100\%$ kvantil rozdelenia a dostaneme tak $-VaR_\alpha$ na nasledujúci deň.

Otázkou ostáva ako v kroku 3 modelovať náhodnú premennú z GLD rozdelenia s odhadnutými parametrami. Generovanie náhodných premenných pre normálne rozdelenie je všeobecne známe:

$$X = \mu + \sigma U$$

kde X je náhodná premenná z normálneho rozdelenia so strednou hodnotou μ a disperziou σ^2 a U je premenná z rovnomerného rozdelenia $R(0, 1)$. Ako však generovať náhodnú premennú z GLD rozdelenia? Pomôže nám nasledovná veta z [1].

Veta 5.2.1. *Nech $U_1, U_2, \dots$ sú nezávislé, rovnomerne rozdelené náhodné premenné na intervale $(0, 1)$, potom:*

$$X_1 = Q_X(U_1), X_2 = Q_X(U_2), \dots$$

(kde Q_X je kvantilová funkcia rozdelenia X), sú nezávislé náhodné premenné s rovnakým rozdelením ako X .

Teda predpokladajme, že máme GLD rozdelenie s parametrami:

$$\lambda_1 = -0.001425662, \quad \lambda_2 = -7.062634117, \quad \lambda_3 = -0.099887312, \quad \lambda_4 = -0.104778616$$

Potom kvantilová funkcia podľa vzťahu (2.1.1) má tvar:

$$Q(y) = Q(y, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = -0.001425662 + \frac{y^{-0.099887312} - (1 - y)^{-0.104778616}}{-7.062634117}$$

Teda ak vygenerujeme náhodné premenné z rovnomerného rozdelenia $U_1, U_2, \dots$ na intervale $(0, 1)$ tak podľa vety (5.2.1)

$$Q(U_1, \lambda_1, \lambda_2, \lambda_3, \lambda_4), Q(U_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4), \dots$$

budú náhodné premenné z $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

Monte Carlo pre portfólio akcií:

V predchádzajúcej časti sme popísali algoritmus, ako vypočítať VaR pre jednu akciu metódou Monte Carlo. Ako však tento algoritmus aplikovať pre portfólio zložené z viacerých akcií? V prípade portfólia zloženého z niekoľkých finančných aktív sa môže pracovať na pozičnom, alebo portfóliovom prístupe.

- Pri **portfóliovom prístupe** hodnota portfólia závisí iba na jednom rizikovom faktore a teda netreba modelovať závislosti medzi jednotlivými výnosmi.
- Naopak pri **pozičnom prístupe** sa hodnota portfólia odvíja od viacerých faktorov, ktoré môžu mať medzi sebou rôzne väzby - teda môžu byť medzi sebou korelované.

My sa budeme zaoberať pozičným prístupom, teda budeme zohľadňovať aj korelácie medzi aktívami, tzn. budeme generovať náhodné vektory z viacrozmerného rozdelenia. Postup výpočtu v prípade portfólia zloženého z niekoľkých akcií pri pozičnom prístupe je nasledovný:

1. Zostavíme portfólio zložené z n akcií
2. Pre jednotlivé akcie určíme ich váhu v rámci portfólia - $w_1, w_2, \dots, w_n$
3. Pomocou historických údajov získame priebeh výnosov jednotlivých akcií a pre každú z nich vypočítame:
 - (a) Vektor stredných hodnôt jednotlivých akcií v portfóliu $\mu = (\mu_1, \mu_2, \dots, \mu_n)$, teda výberovú strednú hodnotu
 - (b) Odhad variačno - kovariančnej matice (VCM) pre jednotlivé výnosy, teda maticu variancií jednotlivých výnosov a kovariancií medzi jednotlivými výnosmi akcií.
4. Generujeme náhodný vektor R z n -rozmerného náhodného rozdelenia (v našom prípade z n -rozmerného GLD rozdelenia a pre porovnanie z n -rozmerného normálneho rozdelenia s výberovou strednou hodnotou μ a variačno-kovariančnou maticou VCM). Tento postup opakujeme k -krát. Dostaneme k simulácií.
5. Pre každú získanú simuláciu z predchádzajúceho kroku vypočítame celkový výnos portfólia R_{portf} :

$$R_{portf} = w^T * R$$
6. Jednotlivé výnosy portfólia zoradíme vzostupne a určíme príslušný $(1 - \alpha) * 100\%$ kvantil rozdelenia a dostaneme tak $-VaR_{99\%}$ na nasledujúci deň.

Generovanie náhodného vektora z GLD rozdelenia: [2], [15]

Tak, ako v prípade jednej akcie sme museli generovať náhodné premenné z daného rozdelenia, tak v prípade portfólia je to podobné, no potrebujeme generovať náhodný vektor z daného rozdelenia s istým vektorom stredných hodnôt a nejakou variančno - kovariančnou maticou.

Nech vektor stredných hodnôt (μ) a variančno - kovariančná matica (VCM) sú známe. Potrebujeme získať m -rozmerný náhodný vektor X z GLD rozdelenia. V prípade

normálneho rozdelenia je postup generovania takéhoto náhodného vektora známy.

$$X = \mu + chol(VCM) * z$$

kde *chol* je choleského rozklad variančno kovariančnej matice, z je náhodný vektor zo štandardizovaného viacrozmerného normálneho rozdelenia $N_m(0, 1)$. Potom teda X je náhodný vektor z m -rozmerného normálneho rozdelenia s vektorom stredných hodnôt μ a variančno - kovariančnou maticou VCM . Generovanie náhodného vektora zo štandardizovaného viacrozmerného normálneho rozdelenia je už vo väčšine programovacích jazykoch implementovaný, tým sa teda nebudeme zaoberať. V programe R je to funkcia *rmultnorm*, ktorá je súčasťou balíka **MSBVAR** [15].

