UNIVERZITA KOMENSKÉHO V BRATISLAVE

FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

STOCHASTICKÉ MODELOVANIE VÝNOSOV POMOCOU ROZDELENÍ PRAVDEPODOBNOSTI S ŤAŽKÝMI CHVOSTAMI

Diplomová práca

2024

Bc. Adriána Leginusová

UNIVERZITA KOMENSKÉHO V BRATISLAVE

FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

STOCHASTICKÉ MODELOVANIE VÝNOSOV POMOCOU ROZDELENÍ PRAVDEPODOBNOSTI S ŤAŽKÝMI CHVOSTAMI

Diplomová práca

Študijný program:	Ekonomicko-finančná matematika a modelovanie
Študijný odbor:	Matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci:	Mgr. Gábor Szűcs, PhD.

Bratislava, 2024

Bc. Adriána Leginusová





Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Adriána Leginusová
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Stochastické modelovanie výnosov pomocou rozdelení pravdepodobnosti s ťažkými chvostami
Stochastic modeling of yields using heavy-tailed probability distributions

Anotácia: Vo viacerých vedeckých prácach bolo v minulosti empiricky ukázané, že sa denné výnosy akcií či akciových indexov neriadia podľa Gaussovho normálneho rozdelenia pravdepodobnosti. Náhodné premenné opisujúce denné výnosy či denné logaritmické výnosy spomínaných typov podkladových aktív zvyčajne majú vyššiu mieru špicatosti, a zároveň aj ťažšie chvosty ako normálne rozdelenie. Cieľom tejto diplomovej práce bude skúmať využiteľnosť vybraných rozdelení pravdepodobnosti s ťažkými chvostami pri modelovaní denných výnosov akcií, akciových indexov alebo iných podobných aktív.

Vedúci: Mgr. Gábor Szűcs, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: doc. Mgr. Igor Melicherčík, PhD.

Dátum zadania: 10.01.2023

Dátum schválenia: 11.01.2023 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
vedúci práce

Abstrakt

Volatilita je jedným z kľúčových ukazovateľov rizikovosti obchodu na finančnom trhu. Predpoklad o jej stacionarite bol vyvrátený, a navyše v mnohých prácach bola demonštrovaná jej časová závislosť a nestacionarita. Táto záverečná práca sa zaoberá modelovaním procesu volatility výmenných kurzov EUR/USD a HUF/EUR pomocou modelov stochastickej volatility s využitím rozdelení pravdepodobnosti s ľahkými a ťažkými chvostami. Výnosy v základnom modeli stochastickej volatility sa ukazujú ako leptokurtické, preto sú na modelovanie výnosov vhodnejšie alternatívy základného modelu stochastickej volatility v podobe modelovania člena chýb pomocou rozdelení pravdepodobnosti s ťažkými chvostami, medzi ktoré patria Studentovo t či GED rozdelenie (z angl. *generalized error distribution*) pravdepodobnosti. Cieľom tejto diplomovej práce je preskúmať vhodnosť a využiteľnosť rozdelení pravdepodobnosti s ťažkými chvostami na modelovanie volatility vo výnosov výmenných kurzoch s využitím Markov Chain Monte Carlo techník a Hamiltonovskej Monte Carlo metódy v prostredí softvéru R.

Kľúčové slová: volatilita, modely stochastickej volatility, rozdelenia pravdepodobnosti s ťažkými chvostami, MCMC, HMC.

Abstract

Volatility stands as a fundamental risk indicator in financial market trading. Traditional assumptions of its stationarity have been refuted, with empirical evidence consistently revealing its time-varying and non-stationary nature. This thesis deals with modeling the volatility process of EUR/USD and HUF/EUR exchange rates using stochastic volatility models that incorporate light-tailed and heavy-tailed probability distributions. The returns in the conventional stochastic volatility model are found to exhibit leptokurtosis, indicating that alternatives employing heavy-tailed distributions for the error terms, such as the Student's t distribution or the Generalized Error Distribution (GED), may offer superior modelling precision. The aim of this thesis is to explore the suitability and utility of heavy-tailed probability distributions for modeling volatility in exchange rate returns using Markov Chain Monte Carlo techniques and Hamiltonian Monte Carlo method within the R software environment.

Keywords: volatility, stochastic volatility models, heavy-tailed probability distributions, MCMC, HMC.

Predhovor

Medzi základné faktory ovplyvňujúce dynamiku finančných trhov sa zaraďuje volatilita, ktorá hrá kľúčovú úlohu pri formovaní cien a miere rizikovosti obchodu. Výskum a odhad časovo závislej volatility je preto nevyhnutnou súčasťou aktivít investora, ktorý sa zapája do obchodovania na finančnom trhu. Vzhľadom k tomu, že dnes už je téma investovania rozšírená aj medzi bežnú populáciu, najmä v podobe akciového investovania, obchodovania s kryptomenami, či dôchodkového sporenia, považujem potrebu poznania konceptu volatility ako nevyhnutnú súčasť správneho riadenia investícií. Výskum, prezentovaný v tejto práci, mi poskytol zaujímavý náhľad do témy modelovania volatility, a to najmä z matematického hľadiska, nakoľko táto práca zahŕňa spojenie niekoľkých matematických konceptov do jednej ucelenej praktickej témy, a to: makroekonomický vývoj výmenného kurzu, stochastické simulačné metódy, Markovove reťazce, bayesovskú štatistiku, diferenciálne rovnice, fyzikálne vlastnosti dynamických systémov, numerické aproximácie, či mnohé iné štatistické témy. Pevne verím, že predkladaná práca bude užitočnou pomôckou a zaujímavou publikáciou pre hĺbavého čitateľa, ktorý sa chce dozvedieť o zaujímavých technikách odhadovania volatility, no najmä metódach, ktoré ho pri čítaní privedú až na obežnú dráhu Zeme.

Touto cestou by som sa chcela poďakovať vedúcemu diplomovej práce, Mgr. Gáborovi Szűcsovi, PhD., za jeho cenné rady a námety, ktoré obohatili túto záverečnú prácu. Osobitná vďaka patrí môjmu priateľovi Martinovi, rodine, kamarátom a známym za nekonečnú podporu pri štúdiu a pri písaní tejto práce.

Obsah

Úvod **9**

1 Volatilita aktíva a jej modelovanie **12**

1.1 Základné prístupy k modelovaniu volatility 13

1.2 Vlastnosti volatility 13

2 Modely stochastickej volatility **15**

2.1 Základný model stochastickej volatility 15

2.2 Koeficient špicatosti výnosov vo všeobecnom modeli stochastickej volatility 16

2.3 Model A – Taylorov model 18

2.4 Model B – Zošikmené normálne rozdelenie 20

2.5 Model C – Studentovo t-rozdelenie 20

2.6 Model D – GED rozdelenie 22

2.7 Ekvivalentný zápis modelov 22

3 Odhadovanie parametrov modelu stochastickej volatility **25**

3.1 Bayesovský prístup k štatistike 25

3.2 Apriórne rozdelenia pravdepodobností v modeloch stochastickej volatility 27

3.3 Prehľad existujúcej literatúry 29

4 Monte Carlo techniky v modeloch stochastickej volatility **31**

4.1 Bayesovský prístup v modeloch stochastickej volatility 31

4.2 Monte Carlo integrovanie 32

4.2.1 Monte Carlo integrovanie v modeloch stochastickej volatility . . 33

5 Metóda Markov Chain Monte Carlo **34**

5.1 Markovove reťazce 34

5.1.1 Markovove reťazce s konečnou množinou stavov 34

5.1.2 Markovove reťazce so všeobecnou množinou stavov 36

5.1.3 Konvergenčná teória Markovových reťazcov 38
5.2 Princíp metódy Markov Chain Monte Carlo 39
5.3 Hamiltonovský Monte Carlo algoritmus 41
5.3.1 Princíp Hamiltonovskej Monte Carlo metódy 41
5.3.2 Hamiltonova Monte Carlo metóda v modeloch stochastickej volatility 44
5.3.3 Hamiltonovská Markovova tranzícia 47
5.3.4 Voľba kinetickej energie 49
5.3.5 Vlastnosti dynamiky Hamiltonových rovníc 49
5.3.6 Výpočet Hamiltonových rovníc 50
5.3.7 Metropolisovo-Hastingsovo akceptačné kritérium 51
5.3.8 Algoritmus HMC metódy 53
5.3.9 No-U-Turn sampler 54

6 Analýza odhadov parametrov z aposteriórneho rozdelenia 56

6.1 Konvergencia Markovovho reťazca k cieľovému rozdeleniu 56
6.2 Efektívna veľkosť vzorky 58
6.3 Odhad štandardnej chyby pre MCMC 61

7 Porovnanie a testovanie modelov stochastickej volatility 62

7.1 Testovanie 62
7.2 Informačné kritériá 63
7.2.1 Informačné kritérium DIC 64
7.2.2 Informačné kritérium WAIC 65

8 Modelovanie stochastickej volatility 67

8.1 Rozhranie RStan 67
8.2 Techniky matematického modelovania 70

8.3 Postup aposteriórnej analýzy výsledkov 71
8.4 Modelovanie výmenného kurzu EUR/USD v roku 2022 72
8.5 Modelovanie výmenného kurzu EUR/USD v roku 2023 80
8.6 Zhrnutie výsledkov modelovania pre výmenný kurz EUR/USD 82
8.7 Modelovanie výmenného kurzu HUF/EUR v roku 2023 85
8.8 Zhrnutie výsledkov modelovania pre výmenný kurz HUF/EUR 90

Záver **93**

Zoznam použitej literatúry **98**

Prílohy

Úvod

Mnohé matematické teórie sa opierajú o prísne predpoklady, ktoré sú v porovnaní s realitou zjednodušené alebo dokonca nereálne. Medzi takéto teoretické aparáty patrí aj Blackov-Scholesov model oceňovania derivátov, ktorý je v praxi využívaný pre svoju efektivitu a jednoduchosť ocenenia európskych opcií, avšak ktorý je aj kritizovaný pre silný predpoklad konštantnosti a časovej nezávislosti volatility. Tento predpoklad má len málo spoločného s reálnym trhom, čo spochybnili aj samotní autori Black a Scholes, ktorí vo svojej publikácií [5] upozornili na potrebu ďalšieho výskumu v oblasti nestacionarity volatility.

V predkladanej diplomovej práci sa venujeme problematike časovej premenlivosti volatility, ktorá je skúmaná a široko využívaná na finančných trhoch ako nástroj odhadovania miery rizikovosti investičných aktív. Predstavíme si modely stochastickej volatility, ktoré sú účinnými nástrojmi na odhadovanie volatility, ktoré sú založené na predpoklade, že sa proces volatility riadi podľa nepozorovateľného náhodného procesu [4]. Dokážeme vlastnosť špicatosti výnosov pre všeobecný model stochastickej volatility a predstavíme jeho ďalšie, v praxi využiteľnejšie alternatívy. Teoretická časť poskytuje náhľad do simulačných Monte Carlo metód, kde čitateľovi priblížime základný prístup odhadovania parametrov pre modely stochastickej volatility pomocou bayesovského prístupu ku štatistike využitím Markov Chain Monte Carlo simulácií. Teóriu Markov Chain Monte Carlo metód postavíme na základoch Markovových reťazcov a na, v súčasnosti najefektívnejšom, Hamiltonovskom Monte Carlo algoritme a jeho vylepšenej NUTS verzii. Praktická časť výskumu pozostáva zo spracovania log-výnosov výmenných kurzov eura a amerického doláru (označujeme ako EUR/USD) a výmenných kurzov maďarského forintu a eura (značíme ako HUF/EUR), tvorby matematických modelov v softvéri R [24] využitím balíčka `rstan` [27], a ich následnej validácie a porovnávania.

V prvej kapitole vysvetlíme potrebu matematického modelovania volatility z praktického hľadiska a uvedieme základné prístupy k jej modelovaniu. Predstavíme si niektoré z vlastností, ktoré sú pre volatilitu typické a ktoré vyžadujeme od našich modelov, pričom vychádzame najmä z publikácií [6], [13], [26], či [32].

V druhej kapitole si predstavíme koncept modelov stochastickej volatility, zadefinujeme si základný, Taylorov model, navrhnutý britským štatistikom a ekonómom

Stephenom J. Taylorom v publikácií [37] z roku 1982, ktorý je založený na opísaní volatility pomocou triedy tzv. modelov stavového priestoru (angl. *state-space models*). Odvodíme koeficient špicatosti výnosov pre všeobecný model stochastickej volatility, na ktorom demonštrujeme potrebu zavedenia ďalších – vhodnejších alternatív k základnému, Taylorovmu modelu. Vymedzíme si pojem rozdelení pravdepodobnosti s ťažkými chvostami a zadefinujeme si jednotlivé modely, spolu s ich vlastnosťami vzhľadom na hodnoty parametrov, pričom vychádzame z publikácií [10], [18], [19], [40].

V tretej kapitole načrtneme prístup k odhadovaniu parametrov modelov stochastickej volatility, čitateľovi priblížime bayesovský prístup k štatistike, vymedzíme si pojmy apriórneho a aposteriórneho rozdelenia pravdepodobnosti a v tomto prístupe zadefinujeme apriórne rozdelenia pravdepodobností parametrov modelov a hodnoty ich parametrov, tzv. hyperparametrov, pričom vychádzame z článkov [10], [18], [19], [20], [29], [40]. História a prístupy k modelovaniu stochastickej volatility zhrnieme v podkapitole 3.3.