Generovanie náhodného vektora z viacrozmerného GLD rozdelenia je však o niečo zložitejšia [2]. Pred samotným generovaním potrebujeme poznať:

- Margiálne distribúcie výnosov jednotlivých akcií v tvare kvantilových funkcií.
- Rank korelácie R_X medzi jednotlivými akciami, pričom rank korelácie dostaneme nasledovným spôsobom:

$$r_{V_i, V_j} = corr[rank(V_i), rank(V_j)]$$

kde *rank* je rank - pozícia prvku, ak vzorka je usporiadaná, V_i sú jednotlivé premenné v pozorovanej vzorke.

Potom algoritmus na generovanie náhodného vektora z GLD rozdelenia je nasledovný:

1. Transformovať korelačnú maticu R_X na korelačnú maticu C_Z ako:

$$C_Z = 2 \cdot \sin\left[\frac{\pi}{6} R_X(i, j)\right]$$

2. Vygenerovať m -rozmerný náhodný vektor Z z normálneho rozdelenia s korelačnou maticou C_Z
3. Konvertovať prvky normálneho náhodného vektora Z na prvky z rovnomerného rozdelenia aplikovaním štandarnej normálnej distribučnej funkcie - Φ .

$$U_i = \Phi(Z_i)$$

Potom $U_i \sim R(0, 1)$.

4. Vypočítať X_i použitím marginálnych kvantilových funkcií:

$$X_i = Q_X(U_i)$$

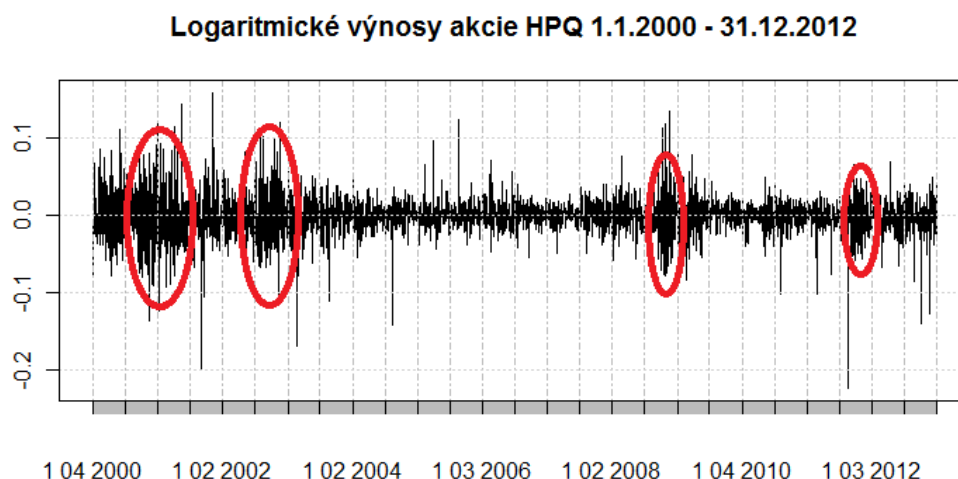
5.2.3 GARCH modely [3], [8], [12]

V predchádzajúcich častiach sme priblížili výpočet VaR pomocou historickej simulácie a pomocou klasickej metódy Monte Carlo. Nevýhodou týchto modelov je že predpokladajú v čase nemennú volatilitu, čo však pre výnosy finančných aktív neplatí. Ešte predtým ako si predstavíme samotnú GARCH metodológiu, tak spomenieme základné charakteristiky finančných aktív, na ktoré je treba myslieť pri výpočte miery rizika.

Charakteristika volatility na finančných trhoch

Volatilita na finančných trhoch má nasledovné charakteristické vlastnosti:

- **Volatility clustering**, resp. zhlukovanie volatility. Znamená to, že vysoká volatilita v čase t je často nasledovaná vysokou volatilitou v čase $t + 1$ a to isté platí aj pre nízke hodnoty volatility (viď. obrázok (5.2.3)).
- **Leverage effect**: je ďalšou charakteristickou vlastnosťou volatility a znamená to, že volatilita reaguje inak na zlé správy na trhu (pokles ceny) a inak na správy dobré (zvýšenie ceny). Pri poklese ceny aktíva je volatilita väčšia ako v prípade rastu ceny



Na modelovanie procesu, ktorý spĺňa hore uvedené vlastnosti je najvhodnejšie aplikovať GARCH metodológiu. GARCH (Generalized Autoregressive Conditional Heteroscedasticity) je model, ktorý modeluje podmienenú varianciu v čase a v porovnaní s ARCH modelmi ešte navyše zahŕňa v modeli aj vplyv minulých hodnôt volatility.

Každý GARCH model pozostáva z dvoch rovníc: jedna rovnica modeluje podmienenú strednú hodnotu, druhá modeluje podmienenú varianciu.

Rovnica podmienenej strednej hodnoty:

Na odhad podmienenej strednej hodnoty sa používa ARMA(p,q) proces. Potom rovnica pre výnosy bude mať tvar:

$$r_t = \mu + \sum_{i=1}^p \phi_i r_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t,$$

kde μ , ϕ_i , θ_j sú parametre modelu, ϵ_t sú disturbancie, r_t sú v našom prípade výnosy. To aký rád modelu, teda aké p, q zvolíme závisí od tvaru autokorelačnej (ACF) a parciálnej autokorelačnej (PACF) funkcie ARMA modelov. Najčastejšie sa používa buď AR(1), MA(1), ARMA(1,1), alebo na modelovanie výnosov sa použije čisto len konštanta + náhodná zložka (biely šum). Teda rovnice pre jednotlivé modely majú nasledovné tvary:

- **AR(1):** $r_t = \mu + \phi_1 r_{t-1} + \epsilon_t$
- **MA(1):** $r_t = \mu + \theta_1 \epsilon_{t-1} + \epsilon_t$
- **ARMA(1,1):** $r_t = \mu + \phi_1 r_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$
- **Konštanta + biely šum:** $r_t = \mu + \epsilon_t$

Premenné ϵ_t sú náhodné zložky (disturbancie, rezíduá) v čase. Predstavujú neočakávaný výnos. Mali by to byť nekorelované náhodné premenné, pre ktoré platí:

$$E(\epsilon_t) = 0$$

$$E(\epsilon_t \epsilon_s) = \begin{cases} 0, & t \neq s \\ \sigma_t^2, & t = s \end{cases}$$

Rovnica podmienenej variancie:

Podstatou GARCH modelov je práve rovnica podmienenej variancie, pretože práve táto rovnica modeluje v čase sa meniacu varianciu výnosov. Pri ARCH modeli rovnica podmienenej variancie má tvar:

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2, \text{ kde } \omega > 0, \alpha_i \geq 0$$

Pridaním q minulých podmienených variancií do predchádzajúcej rovnice dostaneme zovšeobecnený ARCH, teda GARCH model s rovnicou podmienenej variancie tvaru:

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \text{ kde } \omega > 0, \alpha_i \geq 0, \beta_j \geq 0$$

Na to, aby bol proces stacionárny, tak musí platiť:

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$$

Predstavili sme si všeobecne GARCH(p,q) model, avšak v praxi sa najčastejšie používa GARCH(1,1). Modely vyšších rádov sa používajú len veľmi zriedka. Teda pre GARCH(1,1) rovnica podmienenej variancie má tvar:

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

Na začiatku sme spomínali, že GARCH model dokáže zahrnúť jednak zhlukovanie volatility a taktiež leverage effect. Avšak čo sa týka klasického GARCH modelu, ktorý sme si predstavili, tak dokáže zachytiť iba zhlukovanie volatility. Spomínaný leverage effect, teda asymetria vo volatility výnosov sa dá modelovať pomocou dvoch asymetrických GARCH modelov, a to EGARCH (exponencial GARCH) alebo TARCH (Treshold ARCH). V našom prípade nepoužijeme ani jednu z týchto metód ale namiesto toho použijeme kombináciu metódy Monte Carlo a klasickej GARCH metódy. Myšlienka je taká, že najskôr odhadneme z historických dát GARCH model a pomocou neho namodelujeme predikcie volatility a následne budeme postupovať rovnako ako pri metóde Monte Carlo, ale namiesto variancie získanej z historických dát použijeme práve tieto predikcie získané z GARCH modelu. Pri takomto výpočte využitím GLD rozdelenia zachytíme aj asymetriu volatility výnosov a to tým, že GLD nám modeluje aj šikmosť. Teda postupovať budeme nasledovne:

1. Odhadneme GARCH modele pre historické dáta, napr. 3 roky späť.
2. Vypočítame predikcie pre volatility na obdobie odhadu VaR - σ_{GARCH}
3. Pre zvolené „okno“, v našom prípade opäť 350 dní späť určíme príslušné momenty - $\text{mean}(\mu)$, $\text{skewness}(\gamma)$, $\text{kurtosis}(\kappa)$.
4. Generujeme náhodné premenné z GLD rozdelenia so strednou hodnotou μ , varianciou σ_{GARCH} , šikmosťou γ a špicatosťou κ .

5. Určíme príslušný $(1 - \alpha)\%$ kvantil nasimulovaných výnosov.

Na odhad GARCH modelu budeme používať opäť programovacie prostredie R a balíček pre odhadovanie GARCH modelov - fGarch. Pre informácie o používaní balíka fGarch pozri [14].

5.3 Spätné testovanie VaR - Backtesting

V tejto časti si povieme ako testovať vhodnosť použitej metódy na výpočet miery rizika. Teda ak sme už na základe istého modelu odhadli vývoj miery rizika daného portfólia, môžeme pomocou backtestingu zistiť, v akej miere je tento odhad dobrý. Spomenieme dva najznámejšie testy: Kupiecov test a Christoffersenov test.

Kupiecov test: Tento test sa zameriava na pomer počtu prekročení k celkovému počtu odhadovaných VaR a testuje, či sa pohybuje dostatočne blízko hodnoty $(1-\alpha)$.

Označme počet skutočných prekročení VaR, teda také, ktoré skutočná strata presiahla (t.j. zlé predikcie) ako T . Počet všetkých odhadov VaR nech je N . Potom skutočná pravdepodobnosť prekročenia je $\hat{p} = \frac{T}{N}$. Kupiecov test je založený na porovnaní stanovenej hladiny spoľahlivosti p (teda ak počítame 99% VaR, tak $p = 0.01$) so skutočnou pravdepodobnosťou prekročenia VaR, teda s $\hat{p}$.

Teda testuje sa hypotéza: **H0 ... $\hat{p} = p$.**

Testovacia štatistika pre tento test má nasledovný tvar:

$$LRuc = -2 \ln \frac{p^n (1-p)^{T-N}}{\hat{p}^N (1-\hat{p})^{T-N}}$$

Pri platnosti H_0 má štatistika LRuc asymptotické chí-kvadrát rozdelenie s jedným stupňom voľnosti. Teda H_0 zamietame, ak hodnota štatistiky LRuc je väčšia ako kritická hodnota chí-kvadrát rozdelenia s 1 stupňom voľnosti.

Christoffersenov test:

V Kupiecovom teste sme testovali iba skutočnosť, že skutočná pravdepodobnosť prekročenia je porovnateľná so stanovenou hladinou spoľahlivosti. Pri Christoffersenovom teste navyše testujeme, či dané prekročenia sú nezávislé, pretože ak by bol model pre

výpočet VaR správny, tak prekročenia by mali byť nezávislé, tzn., že to či prekročenie nastalo v predchádzajúcom časovom okamihu, nemá vplyv na to, či nastane prekročenie v časovom okamihu nasledujúcom.