Základy Monte Carlo techník si opíšeme v štvrtej kapitole, kde si zadefinujeme bayesovský prístup k štatistike v modeloch stochastickej volatility a predstavíme si základnú teóriu Monte Carlo integrovania, ktorá je užitočná pri výpočte odhadov parametrov získaných touto simulačnou technikou.

Kľúčovou súčasťou tejto diplomovej práce je piata kapitola, ktorá priblíži čitateľovi princíp Markov Chain Monte Carlo metódy, ktorý vybudujeme na základoch Markovových reťazcov a Hamiltonovskej Monte Carlo metóde. Hamiltonovskú metódu, ktorá je založená na fyzikálnych princípoch dynamických systémov, rozšírime podľa publikácie [2] do Hamiltonovskej Monte Carlo metódy, ktorá efektívne poskytuje vzorky z aposteriórneho rozdelenia pravdepodobnosti, z ktorého sa nedá generovať priamo. Pre Hamiltonovskú Monte Carlo metódu predstavíme jej vylepšenú, dynamickú modifikáciu v podobe NUTS algoritmu, ktorý je zabudovaný v knižnici Stan [27], ktorá je pomocou balíčka `rstan` implementovaná v softvéri R [24].

V piatej a šiestej kapitole si predstavíme kritériá a ukazovatele, na základe ktorých je možné posudzovať konvergenciu a kvalitu odhadov jednotlivých parametrov. Splnenie predpokladov pre jednotlivé modely testujeme, pričom metodiku a jednotlivé štatistické testy zhrnieme v siedmej kapitole. Modely z hľadiska kvality, t. j. ako dobre dokážu

popísať modelované procesy, a komplexnosti modelov v závislosti od počtu parametrov porovnávame na základe informačných kritérií DIC a WAIC [10], [39].

Ôsma kapitola zahŕňa praktickú časť tejto práce, teda modelovanie stochastickej volatility pomocou balíčka `rstan`. V tejto kapitole si vysvetlíme proces spracovania dát, techniky matematického modelára a predstavíme si syntax funkcie `stan`, ktorá obsahuje praktickú tvorbu modelu v blokoch, ktoré knižnica Stan vykonáva pomocou jazyka C++. Pre vybrané časové okná výmenných kurzov EUR/USD a HUF/EUR si zhrnieme výsledky modelovania, ktoré analyzujeme z hľadiska konvergencie a chyby odhadov. Pre jednotlivé modely zároveň testujeme splnenie ich predpokladov a na záver ich porovnáme na základe informačných kritérií. Výsledky modelovania adekvátne vyhodnotíme, pričom v prílohách B a C prezentujeme aj tie výsledky, ktoré neboli v hlavnej časti tejto diplomovej práce publikované.

1 Volatilita aktíva a jej modelovanie

Typickou vlastnosťou cien podkladových aktív je ich premenlivosť v čase, ktorá, okrem iného, súvisí tiež s volatilitou. V teórii finančnej matematiky označujeme volatilitu ako štandardnú odchýlku σ náhodných fluktuácií ceny podkladového aktíva. Ak má aktívum tendenciu výrazne meniť svoju cenu, nazývame ho volatilné. Volatilita súčasne predstavuje mieru rizika, pretože s vyššou volatilitou sa zvyšuje aj rizikovosť investície [6].

Medzi najzásadnejšie objavy v oblasti finančnej matematiky sa jednoznačne zaraďuje Blackov-Scholesov model (skr. B-S model) oceňovania finančných derivátov, ktorý je dôležitým nástrojom pre oceňovanie a riadenie rizík finančných derivátov. Z hodín finančnej matematiky však vieme, že B-S model má niekoľko dôležitých predpokladov opisujúcich „ideálny“ trh, ktorý má so skutočnosťou len málo čo spoločné [33]. Jedným z predpokladov je konštantnosť a časová nezávislosť volatility, ktorá vychádza z predpisu ceny, ktorá sa podľa predpokladov B-S modelu riadi geometrickým Brownovým pohybom. Predpoklad o stacionarite volatility spochybnili aj samotní priekopníci moderného finančníctva Black a Scholes, ktorí vyjadrili v roku 1972 nesúhlas s nerealistickým predpokladom stacionárnej volatility [26] podľa [5]: *„ . . . there is evidence of nonstationarity in the variance. More work must be done to predict variances using the information available.“*. Zjednodušenie reality v podobe stacionárnej volatility nám dáva možné vysvetlenie nesúladu medzi cenou opcie, ktoré sa vyskytujú pri B-S modeli [26]. V praxi existujú techniky „hrubého“ odhadu volatility, medzi ktoré sa zaraďuje napríklad princíp implikovanej volatility, ktorý odhadne volatilitu porovnaním teoretickej (ocenenej) hodnote aktíva a jej aktuálnej trhovej ceny [33]. Ako sa však ukazuje, cena akcií môže byť ovplyvnená napríklad očakávaniami o budúcej cene, čo môže vychýliť odhad implikovanej volatility.

V skutočnosti môže mať samotná volatilita fluktuácie, spôsobené napríklad zmenami na trhu, čím dokáže ovplyvňovať cenu daného aktíva. Modelovanie samotnej volatility zohráva preto dôležitú úlohu vo finančníctve pre presnejšie odhadnutie budúcej rizikovosti obchodu.

1.1 Základné prístupy k modelovaniu volatility

Vo všeobecnosti sú známe dva prístupy modelovania volatility [6], [26], [32]. Prvým z nich sú autoregresné modely ARCH (skr. z angl. *autoregressive conditional heteroskedasticity*), či zovšeobecnený GARCH model (skr. z angl. *generalized autoregressive conditional heteroskedasticity*). Základný ARCH model je model typu *one-step-ahead*, čo predstavuje predikciu variancie (matematicky označujeme ako σ^2) na jedno časové obdobie dopredu využívajúc predošlé pozorovania výnosov. Zovšeobecnením tohto prístupu sa dostávame ku GARCH modelom, ktoré na predikciu variancie na jedno časové obdobie dopredu využíva okrem predošlých pozorovaní výnosov aj samotné predošlé odhady variancie. Tieto typy modelov sú obľúbené najmä v oblasti riadenia rizík, pretože dokážu vhodne zachytiť zmeny časovo-závislých štruktúr dát.

Druhý prístup modelovania volatility poskytujú modely stochastickej volatility (matematicky označujeme ako σ) (skr. SV), ktoré modelujú volatilitu ako prvok nasledujúci nepozorovateľný náhodný proces, ktorý sa v teórii modelov stochastickej volatility nazýva tiež latentný náhodný proces [4]. Tento prístup poskytuje viacero výhod oproti klasickým (G)ARCH modelom, medzi ktoré patrí napríklad možnosť spojitého modelovania volatility, ktorý sa využíva najmä vo finančníctve, či zovšeobecnenie jednorozmerného modelovania časových radov na viacrozmerné. V predkladanej práci sa venujeme problematike odhadovania volatility diskrétnymi (autoregresnými) technikami, ktorú bližšie popíšeme v kapitole 2.

1.2 Vlastnosti volatility

Volatilita je objektom výskumu už dlhé desaťročia, ktoré viedli k pozorovaniu a potvrdeniu známych vlastností, ktoré sú pre volatilitu typické (angl. *stylized facts*). Vhodnosť modelov volatility je podmienená schopnosťou dostatočného zachytenia aspoň niektorých týchto vlastností, medzi ktoré patria [6], [13]:

i) **Zhlukovanie volatility** (z angl. *clustering*) predstavuje takú vlastnosť, kedy veľké zmeny ceny aktíva sú nasledované veľkými zmenami a takisto, malé zmeny ceny aktíva sú nasledované malými zmenami. Tieto zmeny môžu byť ľubovoľného charakteru, či už sa jedna o kladnú zmenu ceny v podobe jej nárastu, alebo

negatívnu zmenu, ktorá predstavuje pokles ceny.

ii) **Pretrvávanie volatility** (z angl. *persistence*) predstavuje takú vlastnosť volatility, kde jej dnešná hodnota ovplyvňuje budúce očakávania o hodnote volatility do budúcnosti. Táto vlastnosť (perzistencia) úzko súvisí s vlastnosťou zhlukovania, pretože zhluk, bez ujmy na všeobecnosti, s vysokými zmenami ceny má tendenciu pretrvávania na viac období dopredu.

iii) **Priťahovanie k dlhodobo ustálenej hodnote** (z angl. *mean-reversion*) rovnako ako perzistencia súvisí so zhlukovaním volatility. Obdobie vysokej volatility je nasledované obdobím, kedy sa volatilita ustáli na akejsi „normálnej“ hodnote, rovnako, obdobie nízkych hodnôt volatility nahradí obdobie rastu. Vlastnosť priťahovania k dlhodobo ustálenej hodnote teda predstavuje takú hodnotu, ku ktorej sa volatilita vráti po období zmien. V praxi vyžadujeme od modelovania budúcej volatility (predpovede) konvergenciu k tejto ustálenej hodnote, ktorá nie je nutne konštantná a v priebehu času sa môže meniť.

iv) **Externý vplyv** na volatilitu súvisí s informáciami na trhu podkladového aktíva. Je zrejmé, že ceny aktíva sa nevyvíjajú len z historických cien a istej náhodnosti, ale sú podmienené situáciou na trhu, politikou či inými udalosťami. Spomenuté vlastnosti i) až iii) predstavovali všeobecné empirické znalosti, ktoré boli v minulosti dokázané, avšak platí, že zväčša bývajú podmienené práve externými vplyvmi. Externé vplyvy v prípade výmenných kurzov ako podkladového aktíva môžu byť obchodná bilancia daných dvoch krajín, geopolitická situácia, menová politika centrálnych bánk, hospodárske faktory akými sú nezamestnanosť alebo inflácia, či samotná volatilita komodít akými sú plyn, ropa, zlato a iné.

2 Modely stochastickej volatility

V tejto kapitole si zadefinujeme základnú terminológiu spojenú s modelmi stochastickej volatility. Pre správne pochopenie konceptu modelov stochastickej volatility je potrebné porozumieť ich vzťahom a jednotlivým premenným. V nasledujúcej podkapitole si predstavíme ideu základného tvaru modelu stochastickej volatility, ktorý následne rozšírime do ďalších, zložitejších modelov.

Základom pre modelovanie volatility sú výnosy podkladového aktíva y_t (y z angl. *yield*), ktoré predstavujú relatívnu zmenu ceny ľubovoľného aktíva za skúmanú jednotku času t z množiny časov T, pričom T typicky volíme ako $T = \{1, 2, \ldots, n\}$, kde $n \in \mathbb{N}$ označuje počet sledovaných výnosov. V predkladanej práci sa zaoberáme modelovaním denných logaritmických výnosov výmenných kurzov upravených o strednú hodnotu, ktoré pre jednoduchosť nazývame aj ako *log-výnosy* a sú definované vzťahom:

$$y_t = \ln\left(\frac{p_t}{p_{t-1}}\right) - \bar{y} \quad \text{pre } t \in T,$$

kde p_t predstavuje cenu podkladového aktíva v čase $t \in T$, p_0 značí cenu podkladového aktíva na samom začiatku sledovaného časového obdobia a $\bar{y}$ označuje priemer log-výnosov daný vzťahom:

$$\bar{y} = \frac{1}{n} \sum_{t \in T} \ln\left(\frac{p_t}{p_{t-1}}\right).$$

2.1 Základný model stochastickej volatility

Nech $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$ je vektor log-výnosov podkladového aktíva, pričom na každom potrebnom mieste pripúšťame aj existenciu $(n+1)$. log-výnosu, ako aj existenciu $(n+1)$. členov niektorých ďalších, pomocných procesov (prezentovaných v tejto záverečnej práci). Poznamenávame, že totožný zápis y_t, pre $t \in T$, používame na označenie rôznych typov veličín: či už ide o neznámy log-výnos (definovaný na začiatku druhej kapitoly tejto práce), matematicky modelovanú náhodnú premennú y_t (pozri nižšie) alebo pozorované (číselné) hodnoty denných log-výnosov, pre jednoduchosť vo väčšine prípadov používame to isté označenie (y_t, $t \in T$). Ak to situácia vyžaduje, práve používaný „typ“ veličiny y_t upresňujeme slovným opisom. Ďalej, nech σ_t označuje volatilitu log-výnosu v čase $t \in T$. Základný model stochastickej volatility má

tvar [6], [40]:

$$\begin{aligned} y_t &= \sigma_t\, \epsilon_t, \\ \ln \sigma_{t+1}^2 &= \mu + \phi \left(\ln \sigma_t^2 - \mu \right) + \eta_t, \\ \epsilon_t &\sim \text{iid}, \, \text{E}(\epsilon_t) = 0, \, \text{D}(\epsilon_t) = 1, \\ \eta_t &\overset{\text{iid}}{\sim} \mathcal{N}\left(0;\, \sigma_\eta^2\right), \end{aligned}$$

pričom predpokladáme, že rozdelenia chýb (nazývané aj šum) ϵ_t a η_t sú nezávislé, rovnako rozdelené náhodné premenné (skr. iid z angl. *independent identically distributed*). Od latentného procesu log-variancie $\ln\left(\sigma_t^2\right)$ vyžadujeme, aby bol stacionárny, čo zabezpečíme podmienkou $|\phi| < 1$. V praxi sa často stretávame s ekvivalentným zápisom modelu pomocou označenia:

$$h_t = \ln \sigma_t^2, \tag{1}$$

ktoré predstavuje *proces volatility* (presnejšie proces log-variancie), čím dostávame [10], [20]:

$$\begin{aligned} y_t &= \exp\left(\frac{h_t}{2}\right)\epsilon_t, \\ h_{t+1} &= \mu + \phi\,(h_t - \mu) + \eta_t, \end{aligned}$$

pričom rekurzívny proces volatility $\boldsymbol{h} = (h_1, \ldots, h_n)^\top$ je inicializovaný počiatočnou podmienkou získanou z nepodmieneného rozdelenia h_t:

$$h_1 \sim \mathcal{N}\left(\mu;\, \frac{\sigma_\eta^2}{1-\phi^2}\right). \tag{2}$$

Označme parametrický priestor základného modelu stochastickej volatility symbolom Θ. Tento priestor obsahuje parametre $\boldsymbol{\theta} = \left(\mu, \phi, \sigma_\eta\right)^\top$, kde $\mu \in \mathbb{R}$ predstavuje *úroveň* procesu volatility h_t, parameter $\phi \in (-1, 1)$ jeho *perzistenciu* a $\sigma_\eta^2 > 0$ predstavuje *volatilitu* samotného procesu volatility h_t.

2.2 Koeficient špicatosti výnosov vo všeobecnom modeli stochastickej volatility

Všeobecne známym faktom je, že denné výnosy sa zvyčajne riadia leptokurtickým rozdelením pravdepodobnosti, čo znamená, že ich koeficient špicatosti je väčší ako

koeficient špicatosti takej náhodnej premennej, ktorá sa riadi podľa normálneho rozdelenia pravdepodobnosti. Pre všeobecný model s bližšie nešpecifikovaným rozdelením šumu $\epsilon_t \sim \text{iid}\,(0;\,1)$ vypočítame koeficient špicatosti výnosov pomocou vzťahu [40]:

$$k = \frac{\mathrm{E}\left[y_t^4\right]}{\left(\mathrm{E}\left[y_t^2\right]\right)^2}.$$

Strednú hodnotu nepodmieneného rozdelenia procesu volatility získame úpravou jeho predpisu takýmto spôsobom:

$$\begin{aligned}\mathrm{E}\left[h_{t+1}\right] &= \mathrm{E}\left[\mu + \phi\left(h_t - \mu\right) + \eta_t\right],\\ \mathrm{E}\left[h_{t+1}\right] &= \mu + \phi\mathrm{E}\left[h_t\right] - \phi\mu + \mathrm{E}\left[\eta_t\right].\end{aligned}$$

Ak uvažujeme nepodmienené rozdelenie náhodnej premennej h_t, tak za predpokladu nezávislostí náhodných premenných $h_1, h_2, \ldots, h_t$ pre $t \in T$ platí, že všetky náhodné premenné $h_1, h_2, \ldots, h_t$ majú rovnaké nepodmienené rozdelenie pravdepodobnosti, pričom platia vzťahy $\mathrm{E}\left[h_{t+1}\right] = \mathrm{E}\left[h_t\right]$ a $\mathrm{D}\left(h_{t+1}\right) = \mathrm{D}\left(h_t\right)$ pre $\forall t \in T$. Využitím tejto skutočnosti dostávame:

$$\begin{aligned}\mathrm{E}\left[h_t\right] &= \mu + \phi\mathrm{E}\left[h_t\right] - \phi\mu + \mathrm{E}\left[\eta_t\right],\\ \left(1-\phi\right)\mathrm{E}\left[h_t\right] &= \left(1-\phi\right)\mu,\\ \mathrm{E}\left[h_t\right] &= \mu.\end{aligned}$$

Varianciu nepodmieneného rozdelenia procesu volatility

$$\mathrm{D}\left(h_t\right) = \frac{\sigma_\eta^2}{1-\phi^2}$$

získame obdobným spôsobom využitím vlastnosti variancie, pričom dôkaz je možné spraviť napríklad odvodením vzťahu pre druhý počiatočný moment $\mathrm{E}\left[h_t^2\right]$, a následným využitím formuly $\mathrm{D}\left(h_t^2\right) = \mathrm{E}\left[h_t^2\right] - \left(\mathrm{E}\left[h_t\right]\right)^2$.

Normalita nepodmieneného procesu volatility h_t vyplýva zo základnej vlastnosti autoregresného modelu (skr. AR z angl. *autoregressive*), modelu kĺzavých priemerov (skr. MA z angl. *moving-average*), či autoregresného modelu s kĺzavými priemermi (skr. ARMA z angl. *autoregressive moving-average*), pretože ak si uvedomíme, že predpis časového radu h_t môžeme interpretovať ako AR(1) proces s gaussovským bielym šumom

$$h_{t+1} = \mu + \phi\left(h_t - \mu\right) + \eta_t, \quad \eta_t \sim \mathcal{N}\left(0;\,\sigma_\eta^2\right),$$

pre ktorý platí, že nepodmienené rozdelenie náhodnej premennej h_t pre $\forall t \in T$ má Gaussovo normálne rozdelenie pravdepodobnosti [36], tak dostávame nepodmienené rozdelenie pravdepodobnosti náhodnej premennej h_t

$$h_t \sim \mathcal{N}\left(\mu;\, \frac{\sigma_\eta^2}{1-\phi^2}\right).$$

Na výpočet využijeme predpoklad o nezávislosti η_t a ϵ_t a vlastnosť strednej hodnoty log-normálneho rozdelenia, čím dostávame:

$$\mathrm{E}\left[y_t^4\right] = \mathrm{E}\left[\mathrm{e}^{2h_t}\epsilon_t^4\right] = \mathrm{E}\left[\mathrm{e}^{2h_t}\right]\mathrm{E}\left[\epsilon_t^4\right] = \mathrm{e}^{\left(2\mu+2\frac{\sigma_\eta^2}{1-\phi^2}\right)}\mathrm{E}\left[\epsilon_t^4\right], \tag{3}$$

$$\begin{aligned}\left(\mathrm{E}\left[y_t^2\right]\right)^2 &= \left(\mathrm{E}\left[\mathrm{e}^{h_t}\epsilon_t^2\right]\right)^2 = \left(\mathrm{E}\left[\mathrm{e}^{h_t}\right]\right)^2\left(\mathrm{E}\left[\epsilon_t^2\right]\right)^2 = \mathrm{e}^{\left(2\mu+\frac{\sigma_\eta^2}{1-\phi^2}\right)}\left(\mathrm{E}\left[\epsilon_t^2\right]\right)^2,\\ &= \mathrm{e}^{\left(2\mu+\frac{\sigma_\eta^2}{1-\phi^2}\right)}\left(\mathrm{D}\left(\epsilon_t\right)+\mathrm{E}^2\left[\epsilon_t\right]\right)^2 = \mathrm{e}^{\left(2\mu+\frac{\sigma_\eta^2}{1-\phi^2}\right)}.\end{aligned} \tag{4}$$

Spojením (3) a (4) získavame koeficient špicatosti modelovaných log-výnosov y_t, $t \in T$, ktoré sa riadia podľa základného modelu definovaného v podkapitole 2.1:

$$k = \frac{\mathrm{E}\left(y_t^4\right)}{\left[\mathrm{E}\left(y_t^2\right)\right]^2} = \mathrm{E}\left[\epsilon_t^4\right]\mathrm{e}^{\left(\frac{\sigma_\eta^2}{1-\phi^2}\right)}. \tag{5}$$

2.3 Model A – Taylorov model

V praxi sa stretávame s variáciami základného modelu stochastickej volatility. Ak požadujeme normalitu rozdelenia šumu $\epsilon_t \sim \mathcal{N}\left(0;\, 1\right)$ pre všetky $t \in T$, tak dostávame najznámejší [10], [19], [20], [29], [40], a zároveň aj najjednoduchší, takzvaný Taylorov model stochastickej volatility publikovaný v roku 1982 ekonómom a štatistikom Stephenom J. Taylorom [37], čím dostávame prvý zo štyroch analyzovaných modelov tejto záverečnej práce:

$$\begin{aligned} y_t &= \exp\left(\frac{h_t}{2}\right)\epsilon_t,\\ h_{t+1} &= \mu + \phi\left(h_t-\mu\right)+\eta_t,\\ \epsilon_t &\sim \mathcal{N}\left(0;\, 1\right),\\ \eta_t &\sim \mathcal{N}\left(0;\, \sigma_\eta^2\right),\\ \rho\left(\epsilon_t,\, \eta_t\right) &= 0,\, \forall t \in T, \end{aligned}$$

kde $\rho(\cdot, \cdot)$ označuje Pearsonov korelačný koeficient. Pre štvrtý štandardizovaný moment Taylorovho modelu s Gaussovým rozdelením pravdepodobnosti chybového člena $\epsilon_t \sim \mathcal{N}(0; 1)$ platí $\mathrm{E}\left[\epsilon_t^4\right] = 3$, čím dostávame nerovnosť

$$k = \mathrm{E}\left[\epsilon_t^4\right] \mathrm{e}^{\left(\frac{\sigma_\eta^2}{1-\phi^2}\right)} = 3\mathrm{e}^{\left(\frac{\sigma_\eta^2}{1-\phi^2}\right)} \geq 3,$$

ktorý je rovný trom vtedy a len vtedy, ak $\frac{\sigma_\eta^2}{1-\phi^2} = 0$. V prípade časovo závislej volatility preto platí, že nepodmienené rozdelenie log-výnosov y_t môže byť leptokurtické aj v prípade normality šumu ϵ_t.

Mnohé štúdie ukázali, že špicatosť výnosov nebola dostatočne vhodne zohľadnená Taylorovým modelom, preto boli navrhnuté ďalšie modely s mnohými ďalšími rozdeleniami šumu ϵ_t, ktoré boli bohatšie na počet parametrov alebo na ťažšie chvosty, pričom v praxi sa tieto modely ukázali ako vhodnejšie alternatívy Taylorovho modelu [10], [40].

V našej práci uvažujeme pre prípad log-výnosov s asymetrickým rozložením okolo svojej strednej hodnoty zošikmené normálne rozdelenie; pre extrémnejšie situácie uvažujeme rozdelenia pravdepodobnosti s ťažkými chvostami, v našom prípade Studentovo t-rozdelenie a GED rozdelenie (z angl. *generalized error distribution*) nazývané aj ako zovšeobecnené chybové rozdelenie.

Voľne povedané, pod pojmom *rozdelenia pravdepodobnosti s ťažkými chvostami* uvažujeme také distribúcie, ktoré majú ťažšie (angl. *heavy tails*), hrubšie (angl. *fat tails*) či dlhšie (angl. *long tails*) chvosty v porovnaní s normálnym rozdelením. Trieda ťažkých chvostov teda pozostáva z ďalších podtried, ktoré sú špecifickejšie v porovnaní so všeobecnou triedou ťažkých chvostov [21]. Jednotlivé podtriedy sa zväčša líšia len technickými definíciami, či momentami, ktoré môžu byť napríklad v prípade hrubých chvostov nekonečné, a to aj v prípade parametra polohy [8], [11]. Motiváciou tejto práce je skôr praktické využitie vlastnosti rozdelení s ťažkými chvostami než výskum ich technických charakteristík.

2.4 Model B – Zošikmené normálne rozdelenie

Tretí mode so zošikmeným štandardizovaným normálnym rozdelením má tvar

$$\begin{aligned} y_t &= \exp\left(h_t/2\right)\epsilon_t, \\ h_{t+1} &= \mu + \phi\left(h_t - \mu\right) + \eta_t, \\ \epsilon_t &\sim \widetilde{\mathcal{N}_\nu}(0;\, 1), \\ \eta_t &\sim \mathcal{N}(0;\, \sigma_\eta^2), \\ \rho\left(\epsilon_t,\, \eta_t\right) &= 0,\, \forall t \in T, \end{aligned}$$

kde $\widetilde{\mathcal{N}_\nu}$ označuje zošikmené štandardizované normálne rozdelenie s parametrom tvaru $\nu \in \mathbb{R}$. Funkcia hustoty má tvar

$$p(\epsilon_t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\epsilon_t^2}{2}\right)\left(1 + \operatorname{erf}\left[\nu\left(\frac{\epsilon_t}{\sqrt{2}}\right)\right]\right),$$

kde $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \mathrm{e}^{-x^2}\,\mathrm{d}x$ pre $z \in \mathbb{R}$ predstavuje tzv. chybovú funkciu (angl. *error function*). Pre parameter šikmosti uvažujeme reálne spektrum vzhľadom k tomu, že hodnoty $\nu > 0$ predstavujú pravostranné zošikmenie a $\nu < 0$ ľavostranné [10]. Pre $\nu = 0$ dostávame štandardizované normálne rozdelenie. Model B sa pripodobňuje Taylorovmu modelu A s ľahkými chvostami chybových členov ϵ_t, avšak jeho zvláštnosťou je práve jeho pridaný parameter tvaru, ktorý dokáže zachytiť zošikmenie výnosov.