Testujeme hypotézu: **H0 ... $p = \hat{p}$ a prekročenia sú navzájom nezávislé.**

Štatistika pre Christoffersenov test je nasledovná:

$$LR_{cc} = LR_{uc} + LR_{ind},$$

LR_{uc} ... štatistika z Kupiecovho testu,

$$LR_{ind} = 2(\ln LA - \ln L0),$$

$$LA = (1 - \pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{10}} \pi_{11}^{T_{11}},$$

$$L0 = (1 - \pi)^{T_{00} + T_{10}} \pi^{T_{01} + T_{11}},$$

T_{ij} ... počet pozorovaní, ktoré boli v čase t v stave j a v čase $t - 1$ boli v stave i .

$$\pi_{01} = \frac{T_{01}}{T_{00} + T_{01}}, \pi_{11} = \frac{T_{11}}{T_{10} + T_{11}}, \pi = \frac{T_{01} + T_{11}}{T}.$$

LR_{ind} má chí-kvadrát rozdelenie s jedným stupňom voľnosti, LR_{cc} má chí-kvadrát rozdelenie s dvoma stupňami voľnosti. Teda hypotézu $H0$ zamietame ak hodnota štatistiky LR_{cc} je menšia ako kritická hodnota chí-kvadrát rozdelenia s dvoma stupňami voľnosti.

6 Praktická časť výpočtu VaR

V predchádzajúcej kapitole sme si priblížili niekoľko spôsobov, ako modelovať mieru rizika. V tejto časti využijeme tieto poznatky a ukážeme si ich praktické použitie. Budeme sa venovať hlavne metóde Monte Carlo pomocou normálneho a GLD rozdelenia pre portfólio zložené najskôr z jednej a následne z piatich akcií. V druhej časti aplikujeme kombináciu GARCH modelov s Monte Carlo metódou, avšak v tomto prípade len pre jednu akciu. Dáta pre praktickú časť sme čerpali z [16].

Pri každej metóde budeme postupovať podobným spôsobom. Na začiatku určíme dátovú vzorku, z ktorej budeme odhadovať model. Následne určíme časový úsek, pre ktorý budeme určovať out-of-sample odhady VaR, teda odhady do budúcnosti. Avšak zvolíme taký interval, pre ktorý budeme poznať skutočné výnosy, na ktoré VaR odhadujeme, aby sme mohli nakoniec vykonať backtesting. Vo všetkých metódach počítame jednodňovú VaR s hladinou spoľahlivosti 99%, takže odhadujeme $VaR_{99\%}$.

6.1 Výpočet VaR pomocou metódy Monte Carlo

Tu si ukážeme, aké výsledky dáva metóda Monte Carlo pre výpočet VaR najskôr pre jedno aktívum a následne pre portfólio zložené z piatich akcií. V oboch prípadoch budeme porovnávať prístup s predpokladom, že výnosy majú normálne rozdelenie s prístupom rozdelenia výnosov z GLD rozdelenia. Nakoniec pomocou grafu znázorníme výsledky s porovnaním so skutočnými výnosmi a vykonáme spätné testovanie a porovnáme výsledky pre normálne a GLD rozdelenie.

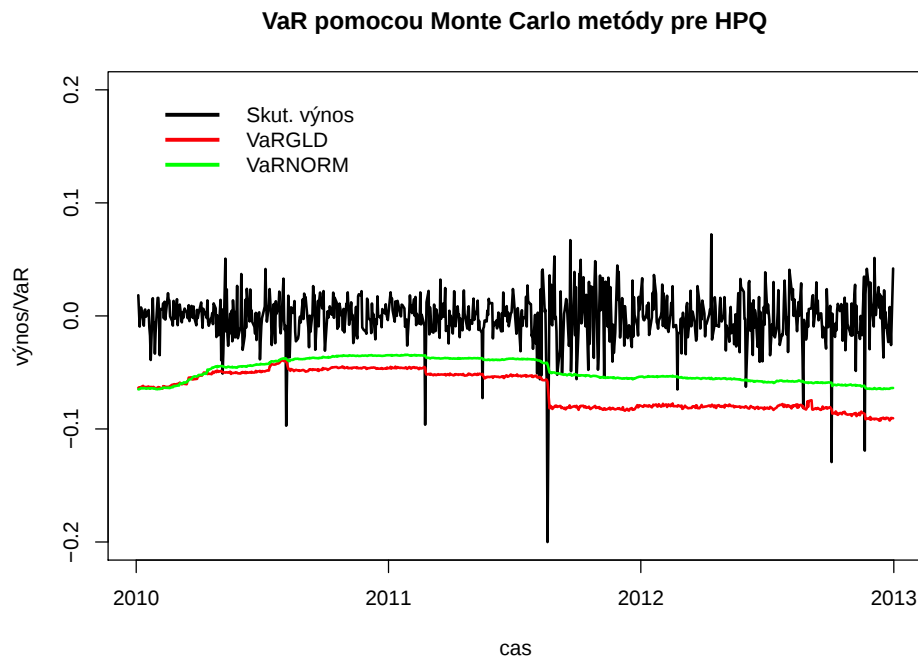
6.1.1 Výpočet VaR pre jedno aktívum

Pre túto časť sme zvolili akciu HPQ. Použili sme historické dáta od 12.8.2008-31.12.2012. Prvých 350 dát sme použili na modelovanie výnosov a výpočet prvej out-of-sample VaR na 351. deň (4.1.2010). Na výpočet hodnoty VaR na nasledujúce dni sme používali vždy historické dáta 350 dní dozadu. V každej iterácii výpočtu VaR sme najstaršie pozorovanie vynechali a pridali pozorovanie z predchádzajúceho dňa. Takto sme postupovali až do konca roku 2012, teda dostali sme tak 754 budúcich hodnôt $VaR_{99\%}$.

Výnosy sme v každom kroku modelovali dvoma spomínanými metódami, pomocou normálneho rozdelenia a pomocou GLD rozdelenia. Predpokladali sme, že VaR

vypočítaná pomocou normálneho rozdelenia bude menšia (VaR je definovaná ako kladné číslo) ako pomocou GLD rozdelenia. Je to hlavne kvôli tomu, že GLD rozdelenia dokáže lepšie zachytiť ťažké chvosty rozdelenia výnosov, teda dokáže zachytiť aj väčšie straty, resp. väčšie výnosy (nás ale zaujímajú hlavne tie väčšie straty) tzn., že maximálna možná strata pomocou GLD by mala väčšia.