2.5 Model C – Studentovo *t*-rozdelenie

V druhom modeli modelujeme chybový člen ϵ_t pomocou Studentovho t-rozdelenia, ktoré nám dáva dostatok voľnosti na zachytenie „extrémnejších“ hodnôt výnosov pomocou jeho ťažkých chvostov. Na zachytenie leptokurtózy sa preto nepodieľa len samotný proces volatility, ale aj chybový člen ϵ_t. Model C definujeme v tvare

$$\begin{aligned} y_t &= \exp\left(h_t/2\right)\epsilon_t, \\ h_{t+1} &= \mu + \phi\left(h_t - \mu\right) + \eta_t, \\ \epsilon_t &\sim t_\nu(0;\, 1), \\ \eta_t &\sim \mathcal{N}(0;\, \sigma_\eta^2), \\ \rho\left(\epsilon_t,\, \eta_t\right) &= 0,\, \forall t \in T, \end{aligned}$$

kde chybový člen ϵ_t modelujeme so Studentovým t-rozdelením so strednou hodnotou rovnou 0, normalizovanou varianciou rovnou 1 a parameterom ν predstavujúcim počet stupňov voľností. V takomto prípade má hustota Studentovho t-rozdelenia pravdepodobnosti predpis s normalizačným členom $(\nu - 2)$ v tvare:

$$p(\epsilon_t) = \frac{1}{\sqrt{\pi\,(\nu-2)}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{\epsilon_t^2}{\nu - 2}\right)^{-\frac{\nu+1}{2}}, \tag{6}$$

kde $\Gamma(\cdot)$ označuje tzv. gama funkciu definovanú vzťahom $\Gamma(z) = \int_0^\infty x^{z-1}\mathrm{e}^{-x}\,\mathrm{d}x$ pre $z > 0$. Z funkcie hustoty vo vzťahu (6) vyplýva spodné ohraničenie pre počet stupňov voľností $\nu > 2$. Pri platnosti podmienky $\nu > 4$ dostávame podľa [40] koeficient špicatosti

$$k = \frac{3\,(\nu-2)}{\nu-4},$$

ktorý je väčší než 3 v prípade, ak $\nu < \infty$.

Ak parameter ν je rovný 1, tak Studentovo t-rozdelenie je ekvivalentné so štandardným Cauchyho rozdelením (ktoré má parameter polohy rovný 0 a parameter škály rovný 1). V takomto prípade je možné zaradiť Studentovo t-rozdelenie do užšej, špecifickejšej podtriedy rozdelení s ťažkými chvostami, a to do podtriedy rozdelení s hrubými chvostami, kedy má „extrémne ťažké, dlhé a hrubé chvosty“ [1]. Dôležitou vlastnosťou chvostov rozdelení patriacich do skupiny rozdelení s hrubými chvostami (*fat tailed distributions*) je vlastnosť asymptotickej invariantnosti škály, ktorá znamená, že chvosty sa správajú konzistentne pri škálovaní celej distribúcie [8]. Konkrétne, pri škálovaní celej distribúcie (zväčšení alebo zmenšení jej variancie) relatívny podiel chvostov v porovnaní s hlavnou časťou distribúcie zostáva nemenný. Ďalšou charakteristikou tejto skupiny je prítomnosť nekonečných momentov. Napríklad, v prípade parametra tvaru $\nu \in (0, 2)$ má Studentovo t-rozdelenie nekonečný druhý moment.

Ak zvyšujeme hodnotu parametra ν, tak dané Studentovo t-rozdelenie sa stále viac a viac „vzďaľuje“ od triedy hrubých chvostov, pričom pri každej konečnej hodnote parametra ν má ešte stále ťažké chvosty. Pre počet stupňov voľnosti $\nu \to \infty$ dostávame normálne rozdelenie so špicatosťou rovnou,

$$\lim_{\nu\to\infty} \frac{3\,(\nu-2)}{\nu-4} = 3.$$

kedy je Studentovo t-rozdelenie zhodné so štandardizovaným Gaussovým normálnym rozdelením, ktoré má ľahké chvosty.

2.6 Model D – GED rozdelenie

Modelom D značíme taký model SV, kde chybový člen ϵ_t má GED rozdelenie so strednou hodnotou 0, jednotkovou varianciou a parametrom tvaru $\nu > 0$ v znení

$$\begin{aligned} y_t &= \exp\left(h_t/2\right)\epsilon_t, \\ h_{t+1} &= \mu + \phi\left(h_t - \mu\right) + \eta_t, \\ \epsilon_t &\sim \mathrm{GED}_\nu(0;\, 1), \\ \eta_t &\sim \mathcal{N}(0;\, \sigma_\eta^2), \\ \rho\left(\epsilon_t,\, \eta_t\right) &= 0,\, \forall t \in T, \end{aligned}$$

s funkciou hustoty

$$p(\epsilon_t) = \frac{\nu}{\lambda 2^{\left(1+\frac{1}{\nu}\right)}\Gamma\left(\frac{1}{\nu}\right)} \exp\left(-\frac{1}{2}\left|\frac{\epsilon_t}{\lambda}\right|^\nu\right),$$

kde

$$\lambda = \left(2^{-\frac{2}{\nu}}\frac{\Gamma\left(\frac{1}{\nu}\right)}{\Gamma\left(\frac{3}{\nu}\right)}\right)^{\frac{1}{2}}.$$

Koeficient špicatosti člena chýb ϵ_t, ktorý sa riadi podľa GED rozdelenia, je definovaný [10], [40]:

$$\mathrm{E}\left[\epsilon_t^4\right] = \frac{\Gamma\left(\frac{1}{\nu}\right)\Gamma\left(\frac{5}{\nu}\right)}{\left[\Gamma\left(\frac{3}{\nu}\right)\right]^2},$$

pričom z nášho uhla pohľadu sú najzaujímavejšie také prípady, kedy parameter tvaru ν nadobúda hodnoty z intervalu (0, 2), čím sa zaraďuje do podtriedy rozdelení pravdepodobnosti s ťažkými chvostami, a to do triedy rozdelení s hrubými chvostami [21]. V takomto prípade je koeficient špicatosti GED rozdelenia je väčší ako 3, čo spolu s exponenciálnym členom uvedenom vo vzťahu (5) dáva leptokurtózu log-výnosov y_t. Pre $\nu = 2$ dostávame normálne rozdelenie [10].

2.7 Ekvivalentný zápis modelov

V tejto časti si zadefinujeme ekvivalentný zápis modelov SV, ktorý, ako sa ukáže neskôr v kapitole 8.1, bude užitočný v praktickej časti tejto práce, pri zápise jednotlivých modelov do špecifického programovacieho jazyka Stan [27] implementovaného v štatistickom softvéri R [24] pomocou balíčka `rstan`.

Na to, aby sme vedeli jednotlivé modely zapísať do jazyka Stan, potrebujeme odvodiť podmienené rozdelenia pravdepodobnosti daných veličín. Bez ujmy na všeobecnosti, uvažujme model A definovaný v podkapitole 2.3. Podmienený proces volatility

$$h_t \,|\, h_{t-1}, \mu, \phi, \sigma_\eta^2,$$

daný rovnicou:

$$h_t = \mu + \phi \left(h_{t-1} - \mu\right) + \eta_t,$$

sa riadi podľa normálneho rozdelenia so strednou hodnotou

$$\begin{aligned} \mathrm{E}\left[h_t \,|\, h_{t-1}, \mu, \phi, \sigma_\eta^2\right] &= \mathrm{E}\left[\mu + \phi \left(h_{t-1} - \mu\right) + \eta_t\right] \\ &= \mu + \phi \left(h_{t-1} - \mu\right) + \mathrm{E}\left[\eta_t\right] \\ &= \mu + \phi \left(h_{t-1} - \mu\right) \end{aligned}$$

a varianciou

$$\begin{aligned} \mathrm{D}\left(h_t \,|\, h_{t-1}, \mu, \phi, \sigma_\eta^2\right) &= \mathrm{D}\left(\mu + \phi \left(h_{t-1} - \mu\right) + \eta_t\right) \\ &= \mathrm{D}\left(\eta_t\right) \\ &= \sigma_\eta^2, \end{aligned}$$

pričom normalita vyplýva z predpokladu $\eta_t \sim \mathcal{N}\left(0;\, \sigma_\eta^2\right)$.

Pre člen chýb ϵ_t, ktorý sa riadi podľa štandardizovaného normálneho rozdelenia $\epsilon_t \sim \mathcal{N}\left(0;\, 1\right)$ uvažujeme podmienené rozdelenie log-výnosov

$$y_t \,|\, h_t,$$

definovaného rovnicou:

$$y_t = \exp\left(h_t/2\right) \epsilon_t,$$

ktoré sa riadi podľa normálneho rozdelenia so strednou hodnotou

$$\mathrm{E}\left[y_t \,|\, h_t\right] = \mathrm{E}\left[\exp\left(h_t/2\right) \epsilon_t\right] = \exp\left(h_t/2\right) \mathrm{E}\left[\epsilon_t\right] = 0$$

a varianciou

$$\mathrm{D}\left(y_t \,|\, h_t\right) = \mathrm{D}\left(\exp\left(h_t/2\right) \epsilon_t\right) = \exp\left(h_t\right) \mathrm{D}\left(\epsilon_t\right) = \exp\left(h_t\right).$$

Model A vieme teda zapísať ako [19]:

$$y_t \,|\, h_t \sim \mathcal{N}\left(0;\, \exp\left(h_t\right)\right), \tag{7}$$

$$h_t \,|\, h_{t-1}, \mu, \phi, \sigma_\eta^2 \sim \mathcal{N}\left(\mu + \phi\left(h_{t-1} - \mu\right);\, \sigma_\eta^2\right), \tag{8}$$

spolu s počiatočným podmieneným rozdelením $h_1 \,|\, \mu, \phi, \sigma_\eta^2 \sim \mathcal{N}\left(\mu;\, \frac{\sigma_\eta^2}{1-\phi^2}\right)$ definovaným vo vzťahu (2). Pre modely B, C a D sa z dvojice vzťahov (7) a (8) zmení iba podmienené rozdelenie $y_t \,|\, h_t$, vzhľadom na rozdelenie pravdepodobnosti člena chýb ϵ_t v závislosti od modelu.

3 Odhadovanie parametrov modelu stochastickej volatility

V klasických štatistických modeloch odhadujeme parametre modelu na základe štatistických metód, medzi ktoré patrí aj metóda maximálnej vierohodnosti (skr. MLE z angl. *maximum likelihood estimation*), ktorej silný predpoklad spočíva práve v tom, že poznáme analytickú funkciu hustoty rozdelenia pravdepodobnosti. Odhady neznámych parametrov skúmaného rozdelenia pravdepodobnosti potom zvyčajne získame riešením optimalizačnej úlohy tak, aby hodnoty týchto parametrov maximalizovali funkciu vierohodnosti. V modeloch stochastickej volatility však narážame na problém, ktorý spočíva v tom, že nevieme explicitne vyjadriť funkciu vierohodnosti, a to najmä kvôli podkladovému stochastickému procesu, a takisto aj kvôli zložitosti modelu. Navyše, v praxi požadujeme okrem odhadu hodnôt parametrov aj niečo navyše, a to odhad latentnej premennej, teda procesu log-volatility, na ktoré nám klasické štatistické metódy na odhad parametrov nestačia. V praxi sa stretávame s dvoma prístupmi, ktorými vieme požadované ciele dosiahnuť [6]:

i) *štatistické charakteristiky,* akými sú momenty rozdelenia alebo funkcia vierohodností, patria medzi klasické metódy založené na pozorovaní teoretických a empirických momentov výnosov, čím však získavame iba odhady parametrov bez odhadu latentného procesu,

ii) *simulačné metódy,* ktoré sa líšia od tradičného prístupu k štatistike; v predkladanej práci využijeme na modelovanie stochastického procesu volatility bayesovský prístup k štatistike, ktorý na odhad parametrov, a rovnako aj realizáciu latentnej premennej využíva Monte Carlo simulácie, ktoré si bližšie popíšeme v kapitole 4. V nasledujúcej časti si priblížime princíp bayesovského prístupu ku štatistike.

3.1 Bayesovský prístup k štatistike

Jedným zo základných krokov pri odhadovaní parametrov rozdelení pravdepodobnosti, a teda parametrov modelu, je výber prístupu k štatistike, ktorý môže byť *frekvenčný* nazývaný aj *klasický*) alebo bayesovský. V našom prípade sa po vzore [10],

[19], [20] či [40] zaoberáme bayesovským prístupom k štatistike, ktorý sa od frekventistického líši, okrem iného, základným predpokladom o parametroch, a to takým, že v bayesovskom prístupe nepokladáme neznámy parameter θ za konštantu, ale za náhodnú premennú.

Uvažujme náhodnú premennú Y s vektorom parametrov $\boldsymbol{\theta}$ z parametrického priestoru Θ a $\boldsymbol{h}$ ako neznámy vektor procesu volatility. Nech $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$ predstavuje vektor pozorovaných dát. Naše predpoklady a očakávania o každom parametre z vektora parametrov $\boldsymbol{\theta}$ v bayesovskej teórii zhrnieme do **apriórnych rozdelení pravdepodobnosti** $f(\boldsymbol{\theta}, \boldsymbol{h})$ vektora náhodných premenných $\boldsymbol{\theta}$, teda každý z parameterov (napr. v prípade Taylorovho modelu: μ, ϕ, σ_η) uvažujeme v bayesovskej teórii ako **náhodnú premennú** z nejakého rozdelenia pravdepodobnosti, pričom parametre jednotlivých rozdelení pravdepodobnosti nazývame hyperparametre. Cieľom je získať také **aposteriórne rozdelenie pravdepodobností** $f(\boldsymbol{\theta}, \boldsymbol{h} \mid \boldsymbol{y})$, ktoré zahŕňa apriórne očakávania o rozdelení pravdepodobnosti spolu s výsledkami pozorovaní pomocou dáta generujúceho rozdelenia nazývaného tiež funkcia vierohodnosti $f(\boldsymbol{y} \mid \boldsymbol{\theta}, \boldsymbol{h})$, čo vieme matematicky zhrnúť [35]:

$$f(\boldsymbol{\theta}, \boldsymbol{h} \mid \boldsymbol{y}) = \frac{f(\boldsymbol{y} \mid \boldsymbol{\theta}, \boldsymbol{h})\, f(\boldsymbol{\theta}, \boldsymbol{h})}{\displaystyle\int_\Theta f(\boldsymbol{y} \mid \boldsymbol{\theta}, \boldsymbol{h})\, f(\boldsymbol{\theta}, \boldsymbol{h})\, \mathrm{d}(\boldsymbol{\theta}, \boldsymbol{h})}\,. \tag{9}$$

Bayesovský prístup k štatistike v praxi znamená, že

i) pre parametre (vektory náhodných premenných) $\boldsymbol{\theta}$ a $\boldsymbol{h}$ vyslovíme naše apriórne očakávania o rozdeleniach pravdepodobnosti $f(\boldsymbol{\theta}, \boldsymbol{h})$ bez toho, aby sme skúmali ich realizácie,

ii) pre dáta $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$, ktoré sú realizáciou náhodnej premennej Y vhodne zvolíme rozdelenie pravdepodobností $f(\boldsymbol{y} \mid \boldsymbol{\theta}, \boldsymbol{h})$, ktoré zahŕňa pozorované dáta $\boldsymbol{y}$ a naše „predstavy“ o parametroch modelu,

a spojením v (9) získame aposteriórnu hustotu s prepočítanými hodnotami hyperparametrov.

3.2 Apriórne rozdelenia pravdepodobností v modeloch stochastickej volatility

Na to, aby sme skompletizovali definície modelov v bayesovskom prístupe ku štatistike, potrebujeme ešte zadefinovať apriórne rozdelenia pravdepodobnosti parametrov jednotlivých modelov stochastickej volatility. Za priekopníka sa považuje [20] po vzore ktorého mnohí ďalší autori, medzi ktorých patria aj [10], [19] či [40], pracujú s nasledujúcimi rozdeleniami pravdepodobnosti.

Pre parameter úrovne μ je celkom prirodzené uvažovať normálne rozdelenie $\mu \sim \mathcal{N}(k_1;\, k_2)$, kde $k_1 \in \mathbb{R}$ a $k_2 > 0$, pričom sme za hyperparametre po vzore [39] položili $k_1 = -10$ a $k_2 = 25$. Podľa empirických výsledkov [19] však voľba konkrétnych hodnôt hyperparametrov rozdelenia μ nie je veľmi ovplyvňujúca.

Čo sa týka parametra perzistencie ϕ „stojaceho“ pri autoregresnom procese, tak od neho vyžadujeme podmienku na ohraničenie $|\phi| < 1$, ktorá zabezpečí stacionaritu procesu volatility. To dosiahneme na základe transformácie štandardného beta rozdelenia $\phi^* \sim \mathrm{B}(l_1; l_2)$, kde $l_1 > 0$ a $l_2 > 0$, pomocou návrhu [20]

$$\phi = 2\phi^* - 1, \tag{10}$$

pričom po transformácii dostávame ohraničenia na parametre $l_1 > \frac{1}{2}$, $l_2 > \frac{1}{2}$. Je zrejmé, že transformáciou (10) premennej definovanej na intervale $\phi^* \in (0,\, 1)$ dostaneme požadované ohraničenie $\phi \in (-1,\, 1)$ zabezpečujúce stacionaritu procesu.

Z vlastnosti prvých momentov beta rozdelenia je známa jej stredná hodnota a variancia daná kombináciou parametrov l_1, l_2:

$$\mathrm{E}[\phi^*] = \frac{l_1}{l_1 + l_2}, \qquad \mathrm{D}(\phi^*) = \frac{l_1 l_2}{(l_1 + l_2)^2 (l_1 + l_2 + 1)},$$

z čoho vieme ľahko na základe vlastností strednej hodnoty a variancie odvodiť prvé dva momenty pre transformovanú premennú ϕ:

$$\mathrm{E}[\phi] = \frac{2l_1}{l_1 + l_2} - 1, \qquad \mathrm{D}(\phi) = \frac{4l_1 l_2}{(l_1 + l_2)^2 (l_1 + l_2 + 1)}. \tag{11}$$

Zo strednej hodnoty a variancie vyplýva, že stredná hodnota závisí od pomeru hyperparametrov l_1, l_2, pričom stredná hodnota je kladná, ak $l_1 > l_2$, a naopak, záporná, ak $l_1 < l_2$. Zo vzťahu variancie vyplýva, že čím volíme väčšie hodnoty hyperparametrov

l_1, l_2, tak tým variancia klesá. Na základe odporúčania publikovaného v [20] volíme hyperparametre $l_1 = 20$ a $l_2 = 1{,}5$, čím získavame strednú hodnotu rovnú 0,86 a varianciu na úrovni 0,11, a tým pádom nízku pravdepodobnosť záporných hodnôt pre parameter perzistencie.

Pre volatilitu σ_η^2 procesu volatility je najznámejšou voľbou [20], [39], [40] apriórneho rozdelenia inverzné gama rozdelenie $\sigma_\eta^2 \sim \text{IG}\left(\frac{\sigma_r}{2}; \frac{S_\sigma}{2}\right)$ s parametrami tvaru a škály v uvedenom poradí, pričom $\sigma_r > 0$ a $S_\sigma > 0$.. V našom prípade sme najlepšie výsledky modelovania dosiahli pri voľbe apriórneho rozdelenia gama $\sigma_\eta^2 \sim \text{G}\left(\frac{1}{2}; \frac{1}{2B_{\sigma_\eta}}\right)$ po vzore [14], [19], ktorí argumentujú, že použitie tohto rozdelenia umožňuje efektívne modelovanie aj pre prípady, kedy sú hodnoty σ_η^2 veľmi blízke nule. Takéto modelovanie umožňuje odstrániť nežiaduci šum pri odhade latentnej premennej premennej. Hodnotu B_{σ_η} sme položili rovnú 20.

Posledným parametrom ν, ktorý sa vyskytuje v modeli B s chybovým členom riadiacim sa zošikmeným normálnym rozdelením, kde ν predstavuje parameter tvaru, rovnako v modeli C ako počet stupňov voľnosti chybového člena ϵ_t, ktoré sa riadi Studentovým t-rozdelením, a taktiež v modeli D, kde pre $\epsilon_t \sim \text{GED}_\nu$ predstavuje parameter tvaru. Vo všetkých troch prípadoch teda závisí tento parameter od rozdelenia pravdepodobností chybového člena ϵ_t, pričom predpokladáme nezávislosť od ϕ a σ_η^2.

Pre model B so zošikmeným normálnym rozdelením uvažujeme podľa [10] normálne rozdelenie $\nu \sim \mathcal{N}(0; 4)$ so strednou hodnotou 0 a varianciou 4. V modeli C so Studentovým t-rozdelením potrebujeme pre potreby konečného druhého momentu pre $\boldsymbol{y}$ zabezpečiť ohraničenie $\nu > 2$. Podľa návrhu recenzenta článku [18] volíme exponenciálne rozdelenie v tvare $\nu - 2 \sim \mathcal{E}(\lambda)$. Hodnotou hyperparametra položíme $\lambda = \frac{1}{3}$. Nakoniec, pre model D s GED rozdelením uvažujeme pre parameter tvaru škálované inverzné χ^2-rozdelenie v tvare $\nu \sim \text{Inv-}\chi^2(10; 0{,}05)$ [10].

Napriek tomu, že je bayesovský prístup v praxi široko uplatniteľný a využívaný, tak má aj svoju často kritizovanú stránku, ktorou je voľba apriórneho rozdelenia, ktorá závisí od empirických skúsenosti modelára, ktoré môže byť chybné. Dôležitou súčasťou je preto bohatosť dátovej zložky, teda informácie obsiahnutej v dátach, ktorá pri dostatočne veľkej dátovej vzorke preváži možné chybné predpoklady modelára [6].

3.3 Prehľad existujúcej literatúry

Modelovanie finančných časových radov je dôležitou súčasťou pri obchodovaní s komoditami, cennými papiermi či inými aktívami. Neustály pohyb na trhu vplýva do hlbokých sfér, ktoré finančníci potrebujú monitorovať a rovnako aj predvídať do budúcna, preto sa modelovanie vo finančníctve stalo nevyhnutnosťou. Medzi takéto sféry patrí aj volatilita, ktorá, ako sa ukázalo, nie je konštantná, ale naopak, má svoj vývoj v čase. Dôležitosť modelovania časovo závislej volatility vyplýva z potreby poznania rizika, ktoré investícia v sebe zahŕňa, a ktoré je s volatilitou úzko späté.

Historicky sa koncepty modelovania volatility vyvíjali od roku 1982, kedy americký ekonóm a štatistik Robert F. Engle [12] predstavil základy modelovania časovo závislej volatility, známej ako pod triedou ARCH, ktorú v roku 1986 rozšíril a zovšeobecnil dánsky ekonóm Tim Bollerslev [7]. Druhú vetvu modelovania časovo závislej volatility predstavil Stephen J. Taylor v roku 1982, ktorý vo svojej práci [37] popisoval volatilitu pomocou triedy modelov stavového priestoru a proces volatility modeloval ako logaritmus štvorcovej volatility $h_t = \ln \sigma_t^2$, ktorý uvádzame v časti 2.1. Hoci mnohí výskumníci, vrátane S. Kima, N. Shepharda a S. Chiba (pozri [20]), uprednostňujú Taylorov prístup k modelovaniu, obmedzená výpočtová kapacita vtedajších počítačov priala skôr (G)ARCH modelom. Log-volatilita je v prípade modelov stochastickej volatility uvažovaná ako Markovov proces, čím sa log-volatilita stáva stochastickým procesom a poskytuje hodnotnú alternatívu modelom, ktoré predpokladajú konštantnosť a časovú nezávislosť tejto veličiny.

Existuje niekoľko metód na odhadovanie modelov, avšak najväčší prínos predstavujú modely založené na simulačných technikách. Medzi tieto techniky patrí aj metóda Monte Carlo Markov Chain, ktorá je považovaná za najefektívnejšiu simulačnú metódu na odhadovanie parametrov modelu a realizáciu latentných premenných. Prvá myšlienka MCMC simulácií, teda Metropolisovho-Hastingsovho (skr. MH) algoritmu, sa objavila vo vedeckej štúdii Nicholasa Metropolisa a kolektívu autorov v roku 1953 [22], ktorý ho vytvoril za účelom simulovania energetických úrovní atómov v kryštalickej štruktúre. Neskôr, v roku 1970 bola Metropolisova metóda zovšeobecnená Williamom Hastingsom [16], ktorý ju rozšíril ako nástroj na použitie v štatistických úlohách s cieľom simulovať vzorky z aposteriórneho rozdelenia pravdepodobnosti. Alternatívu

Metropolisovmu-Hastingsovmu algoritmu poskytuje Gibbsov algoritmus predstavený v roku 1984 bratmi Stuartom a Donaldom Gemanom [15]. Viac informácií k modelom založeným na simulačných technikách je možné nájsť v [6].

Medzi najnovšie objavy v MCMC metódach sa zaraďuje Hamiltonova Monte Carlo metóda, dostupná napríklad v publikácií Radforda M. Neala [23], či Michaela Betancourta [2], ktorá poskytuje efektívne a cielené riešenia pri odhadovaní parametrov a latentnej premennej v modeloch stochastickej volatility.

V súčasnosti existuje adaptívna nadstavba Hamiltonovskej Monte Carlo metódy, nazývaná No-U-Turn Sampler, ktorá bola predstavená v roku 2014 autormi M. D. Hoffmanom a A. Galemanom [17].

V súčasnosti existuje niekoľko softvérov, ktoré predstavujú účinný nástroj pre stochastické modelovanie pomocou MCMC metódy, medzi ktoré patrí knižnica Stan [27] zabudovaná, okrem iného, aj v softvéri R [24] s balíčkom `rstan`, ktorá poskytuje praktické modelovanie pomocou bayesovského prístupu k štatistike pomocou Hamiltonovej Monte Carlo metódy či dokonca jej No-U-Turn Sampler nadstavby, s ktorým pracujeme v našej diplomovej práci.