Na nasledujúcom obrázku máme graf výnosov akcie HPQ v čase znázornený čiernou krivkou, zelenou krivkou je VaR vypočítané pomocou normálneho rozdelenia a červenou je VaR pomocou GLD rozdelenia. Tak ako sme predpokladali, tak VaR pomocou GLD je väčšia, teda je nižšie, ako VaR pomocou normálneho rozdelenia. Pri generovaní VaR bolo v tomto prípade pre obe metódy použitých 100000 simulácií z daného rozdelenia. Pozrime sa na pravdepodobnosti prekročenia pre jednotlivé metódy. V prípade



normálneho rozdelenia došlo k prekročeniu VaR 17 krát ($p_{\text{prekrocenia}} = 2.25$), pričom pri GLD len 10 krát ($p_{\text{prekrocenia}} = 1.32$) pri uvažovanej hladine spoľahlivosti 99%. Vidíme, že rozdiel je vcelku výrazný.

Čo nám hovoria spätné testovania?

1. Kupiecov test(p-value)

- (a) 0.298% - Normálne rozdelenie
- (b) 39.33% - GLD rozdelenie:

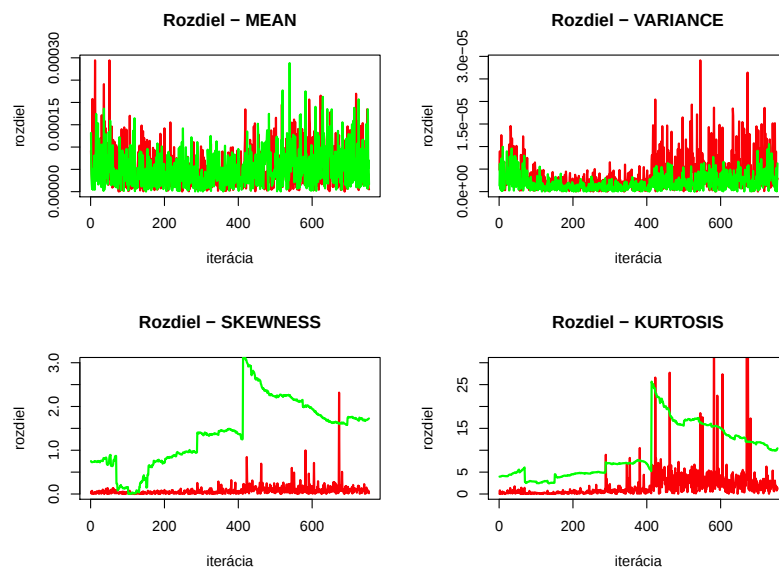
2. Christoffersenov test(p-value)

(a) 0.844% - Normálne rozdelenie

(b) 20.26% - GLD rozdelenie:

Vidíme, že v prípade GLD oba testy prijali nulovú hypotézu, v prípade normálneho rozdelenia oba testy nulovú hypotézu zamietli. Teda iba v prípade GLD rozdelenia platí, že pravdepodobnosť neprekročenia hodnoty VaR je dostatočne blízko hladiny spoľahlivosti a jednotlivé prekročenia sú nezávislé.

Pri Monte Carlo metóde výpočtu VaR sa postupuje tak, že v každom výpočte generujeme niekoľko náhodných premenných z nejakého rozdelenia. V našom prípade to bolo GLD a normálne rozdelenie. Správnosť modelu by sa dala určovať aj pomocou správnosti generovania náhodných premenných. Môžeme porovnávať skutočné momenty danej dátovej vzorky(historické výnosy) a momenty generovaných premenných v jednotlivých iteráciách výpočtu VaR. V nasledujúcich grafoch sú odchýlky jednotlivých momentov pre normálne(zelenou) a GLD rozdelenie(červenou) od skutočných momentov vzorky.



Vidíme, že rozdiely pre prvé dva momenty sú porovnateľné, avšak významnejšie momenty ako sú šikmosť a špicatosť lepšie dokáže popísať GLD rozdelenie, čo je logické, pretože GLD je štvorparametrové a berie do úvahy aj šikmosť a špicatosť na rozdiel od normálneho, ktoré má tieto dva momenty pevne dané. Pre porovnanie: 744 krát bol rozdiel v šikmosti menší pre GLD rozdelenie a 10 pre normálne. Pre špicatosť je to:

740:14. To čo najviac ovplyvňuje konce rozdelenia je práve špicatosť a teda vidíme že výnosy v tomto zmysle naozaj lepšie popisuje GLD rozdelenie a teda aj výpočet VaR je presnejší.

Nevýhodou Monte Carlo metódy je to, že veľmi pomaly reaguje na zmenu hektickejšieho obdobia na pokojnejšie. Teda ak model odhadujeme z hektického obdobia a odhadujeme VaR na pokojnejšie obdobie, tak veľmi vysoká VaR môže dominovať počas dlhšieho obdobia a naopak. Je to dôsledkom zhľukovania volatility ako sme spomínali v predchádzajúcej kapitole. Riešením tohto problému sú GARCH modely, ktoré si ukážeme neskôr.