4 Monte Carlo techniky v modeloch stochastickej volatility

V praxi sa stretávame s problémom analytického vyjadrenia aposteriórnej funkcie hustoty, ktorá súvisí napríklad so zložitejšími tvarmi modelov či komplikovanejšími rozdeleniami pravdepodobnosti. V takýchto situáciach sa matematickí modelári spoliehajú na simulačné techniky Monte Carlo, ktoré na základe generovania náhodných hodnôt podľa zadaných distribúcií a frekvencie týchto hodnôt charakterizujú základné štatistické ukazovatele tejto náhodnej vzorky, čím ďalej vedia odhadovať vlastnosti parametrov a iné štatistické ukazovatele [34]. Neznáme parametre sa teda generujú veľkým množstvom simulácií, ktoré predstavujú možné realizácie systému resp. vzorky z pravdepodobnostného rozdelenia odhadovaných parametrov, z ktorého počítame základné štatistické ukazovatele, ktoré nám čo to napovedia o pravdepodobnostnom rozdelení dát. Z prednášok zo štatistiky vieme, že medzi základné ukazovatele patria

i) **stredná hodnota (očakávaná hodnota)**, ktorá hovorí o centralite rozdelenia pravdepodobnosti, a teda o tom, k akej hodnote sú realizácie „priťahované“,

ii) **štandardná odchýlka (resp. rozptyl)** predstavuje mieru rozptýlenia dát

a ďalšie, ktoré budeme používať v kapitole 6 tejto diplomovej práce.

4.1 Bayesovský prístup v modeloch stochastickej volatility

Ako sme už spomínali v časti 3, dôležitou súčasťou simulačných techník pri odhadovaní parametrov modelu stochastickej volatility je práve získanie odhadov latentnej premennej. Využitím bayesovského prístupu a simulačných techník získavame silný nástroj ako na odhad parametrov $\boldsymbol{\theta}$, tak aj na odhad realizácie procesu volatility $\boldsymbol{h} = \{h_t\}_{t=1}^{T}$ z log-výnosov $\boldsymbol{y} = \{y_t\}_{t=1}^{T}$ pre skúmanú jednotku času t z množiny T, kde $T = \{1, 2, \ldots, n\}$, pričom $n \in \mathbb{N}$ označuje počet sledovaných log-výnosov. V nasledujúcej časti si predstavíme Monte Carlo techniky podľa [6].

Odhad aposteriórnej hustoty získame použitím bayesovského štatistického prístupu zo vzťahu (9). Integrál v menovateli je zvyčajne náročný na výpočet, pretože

sa pri viacparametrových modeloch jedná o viacrozmerný integrál. Navyše, pre fixné $\boldsymbol{y}$ je integrál nezávislý od integračných premenných, čím má v konečnom dôsledku len škálovací efekt na výslednú aposteriórnu hustotu. Z tohto dôvodu sa v praxi stretávame so zjednodušeným, proporcionálnym tvarom, ktorý budeme uvažovať aj v tejto diplomovej práci:

$$f\left(\boldsymbol{\theta}, \boldsymbol{h} \,|\, \boldsymbol{y}\right) \propto f\left(\boldsymbol{y} \,|\, \boldsymbol{\theta}, \boldsymbol{h}\right) f\left(\boldsymbol{\theta}, \boldsymbol{h}\right) ,$$

ktorý má pomerne jasnú interpretáciu. Aposteriórnu hustotu získame súčinom našich predpokladov o hyperparametroch (a stavovej premennej) v apriórnej hustote $f\left(\boldsymbol{\theta} \,|\, \boldsymbol{h}\right)$, spolu s funkciou vierohodnosti, ktorú predstavuje člen $f\left(\boldsymbol{y} \,|\, \boldsymbol{\theta}, \boldsymbol{h}\right)$ po tom, ako naše odhady spojíme s modelovanou štruktúrou dát. Aposteriórnu hustotu vieme ďalej pomocou Bayesovho vzorca proporcionálne rozložiť do tvaru

$$f\left(\boldsymbol{\theta}, \boldsymbol{h} \,|\, \boldsymbol{y}\right) \propto f\left(\boldsymbol{y} \,|\, \boldsymbol{\theta}, \boldsymbol{h}\right) f\left(\boldsymbol{h} \,|\, \boldsymbol{\theta}\right) f\left(\boldsymbol{\theta}\right) ,$$

kde $f\left(\boldsymbol{y} \,|\, \boldsymbol{\theta}, \boldsymbol{h}\right)$ je funkcia vierohodnosti $\boldsymbol{y}$ pri danom $\boldsymbol{\theta}$ a $\boldsymbol{h}$, ďalej funkcia $f\left(\boldsymbol{h} \,|\, \boldsymbol{\theta}\right)$ predstavuje hustotu rozdelenia $\boldsymbol{h}$ pri danom $\boldsymbol{\theta}$ a nakoniec $f\left(\boldsymbol{\theta}\right)$ označujeme ako apriórnu hustotu $\boldsymbol{\theta}$.

4.2 Monte Carlo integrovanie

Medzi užitočné aplikácie Monte Carlo metódy sa zaraďuje MC integrovanie, ktoré je na rozdiel od numerického integrovania založené na integrovaní v náhodných bodoch x_i pre $i \in 1, \ldots, K$ z požadovaného (cieľového) rozdelenia pravdepodobnosti $f\left(x\right)$. Integrand $q\left(x\right)$ je rozložený na súčin cieľovej funkcie $f\left(x\right)$ a vhodne zvolenej pomocnej funkcie $g\left(x\right)$. Cieľová hustota je aproximovaná histogramom, ktorý vznikol simuláciami z vybraného rozdelenia a hodnota integrálu je aproximovaná výberovým priemerom funkcie $g\left(x\right)$ vyčíslených vo funkčných hodnotách simulácií [6]:

$$I = \int_{\Omega} q\left(x\right) \,\mathrm{d}x = \int_{\Omega} g\left(x\right) f\left(x\right) \,\mathrm{d}x = \mathrm{E}\left[q\left(x\right)\right] \approx \frac{1}{N} \sum_{n=1}^{N} q\left(x_n\right) = \hat{I}_N, \tag{12}$$

kde I je hodnota hľadaného integrálu, $x_1, \ldots, x_n$ predstavuje náhodný výber z rozdelenia s hustotou $f\left(x\right)$ na oblasti Ω. Navyše pre náhodnú premennú $q\left(x\right)$ so strednou hodnotou $\mathrm{E}\left[q\left(x\right)\right]$ podľa zákona veľkých čísel konverguje výberový priemer $\hat{I}_N$ skoro všade ku svojej strednej hodnote $\mathrm{E}\left[q\left(x\right)\right] = I$ pre $N \to \infty$.

4.2.1 Monte Carlo integrovanie v modeloch stochastickej volatility

Alfou a omegou pri výpočtoch v štatistike je vždy nejaký druh integrálu, či sa jedná o počítanie pravdepodobnosti z funkcie hustoty alebo výpočet momentov nejakého rozdelenia pravdepodobnosti. V modeloch SV sa stretávame s potrebou integrovania pri odhadovaní parametrov z aposteriórneho rozdelenia pravdepodobnosti medzi ktoré patrí aj výpočet strednej hodnoty [6]

$$\hat{g}\left(\boldsymbol{\theta}, \boldsymbol{h}\right) = \mathrm{E}_f\left[g\left(\boldsymbol{\theta}, \boldsymbol{h}\right)|\boldsymbol{y}\right] = \int g\left(\boldsymbol{\theta}, \boldsymbol{h}\right) f\left(\boldsymbol{\theta}, \boldsymbol{h}|\boldsymbol{y}\right) \mathrm{d}\boldsymbol{h}\mathrm{d}\boldsymbol{\theta} \approx \frac{1}{N}\sum_{n=1}^{N} g\left(\theta^{(n)}, h^{(n)}\right), \quad (13)$$

pričom $\left(\theta^{(n)}, h^{(n)}\right)$ pre $n \in \{1, \ldots, N\}$ predstavuje náhodný výber z integrovateľnej funkcie združenej aposteriórnej hustoty $f\left(\boldsymbol{\theta}, \boldsymbol{h}|\boldsymbol{y}\right)$.

Vhodným zvolením funkcie $g\left(\boldsymbol{\theta}, \boldsymbol{h}\right)$ vo vzťahu (13) získavame pre

i) voľbu $g\left(\boldsymbol{\theta}, \boldsymbol{h}\right) = \theta_i$ bodový odhad parametra θ_i z k-prvkového vektora parametrov $\boldsymbol{\theta}$

$$\hat{\theta}_i = \mathrm{E}_f\left[\theta_i|\boldsymbol{y}\right] \approx \frac{1}{N}\sum_{n=1}^{N} \theta_i^{(n)}, \qquad i = 1, \ldots, k,$$

ii) voľbu $g\left(\boldsymbol{\theta}, \boldsymbol{h}\right) = h_t$ odhad realizácií stavovej premennej h_t v čase t

$$\hat{h}_t = \mathrm{E}_f\left[h_t|\boldsymbol{y}\right] \approx \frac{1}{N}\sum_{n=1}^{N} h_t^{(n)}, \qquad t = 1, \ldots, T.$$

Kľúčom pre získanie bodových odhadov je získanie náhodného výberu z aposteriórnej hustoty. Modely stochastickej volatility však patria do triedy modelov, z ktorých (pre vyššie spomínané dôvody) nie je možné generovanie náhodných výberov priamo. Z tohto dôvodu vzniklo mnoho bayesovských simulačných techník, ktoré aproximujú aposteriórnu hustotu a generujú z nej náhodný výber [6]. Hlavná myšlienka spočíva v použití novej, pomocnej hustoty, z ktorej sa dá generovať náhodný výber, pričom tento výber je ďalej upravovaný v závislosti od typu metódy. Pre zložitejší tvar aposteriórnej hustoty, či vysokú dimenziu, klasické simulačné metódy ako sú napríklad *Importance Sampling*, *Acceptance-Rejection Sampling* či *Adaptive Rejection Sampling*, nestačia, preto sa im bližšie nebudeme venovať, pričom bližšie informácie o týchto metódach je možné nájsť napríklad v publikácií [6].

5 Metóda Markov Chain Monte Carlo

Metóda Markov Chain Monte Carlo (ďalej označované ako MCMC) patrí medzi simulačné metódy, ktorá poskytuje vzorky z cieľového rozdelenia, z ktorého sa nedá generovať priamo. MCMC negeneruje náhodný výber tak, ako je to u bayesovských simulačných techník ale realizácie sú získane z Markovovho reťazca [6], [39]. Nasledujúca kapitola prestavuje náhľad do teórie MCMC, pričom pre jej úplné pochopenie si najskôr predstavíme základy teórie Markovových reťazcov.

5.1 Markovove reťazce

V nasledujúcej časti si voľne predstavíme základnú teóriu Markovových reťazcov a ich vlastnosti, na ktorých je vybudovaná metóda MCMC. Vzhľadom na obmedzený rozsah tejto diplomovej práce neuvádzame rozsiahlu teóriu Markovových reťazcov. Obmedzíme sa len na základné definície a tvrdenia Markovových reťazcov pre diskrétny čas, ktoré voľne čitateľovi predstavíme. Dôkazy či ďalšie rozšírenia teórie Markovových reťazcov je možné nájsť napríklad v publikáciách [6], [25], [38].

5.1.1 Markovove reťazce s konečnou množinou stavov

Markovov reťazec definujeme ako postupnosť náhodných prvkov $X_1, \ldots, X_T$, kde $T \in \mathbb{N}_0$ nadobúdajúcich hodnoty z konečnej množiny prvkov $\boldsymbol{\chi}$, pre ktoré má Markovov reťazec tzv. „zabúdaciu vlastnosť“, kedy podmienené rozdelenie nadobudnutia stavu X_{n+1} pri daných stavoch $X_1, \ldots, X_n$ závisí výhradne len na stave X_n. Pri Markovových reťazcoch sa môžeme stretnúť s množinou stavov, ktorá je konečná alebo spočítateľná. V takom prípade vieme Markovov reťazec matematicky vyjadriť ako [6], [25]:

$$P\left(X_{n+1} = x_{n+1} \,|\, X_n = x_n,\, X_{n-1} = x_{n-1},\, \ldots, X_1 = x_1\right) = P\left(X_{n+1} = x_{n+1} \,|\, X_n = x_n\right),$$

pre všetky časy $\forall n \in T$ a pre všetky stavy z množiny stavov $x_i \in \boldsymbol{\chi}$. Zabúdacia vlastnosť sa nazýva Markovova vlastnosť a predstavuje situáciu, kedy budúci stav závisí len od aktuálneho stavu a nie od predchádzajúcich stavov, za predpokladu, že predchádzajúce stavy sú známe.

Markovov reťazec je definovaný v prvom stave X_1 pomocou začiatočného rozdelenia pravdepodobnosti $\boldsymbol{p}(1)$ (angl. *initial distribution*) a pravdepodobnosťami prechodu medzi stavmi (angl. *transition probabilities*) $P(X_{n+1} \mid X_n)$. Pravdepodobnosť prechodu medzi stavmi $P\left(X_{n+1} = x_j \mid X_n = x_i\right)$ predstavuje pravdepodobnosť, kedy sa reťazec za jednotku času posunie zo stavu x_i do stavu x_j. Pravdepodobnostné rozloženie stavov v reťazci v čase $n \in T$ je dané vektorom [6]:

$$\boldsymbol{p}(n) = (p_n(x),\, x \in \boldsymbol{\chi}),\ \text{kde}\, p_n(x) = P(X_n = x), \forall x \in \boldsymbol{\chi}.$$

Markovov reťazec sa nazýva **homogénny** alebo **stacionárny v čase**, ak pravdepodobnosti prechodu nezávisia od času, t. j. [6]:

$$p_{ij} = P\left(X_{n+k+1} = x_j \mid X_{n+k} = x_i\right) = P\left(X_{n+1} = x_j \mid X_n = x_i\right),$$

kde k je prirodzené číslo. Pravdepodobnosti prechodu potom p_{ij} vieme usporiadať do matice pravdepodobností prechodu $\mathcal{P} = \left(p_{ij}\right)_{x_i, x_j \in \boldsymbol{\chi}}$.

Poznámka. V tejto práci sa obmedzíme na homogénne Markovove reťazce.

Kľúčovou súčasťou teórie Markovových reťazcoch v MCMC metódach je skúmanie jeho asymptotického správania pravdepodobnosti. Za predpokladu existencie takého rozdelenia π, pre ktoré platí [6], [25]:

$$\lim_{n \to \infty} \boldsymbol{p}(n) = \boldsymbol{\pi},$$

Markovov reťazec konverguje k náhodnej premennej X_n pre $n \to \infty$. Takéto rozdelenie $\boldsymbol{\pi}$ sa nazýva **limitné**, resp. rovnovážne, a spĺňa vzťah $\boldsymbol{\pi}\mathcal{P} = \boldsymbol{\pi}$. Limitné rozdelenie je aj stacionárne rozdelenie, avšak v prípade, že limitné rozdelenie existuje, tak je jediné stacionárne rozdelenie. Ak pre Markovov reťazec existuje práve jedno limitné rozdelenie, tak ho nazývame ergodický.

Dôležitým prístupom v MCMC metódach je **vlastnosť reverzibility** Markovovho reťazca. Markovov reťazec sa nazýva reverzibilným, ak je schopný pohybovať sa oboma smermi v rámci svojich stavov (dopredu, ale aj dozadu), a to s rovnakou pravdepodobnosťou. Ak hovoríme o rozdelení pravdepodobnosti Markovovho reťazca, ktorý má vlastnosť reverzibility, tak hovoríme (pre jednoduchosť) o ľubovoľných dvoch stavoch z tohto reťazca X_i a X_{i+1}, pre ktoré zmena poradia spätne zo stavu X_{i+1} do stavu

X_i neznamená zmenu pravdepodobnosti prechodu medzi týmito stavmi. V skutočnosti na to, aby sa Markovov reťazec nazýval reverzibilným potrebujeme, aby vlastnosť reverzibility platila pre celý reťazec, pričom reverzibilita implikuje stacionaritu, z ktorej sa overuje, či dané rozdelenie je aj limitným rozdelením. Stacionarita znamená stabilitu reťazca smerom dopredu, avšak reverzibilita predstavuje silnejšiu vlastnosť, a to stabilitu reťazca v oboch smeroch, čo sa neskôr v časti 5.3 ukáže ako dôležitá vlastnosť pri MCMC metódach [6].

Pri Monte Carlo Markovových reťazcoch sa zvyčajne stretávame s charakteristikou všeobecnej množiny stavov. V nasledujúcej časti si preto stručne predstavíme základnú teóriu Markovových reťazcov pre všeobecnú množinu stavov, pričom sa v mnohých prípadoch vlastností Markovových reťazcov odvolávame na odbornú literatúru.

5.1.2 Markovove reťazce so všeobecnou množinou stavov

V prípade MCMC techník bolo pre potreby modelovania so spojitými rozdeleniami pravdepodobnosti nutné rozšíriť spočítateľnú množinu stavov na všeobecnú množinu stavov $\boldsymbol{\chi}$, na ktorej existuje σ-algebra, označíme ju zápisom $\mathcal{F}(\boldsymbol{\chi})$. Markovov reťazec pre všeobecnú množinu stavov je definovaný ako [6], [25]:

$$\begin{aligned} &P\left(X_{n+1} \in A \,|\, X_n = x_n,\, X_{n-1} = x_{n-1},\, \ldots, X_0 = x_0\right) = \\ &= P\left(X_{n+1} \in A \,|\, X_n = x_n\right), \qquad \forall\, n \in T,\, \forall\, A \in \mathcal{F}(\boldsymbol{\chi}), \forall\, x_j \in \boldsymbol{\chi}, \end{aligned}$$

pričom počiatočné rozdelenie pravdepodobnosti je určené pomocou pravdepodobnostnej miery μ, ktorá je definovaná na sigma algebre $\mathcal{F}(\boldsymbol{\chi})$ a pravdepodobnosti prechodu medzi stavmi sú dané prechodovým jadrom (*transition kernel*) daným

$$P(x,\, A) = P\left(X_{n+1} \in A \,|\, X_n = x\right), \qquad \forall\, x \in \boldsymbol{\chi},\, \forall\, A \in \mathcal{F}(\boldsymbol{\chi}).$$

Pravdepodobnostnú mieru π, definovanú na $(\boldsymbol{\chi}, \mathcal{F}(\boldsymbol{\chi}))$, ktorá spĺňa [6], [25]:

$$\pi(A) = \int_{\boldsymbol{\chi}} \pi(x)\, \mathrm{d}P(x,\, A),$$

nazývame **stacionárnou mierou Markovovho reťazca**. Rovnako, ako v prípade Markovových reťazcov s konečnou množinou stavov, tak aj teraz platí, že ak Markovov reťazec dosiahne limitnú mieru, tak táto miera je aj stacionárnou mierou.

Pre overenie stacionarity miery sa podobne, ako v prípade množín s konečným počtom prvkov, používa vlastnosť reverzibility, ktorá za podmienky existencie pravdepodobnostnej miery π, pre ktorú prechodové jadro spĺňa [6], [25]:

$$\int_A \pi(x)\,\mathrm{d}P(x,\,B) = \int_B \pi(x)\,\mathrm{d}P(x,\,A), \quad \forall\, A, B \in \mathcal{F}(\boldsymbol{\chi}) \tag{14}$$

sa Markovov reťazec nazýva reverzibilný. Ak v (14) položíme $A = \boldsymbol{\chi}$, tak stacionarita π je dôsledkom reverzibility reťazca

$$\int_{\boldsymbol{\chi}} \pi(x)\,\mathrm{d}P(x,\,B) = \int_B \pi(x) = \pi(B), \quad \forall\, B \in \boldsymbol{\chi}. \tag{15}$$

Podobne ako v prípadne konečnej množiny stavov, tak aj pre všeobecnú množinu stavov voľne uvedieme základný koncept ergodických Markovových reťazcov, potrebných pre konvergenciu konvergenčnú teóriu, pričom presné znenia jednotlivých vlastností je možné nájsť v publikáciách [6], [38].

K zabezpečeniu konvergencie Markovovho reťazca je potrebné vyžadovať, aby bol ergodický, čo v prípade všeobecnej množiny stavov znamená Harris-rekurentnosť a aperiodicitu. Na definíciu Harris-rekurentnosti potrebujeme poznať charakteristiku ϕ-neredukovateľnosti, pre ktorú je Markovov reťazec ϕ-**neredukovateľný** vtedy a len vtedy, keď existuje taká miera ϕ na $(\boldsymbol{\chi}, \mathcal{F}(\boldsymbol{\chi}))$, pre ktorú je pravdepodobnosť nadobudnutia stavu z množiny A s nenulovou mierou (vzhľadom na mieru ϕ) nenulová pre ľubovoľný začiatočný bod $x \in \boldsymbol{\chi}$ [6].

Aperiodicita Markovovho reťazca predstavuje vlastnosť náhodného pohybu medzi jednotlivými stavmi bez nejakého pravidelného vzoru. Pre ϕ-neredukovateľný a aperiodický Markovov reťazec platí vlastnosť konvergencie vzhľadom na mieru π skoro všade pre $\forall\, x \in \boldsymbol{\chi}$. Konvergencia „skoro všade“ však nestačí, pretože Markovov reťazec je inicializovaný bodom, čo predstavuje množinu s nulovou mierou [6].

Daný ϕ-neredukovateľný reťazec stavov $\{X_t\}_{t\in T}$ nazývame **Harris-rekurentný**, ak pre množiny A nenulovej miery (vzhľadom na ϕ) sú všetky množiny A Harris-rekurentné, teda pre každú množinu A platí [6]:

$$P\left(\{X_t\}_{t\in T} \in A \text{ nekonečne často} \,|\, X_0 = x\right) = 1, \quad \forall x \in A.$$

Markovov reťazec nazývame **ergodický**, ak je Harris-rekurentný a aperiodický.

V takom prípade pre stacionárnu mieru pravdepodobností π platí [6], [38]:

$$\lim_{n\to\infty} \|P^n(x, \cdot) - \pi(\cdot)\| = 0, \qquad \forall\, x \in \boldsymbol{\chi}, \tag{16}$$

kde $P^n(x, \cdot)$ predstavuje prechodové jadro n-tého rádu, ktoré popisuje pravdepodobnosti prechodu medzi stavmi i a j v Markovovom reťazci po n krokoch a $\|\cdot\|$ označuje celkovú variačnú vzdialenosť, ktorej definíciu je možné nájsť napríklad v publikácií [6] na strane 33. Zo vzťahu (16) vyplýva, že vzdialenosť medzi distribúciami sa medzi prechodovým jadrom $P^n(x, \cdot)$ a stacionárnou distribúciou $\pi(\cdot)$ zmenšuje do bodu, v ktorom sa tieto dve distribúcie stávajú ekvivalentnými.

5.1.3 Konvergenčná teória Markovových reťazcov

Konvergenčná teória markovových reťazcov poskytuje matematický rámec pre analýzu spávania sa Markovovho reťazca v ustálenom – stacionárnom rozdelení, ktoré je cieľom MCMC analýzy. Pre homogénne Markovove reťazce pri splnení podmienok konvergencie platia matematické tvrdenia – zákon veľkých čísel a centrálna limitná veta [6].

Pre realizácie $x^{(1)}, \ldots, x^{(n)}$ Markovovho reťazca $\{X_t\}_{t\in T}$ definujme [6]:

$$\mu = \mathrm{E}_\pi\left[g(X)\right] = \int_{\boldsymbol{\chi}} g(x)\, \pi(\mathrm{d}x),$$

kde E_π označuje strednú hodnotu počítanú vzhľadom na cieľové rozdelenie pravdepodobnosti π. Odhad μ definujeme predpisom:

$$\hat{\mu}_N = \frac{1}{N}\sum_{n=1}^{N} g\left(x^{(n)}\right).$$

Pre integrovateľnú funkciu $g(x)$ takú, že $\mathrm{E}_\pi\left[|g(x)|\right] < \infty$, a ergodickým Markovovym reťazcom so stacionárnym rozdelením pravdepodobnosti π pre ľubovoľné počiatočné rozdelenie funkcia konverguje $\hat{\mu}_N$ podľa zákona veľkých čísel [38]:

$$\hat{\mu}_N \overset{\text{s. v.}}{\to} \mu$$

pre $N \to \infty$ skoro všade (označujeme skratkou s. v.), pričom zákon veľkých čísel je platný pre akýkoľvek ergodický Markovov reťazec.

Pre uplatnenie centrálnej limitnej vety ergodicita Markovovho reťazca nestačí, a teda vyžaduje si ďalšie podmienky na ergodicitu reťazca, ktorú v tejto práci nejdeme

pre jej komplexnosť a rozsah uvádzať, no je možné ju nájsť napríklad v [38]. Za splnených predpokladov existuje podľa centrálnej limitnej vety číslo σ, pre ktoré konverguje rozdelenie

$$\sqrt{N}\left(\hat{\mu}_N - \mu\right)$$

slabo k normálnemu rozdeleniu so strednou hodnotou 0 a varianciou σ^2 pre ľubovoľné počiatočné rozdelenie a pre $N \to \infty$, pričom ďalšie podrobnosti a analytický tvar variancie σ^2 je možné nájsť napríklad v publikáciách [6] a [38].

5.2 Princíp metódy Markov Chain Monte Carlo

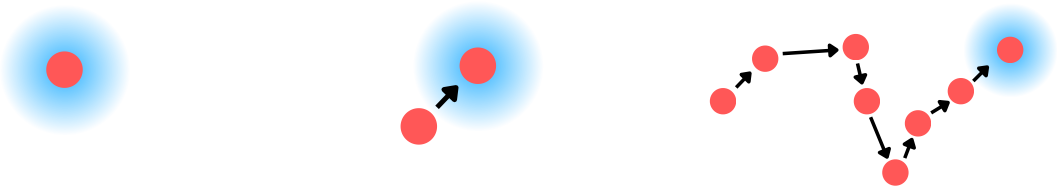
Základný princíp metód MCMC spočíva vo vytvorení takého prechodového jadra (angl. *transition kernel*), aby množina stavov Markovovho reťazca bola zhodná s nosičom cieľového rozdelenia a aby cieľové rozdelenie bolo limitným rozdelením tohto prechodového jadra [6]. Pre jednoduchšiu interpretáciu si predstavme, že prechodové jadro $T\left(x' \,|\, \cdot\right)$ je také rozdelenie pravdepodobnosti, ktorého podmienená funkcia hustoty $\tau\left(x|x'\right)$ na základe zadaného počiatočného stavu x', definuje bod x zo súboru náhodných bodov, do ktorého sa s najväčšou istotou systém posunie, pričom v novom bode x získavame novú distribúciu bodov. Opakovaním tohto procesu získame Markovov reťazec, ktorý sám o sebe nie je ničím výnimočným, pretože prakticky predstavuje iba reťazec náhodných bodov. Ak však princíp prechodového jadra skonštruujeme tak, že bude dosahovať cieľove rozdelenie $\pi\left(x\right)$

$$\pi\left(x\right) = \int_{\Omega} \tau\left(x \,|\, x'\right) \pi\left(x'\right) \,\mathrm{d}x' \,, \tag{17}$$

tak Markovov reťazec má tendenciu približovať sa k tomuto rozdeleniu, a navyše ním aj prechádzať, a to bez ohľadu na počiatočný stav v parametrickom priestore [2].