6.1.2 Výpočet VaR pre portfólio akcií

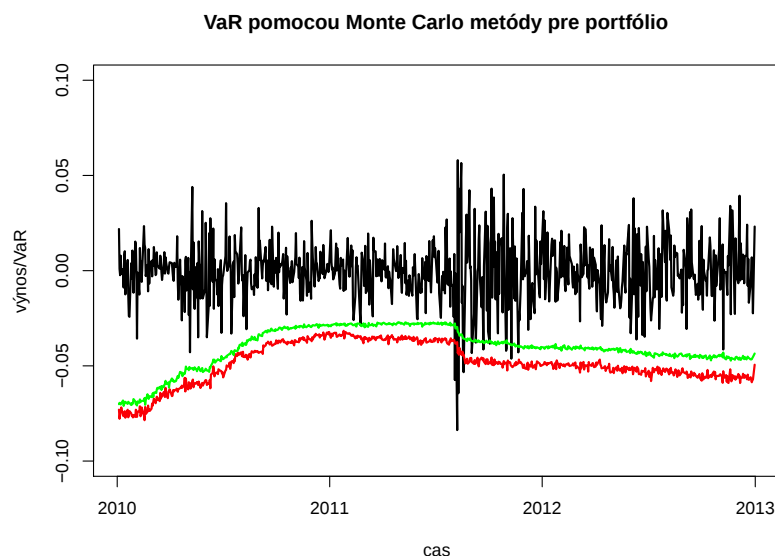
Na modelovanie VaR portfólia sme vybrali akcie:

- Hewlett-Packard Company (HPQ)
- International Business Machines Corporation (IBM)
- Nokia Corporation (NOK)
- Bank of America Corporation (BAC)
- The Coca-Cola Company (KO)

Hodnoty výnosov boli získané z historických dát z obdobia od 12.8.2008 do 31.12.2012, pričom na odhad modelu sme vždy použili historických 350 výnosov. Váhy jednotlivých akcií v našom portfóliu sme pre jednoduchosť zvolili rovnomerne, teda: $w_1 = w_2 = w_3 = w_4 = w_5 = \frac{1}{5}$.

Pri výpočte VaR sme postupovali na základe pozičného prístupu, teda zohľadnili sme aj korelácie medzi jednotlivými akciami, teda generovali sme náhodné vektory z viacrozmerného rozdelenia. Opäť predpokladáme, že VaR získaná pomocou normálneho rozdelenia bude menšia ako pomocou GLD rozdelenia.

Na nasledujúcom obrázku máme skutočné výnosy portfólia (čiernou) za obdobie od kedy sme odhadovali VaR, zelenou je VaR pomocou normálneho rozdelenia, červenou pomocou GLD rozdelenia. V každej iterácii výpočtu VaR sme použili len 10 000 (pri jednej akcii to bolo 10-krát viac) generovaní výnosov portfólia kvôli časovej náročnosti generovania náhodných vektorov z viacrozmerného rozdelenia. Z grafu teda vidíme, že



VaR pomocou GLD je zväčša nižšie ako VaR pomocou normálneho rozdelenia, čo sme aj očakávali. Pre lepšie porovnanie sme vykonali backtesting pre obe získané VaR.

1. Kupiecov test(p-value)
 - (a) 2.24% - Normálne rozdelenie
 - (b) 23.19% - GLD rozdelenie:
2. Christoffersenov test(p-value)
 - (a) 5.32% - Normálne rozdelenie
 - (b) 40.11% - GLD rozdelenie:

Teda Kupiecov test pre normálne rozdelenie zamietol nulovú hypotézu, Christoffersenov test nulovú hypotézu tesne nezamietol. Avšak pre VaR vypočítané pomocou GLD rozdelenia sú obe štatistiky podstatne lepšie ako pre normálne rozdelenie. V prípade GLD rozdelenia oba testy vyšli pozitívne. Pozrime sa konkrétne na jednu iteráciu a porovnajme momenty skutočných výnosov portfólia a výnosov vygenerovaných pomocou GLD a normálneho rozdelenia (počet iterácií - 10 000).

- Skutočné momenty: 0.0013487821, 0.0001899966, -0.2994180703 , 4.8534453724
- Momenty z GLD : 0.0013824465, 0.0002057337, -0.2334501368 , 4.8570969471
- Momenty z Norm : 0.0013444971, 0.0001911421, -0.0041223875 , 3.0115664273

Teda vidíme opäť, že prvé dva momenty približne rovnako dobre popisuje normálne aj GLD rozdelenie. Tretí a štvrtý moment podstatne lepšie zachytilo GLD rozdelenia.

6.2 VaR pomocou GARCH modelov

V tejto časti opäť aplikujeme metódu Monte Carlo, avšak v tomto prípade aj s využitím GARCH modelov. Tento model aplikujeme iba na portfólio zložené z jednej akcie.

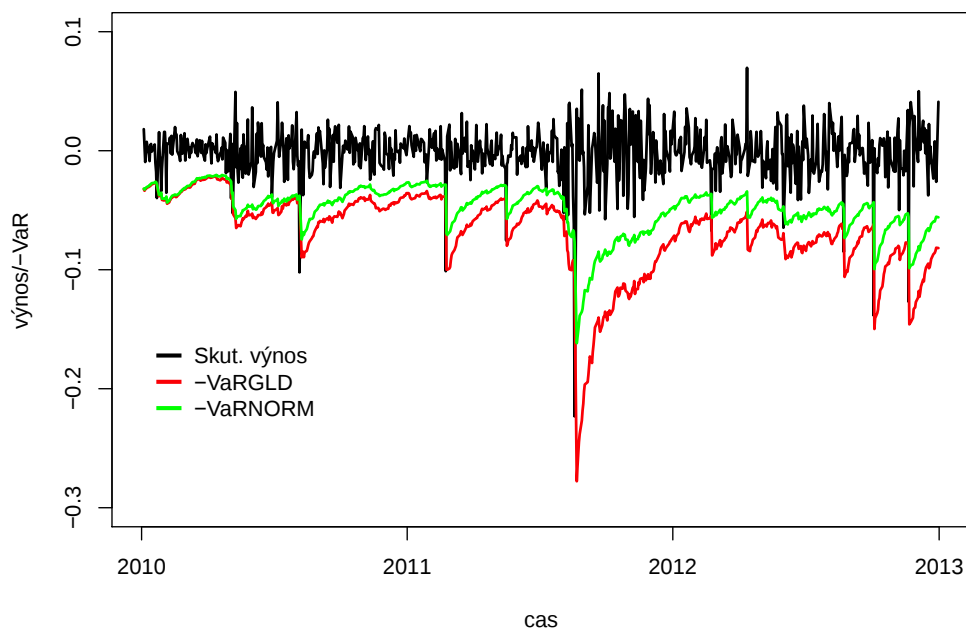
Ako sme už v teoretickej časti VaR spomenuli, tak najskôr odhadneme GARCH model pre výnosy z historických dát 3 roky dozadu. Následne určíme predikcie variancií do budúcnosti a potom postupujeme rovnako ako v prípade Monte Carlo, avšak s tým rozdielom, že namiesto variancie získanej z historických 350 dát (350 je okno aj v tomto prípade) použijeme predikciu určenú z GARCH modelu. Teda v prípade GLD rozdelenia generujeme náhodné premenné z GLD rozdelenia so strednou hodnotou určenou z historických dát, varianciou určenou z GARCH modelu, šikmosťou a špicatosťou určenými z historických dát 350 dní späť. V prípade normálneho rozdelenia je šikmosť a špicatosť daná pevne - šikmosť 0, špicatosť 3. Výhodou tohto modelu v porovnaní s klasickým Monte Carlo je v tom, že model sa dokáže rýchlejšie adaptovať na zmeny volatility výnosov. Navyše plus GLD oproti normálnemu rozdeleniu je to, že GLD dokáže zachytiť aj šikmosť a ťažšie chvosty rozdelenia výnosov.

K tejto časti sme opäť vybrali akciu HPQ a mieru rizika pre výnosy tejto akcie sme určovali na rovnaké obdobie ako v predchádzajúcom modeli, aby sme mohli jednotlivé metódy porovnať. Teda odhad VaR začína od 4.1.2010 a končí 31.12.2012. Na modelovanie volatility pomocou GARCH modelu sme použili dáta 3 roky dozadu - od 1.1.2007-31.12.2009. Na odhad strednej hodnoty, šikmosti a špicatosti sme v každej iterácii použili opäť 350 historických najnovších pozorovaní. Pre modelovanie strednej hodnoty v GARCH modeli sme sa na základe autokorelačnej a parciálnej autokorelačnej funkcie rozhodli pre model zložený iba z konštanty a náhodnej zložky. Keďže najčastejšie sa používa GARCH(1,1) model, tak na modelovanie volatility sme použili práve tento konkrétny GARCH:

$$\sigma_{t+1}^2 = \omega + \alpha\epsilon_t^2 + \beta\sigma_t^2$$

Teda na modelovanie volatility sme použili len disturbancie jeden krok dozadu a varianciu tiež iba jeden krok dozadu.

VaR pomocou GARCH pre HPQ

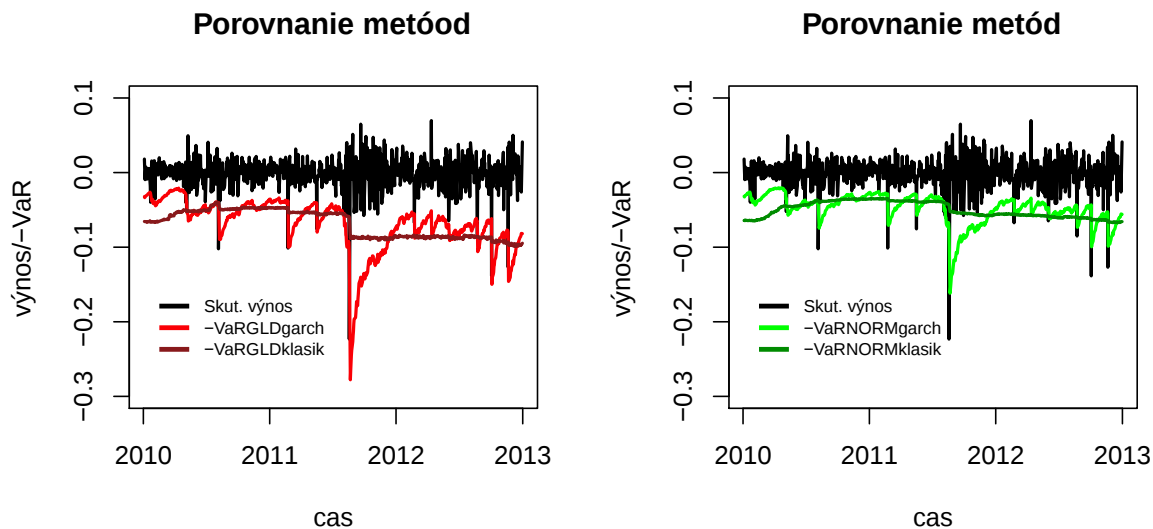


Vidíme, že opäť VaR pomocou GLD rozdelenia je väčšia ako pomocou normálneho rozdelenia. Takisto môžeme pozorovať, že tentokrát VaR dokáže lepšie reagovať na zmeny vo volatilitě ako pri klasickej Monte Carlo metóde. To spôsobuje práve volatilita modelovaná pomocou GARCH modelu. Pravdepodobnosť prekročenia v prípade GLD vyšla 1.46, pričom v prípade normálneho rozdelenia až 2.12. Pozrime sa aj na backtesting:

1. Kupiecov test(p-value)
 - (a) 0.72% - Normálne rozdelenie
 - (b) 23.75% - GLD rozdelenie:
2. Christofersenov test(p-value)
 - (a) 0.19% - Normálne rozdelenie
 - (b) 42.29% - GLD rozdelenie:

Opäť oba testy, využitím GLD rozdelenia, nulovú hypotézu (že pravdepodobnosť neprekročenia hodnoty VaR je dostatočne blízko hladiny spoľahlivosti) prijímajú, v prípade normálneho rozdelenia, naopak, oba testy nulovú hypotézu zamietajú. Teda znova prístup pomocou GLD vyšiel lepšie ako prístup s normálnym rozdelením.

Teraz porovnajme dve použité metódy. Na nasledujúcich obrázkoch máme vedľa seba porovnania oboch metód pre GLD rozdelenia a vedľa pre normálne rozdelenie. tmavšou farbou je vždy metóda klasickej Monte Carlo metódy a bledšou je 2. spomínaná metóda pomocou GARCH modelu. Vidíme teda, že pri klasickej Monte Carlo metóde



sa model nedokáže tak rýchlo adaptovať na prudké zmeny, ako model druhý, kde sme využili modelovanie volatility pomocou GARCH modelu.

V tejto metóde výpočtu VaR sme predpokladali klasický GARCH model s predpokladom normálneho rozdelenia disturbancií a až následne sme v metóde Monte Carlo využili generovanie náhodných premenných z GLD rozdelenia s využitím variancie získanej z GARCH modelu. Rozšírením tohoto modelu by bol GARCH model založený na predpoklade GLD rozdelenia disturbancií, prípadne modelovanie šikmosti a špicatosti v čase, tak ako volatilitu. Avšak aj napriek nie úplne korektným predpokladom nám dal použitý model vcelku pekné výsledky, ktoré zodpovedajú teoretickým predpokladom.

Záver:

Hlavnou náplňou našej diplomovej práce bolo predstaviť zovšeobecnené lambda rozdelenie na popis distribúcie výnosov a následne poznatky využiť na výpočet miery rizika - VaR pre akciové portfólio.

V prvých častiach sme sa venovali teórii pre zovšeobecnené lambda rozdelenie, predstavili jeho vlastnosti a odhad parametrov pomocou metódy momentov. V ďalších častiach sme sa venovali výpočtu miery rizika aplikovaním tohto rozdelenia a porovnaním s často používaným normálnym rozdelením. Prínosom tejto časti bolo, že GLD dokáže lepšie popísať distribúciu výnosov finančných aktív ako normálne rozdelenie vďaka jeho štyrom parametrom. Na rozdiel od normálneho rozdelenia dokáže zachytiť aj také vlastnosti ako sú ťažké chvosty, či šikmosť rozdelenia výnosov. Tieto vlastnosti sa podľa očakávaní prejavili vo výpočte VaR. VaR pomocou normálneho rozdelenia bola podhodnotená, teda normálne rozdelenie nedokázalo zachytiť extrémne udalosti tak dobre ako GLD rozdelenie.

Na výpočet VaR sme použili najskôr klasickú metódu Monte Carlo, ktorá nedokázala pohotovo reagovať na zmeny vo volatilitě výnosov akcií. Diplomovú prácu sme preto ešte rozšírili o výpočet VaR pomocou GARCH metodológie a to tak, že sme použili kombináciu metódy Monte Carlo a GARCH modelov, no iba pre jednu akciu. Výsledky z tejto metódy nám ukázali, že použitím GARCH modelov sme docielili to, že takýto výpočet VaR dokázal lepšie reagovať na prudké zmeny vo volatilitě, čo klasické Monte Carlo nedokázalo.

Ďalším rozvinutím našej práce by bola kombinácia Monte Carlo s GARCH modelom pre portfólio zložené z viac akcií. Pri modelovaní GARCH modelu sme predpokladali disturbancie z normálneho rozdelenia. Možnou alternatívou by bolo modelovanie GARCH s využitím rozdelenia disturbancií z GLD rozdelenia.

Literatúra

- [1] Zaven A. Karian and Edward J. Dudewicz: *Fitting Statistical Distributions, The Generalized Lambda Distribution and Generalized Bootstrap Methods*, ISBN 1-584488-069-4
- [2] Andre Lange, Christoph Sohrmann, Roland Jancke, Joachim Haase, Binjie Cheng, Urban Kovac, and Asen Asenov, *A General Approach for Multivariate Statistical MOSFET Compact Modeling Preserving Correlations*, European Solid-State Device Research Conference IEEE, Sept. 13 -15, 2011.
- [3] Krátka, J. (2006): *GARCH modely a Value-at-Risk aplikácie*. Diplomová práca, FMFI UK
- [4] Butora P. (1998): *Nelineárne rovnice a ich použitie v inžinierskych problémoch*, Diplomová práca, Žilinská univerzita
- [5] Janková, K.: *Pravdepodobnosť a štatistika 1-MAT-180*, Poznámky k prednáškam FMFI UK
- [6] Engle, R. (2002): *Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models*, In Fourthcoming Journal of Business and Economic Statistics
- [7] Lars Karlsson: *GARCH Modelling: Theoretical Survey, Model Implementation and Robustness Analysis*, Master Thesis. Dostupné na internete: <http://www.fattails.com/research/Masters Thesis - GARCH modeling.pdf>
- [8] Tesárová, V. (2012): *Value at Risk: GARCH vs. Stochastic Volatility Models: Empirical Study* Diploma thesis, Faculty of Social Sciences Institute of Economic Studies
- [9] Dennis, J.E. Jr and Schnabel, R.B. (1996), *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Siam
- [10] Berend Hasselman, (2013): *R - Package 'nleqslv' - Solve system of non linear equation*. Dostupné na internete: <http://cran.r-project.org/web/packages/nleqslv/nleqslv.pdf>

- [11] Chalabi Y., Scot D., Würtz D., *The Generalized Lambda Distribution as an Alternative to Model Financial Returns*: Dostupné na internete:
<https://www.rmetrics.org/sites/default/files/glambda.pdf>
- [12] P.Jurca, (2013): *Kvantitatívne metódy v riadení rizík - študijné materiály*, FMFI UK. Dostupné na internete:
<http://www.iam.fmph.uniba.sk/institute/jurca/qrm/qrm.htm>
- [13] Martines, J.M. (2000): *Practical quasi-Newton methods for solving nonlinear systems*. Dostupné na internete:
<http://www.ime.unicamp.br/~martinez/survcam.pdf>
- [14] Diethelm Wuertz and Yohan Chalab, (2013): *Package 'fGarch' - Autoregressive Conditional Heteroskedastic Modelling* Dostupné na internete:
<http://cran.r-project.org/web/packages/fGarch/fGarch.pdf>
- [15] Patrick T. Brandt, (2013): *Package 'MSBVAR' - Markov-Switching, Bayesian, Vector Autoregression Models* Dostupné na internete:
<http://cran.r-project.org/web/packages/MSBVAR/MSBVAR.pdf>
- [16] <http://finance.yahoo.com/>