V teórii Markovových reťazcov sme sa stretli s princípom, kedy sme pri známom prechodovom jadre hľadali jeho limitné rozdelenie. MCMC metódy sú založené na opačnom princípe, kedy pre známe limitné rozdelenie hľadáme prechodové jadro, pričom v nasledujúcej časti čerpáme z [6].

Na vytvorenie prechodového jadra $T\left(x' \,|\, \cdot\right)$ sa využíva pomocné prechodové jadro $\Phi\left(x' \,|\, \cdot\right)$, pričom $\Phi\left(x' \,|\, \cdot\right)$ je také rozdelenie pravdepodobnosti, ktoré umožňuje generovanie náhodných vzoriek, s takou funkciou hustoty $\phi\left(x \,|\, x'\right)$, ktorú je možné vyčísliť



i) ii) iii)

Obr. 1: i) Markovov reťazec v úvodnej fáze pozostáva z počiatočného bodu (červený), ktorý sa nachádza v parametrickom priestore, pričom prechodové jadro (modré), predstavuje podmienenú hustotu, ktorá definuje pravdepodobnosť posunu do nového bodu x za predpokladu, že proces sa nachádza v bode x'. ii) Generovaním z tejto hustoty sa proces presunie do nového bodu x, čím vznikne nová distribúcia bodov. iii) Opakovaním vzniká Markovov reťazec. (*zdroj*: vlastné spracovanie podľa [2])

v akomkoľvek bode priestoru $\boldsymbol{\chi}$. V praxi sa môže jednať aj o komplikované rozdelenie, ktoré však musí spĺňať túto vlastnosť. Na to, aby sme sa mohli posunúť zo stavu x', potrebujeme vygenerovať z pomocného prechodného jadra $\Phi\left(x' \,|\, \cdot\right)$ možných kandidátov, pričom posun do nového stavu x schválime, resp. zamietneme, pomocou akceptačného kritéria (angl. *acceptance probability*) $\alpha\left(x' \,|\, x\right)$. Samotné prechodové jadro potom pozostáva zo súčtu pravdepodobnosti nadobudnutia nového stavu x a pravdepodobnosti zotrvania v stave x' dané [6]

$$T\left(x' \,|\, \Omega\right) = \int_{\Omega} \phi\left(x' \,|\, x\right) \alpha\left(x' \,|\, x\right) \,\mathrm{d}x + I_{\Omega}\left(x'\right) \left(1 - \int_{\chi} \phi\left(x' \,|\, x\right) \alpha\left(x' \,|\, x\right) \,\mathrm{d}x\right),$$

kde $\boldsymbol{\chi}$ predstavuje podmnožinu $\mathbb{R}^d$ a $\Omega \subseteq \boldsymbol{\chi}$, $I_{\Omega}\left(x'\right)$ je indikátor množiny Ω a hustota $\phi\left(x' \,|\, x\right) \alpha\left(x' \,|\, x\right)$ definuje hustotu pravdepodobnosti posunu systému zo stavu x' do bodu x. Od akceptačného kritéria vyžadujeme splnenie reverzibility cieľového rozdelenia $\pi\left(\cdot\right)$ vzhľadom na prechodové jadro, čo zabezpečuje ustálenie rozloženia pravdepodobností v Markovovom reťazci podľa cieľového rozdelenia. Pre splnenie vlastnosti limitného rozdelenia požadujeme od Markovovho reťazca taktiež aj splnenie podmienok konvergencie, čiže ϕ-neredundantnosť a rovnako aj aperiodicita. Za týchto predpokladov spĺňa Markovov reťazec zákon veľkých čísel, čo znamená, že charakteristiky cieľo-

vého rozdelenia môžeme vypočítať pomocou metódy Monte Carlo pomocou vzťahu (12) s použitím náhodného výberu $f(x)$, pre realizácie Markovovho reťazca $x^{(1)}, \ldots, x^{(N)}$ [6]:

$$\mu \approx \frac{1}{N} \sum_{n=1}^{N} g\left(x^{(n)}\right) = \hat{\mu}_N.$$

Vlastnosť konvergencie ku skutočným odhadom je v skutočnosti najmä teoretickou záležitosťou, ktorá je pre obmedzené výpočtové kapacity ťažko uplatniteľná v praxi, keďže počet tranzícií (iterácií v MCMC) nie je možne uplatniť pre $N \to \infty$. Na získanie dôveryhodných odhadov pri obmedzenom počte tranzícií v MCMC metóde je potrebné pochopiť, ako Markovov reťazec funguje a ako z neho získať čo najviac, a zároveň čo najmenej vychýlených odhadov, čomu sa venujeme v kapitole 6. Predtým sa však zameriame na „srdce“ samotných MCMC metód, ktorým je algoritmus na vytvorenie prechodového jadra.

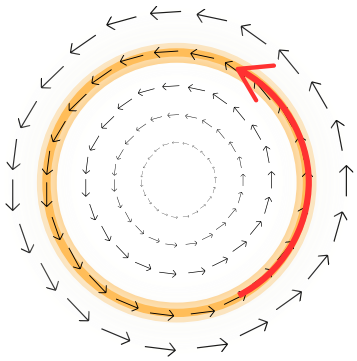
5.3 Hamiltonovský Monte Carlo algoritmus

Pripomeňme si, že to, čo od MCMC techník potrebujeme, je preskúmať aposteriórne cieľové rozdelenie a vzhľadom k typicky náročnému tvaru tohto rozdelenia chceme získať z neho istú aproximáciu. MCMC generuje vzorky z aposteriórneho rozdelenia postupným pohybom Markovovho reťazca, ktorý zostrojíme pomocou pomocného prechodového jadra. Doteraz bolo vyvinutých niekoľko algoritmov, ktoré sú však modifikáciami dvoch základných prístupov, ktorými sú Metropolisov-Hastingsov a Gibbsov algoritmus. V našej práci sa venujeme výhradne Hamiltonovej Monte Carlo metóde, ktorá sa v súčasnosti dostáva do popredia a predstavuje efektívny hybrid zo spomínaných metód. V nasledujúcej časti si predstavíme základné princípy, na ktorých stojí HMC, ktoré zosumarizujeme do matematických tvrdení. Vysvetlíme si výhody HMC oproti iným metódam a nakoniec túto metódu rozšírime o doplnok, pričom principiálne vychádzame z [2].

5.3.1 Princíp Hamiltonovskej Monte Carlo metódy

V časti 5.1 sme si predstavili teóriu Markovových reťazcov, ktorá je v očiach matematika silným nástrojom využiteľným v rôznych oblastiach matematiky. V praxi

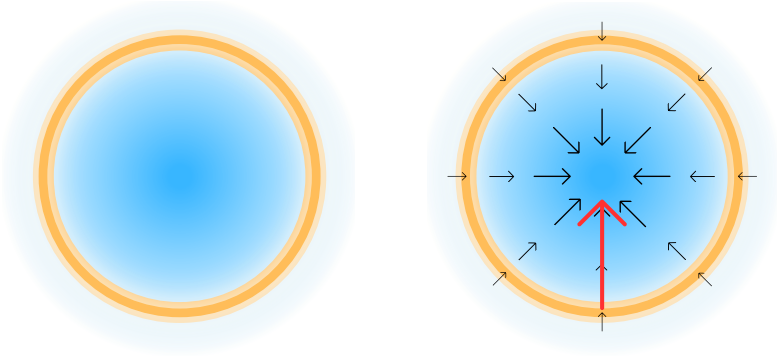
však problém Markovových reťazcov často spočíva v obmedzených výpočtových možnostiach, patologických prípadoch „nepekných“ cieľových funkcií či v neefektívnych algoritmoch. To, čo od Markovovho reťazca potrebujeme, je, aby sa reťazec po dosiahnutí cieľového rozdelenia v ňom efektívne a vhodne pohyboval. Takáto požiadavka sa môže zdať ako nemožná vzhľadom na istú náhodnosť, ktoré so sebou Markovove reťazce nesú. Riešenie v skutočnosti spočíva vo využití vektorového poľa, ktoré predstavuje smer každého bodu nachádzajúceho sa v parametrickom priestore, pričom ak docielime zosúladenie smeru bodov vo vektorov poli tak, aby prechádzalo cieľovým rozdelením, tak dosiahneme efektívne a cielené skúmanie neznámych oblastí cieľového rozdelenia ako zobrazujeme na obrázku 2.



Obr. 2: Grafické znázornenie vektorového priestoru (čierne šípky) a ľubovoľného cieľového rozdelenia (oranžové). Našou snahou je skonštruovať vektorové pole a nájsť takú trajektóriu, ktorá prechádza cieľovým rozdelením (červená šípka). (*zdroj*: vlastné spracovanie podľa [2])

Pre úplné pochopenie princípu vektorového poľa si predstavme funkciu ako cieľové rozdelenie v parametrickom priestore, pričom tejto funkcii potrebujeme pridať ďalšiu informáciu, ktorou je smer, ktoré predstavuje vektorové pole. Pre ľubovoľný bod vo vektorovom priestore existuje smer, ktorým sa posunieme do nového bodu, pre ktorý taktiež existuje smer, pričom tento postup nás posúva k novým stavom procesu, čím postupne získavame trajektóriu. Skonštruovanie takéhoto vektorového poľa, ktoré prechádza cieľovým rozdelením nie je jednoduchá záležitosť. Pre jednoduchosť si graficky načrtnime na obrázku 3 situáciu s cieľovým rozdelením $\pi(x)$ označeným oranžovou farbou a jeho hustotu $\pi(x')$ označenou modrou farbou. Prvý krok k skonštruovaniu požadovaného vektorového poľa pozostáva z využitia gradientu funkcie hustoty cieľového rozdelenia. Geometrický význam gradientu spočíva v ukazovaní smeru najväčšieho

rastu funkcie, v našom prípade hustoty, čo nám však nepomôže, pretože vektorové pole bude smerovať dovnútra k najväčšej distribúcii bodov rozdelenia. Keďže funkcia hustoty je silne previazaná s parametrizáciou, teda strednou hodnotou a inými parametrami, tak gradient bude smerovať k oblastiam závislých od hodnôt parametrov tohto rozdelenia. Naším cieľom teda je transformovať smer vektorového poľa gradientov, a pritom zabezpečiť nezávislosť vzhľadom na parametre tak, aby trajektória, ktorú takto získame, prechádzala cieľovým rozdelením.

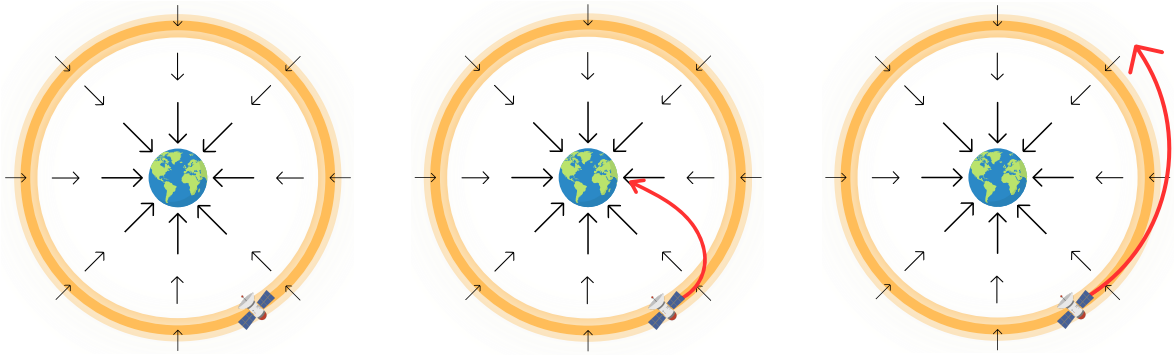


i) ii)

Obr. 3: i) Uvažujme ľubovoľnú cieľovú funkciu (oranžová) a jej funkciu hustoty (modrá) v parametrickom priestore.
ii) Na skonštruovanie vektorového priestoru využijeme gradient funkcie hustoty (čierne šípky), ktoré smerujú k miestam najväčšieho rastu funkcie hustoty, v našom prípade miestam s najväčšou hustotou (červená šípka).
(*zdroj*: vlastné spracovanie podľa [2])

Naše požiadavky na zostrojenie takéhoto vektorového poľa sa môžu zdať príliš abstraktné a ťažko pochopiteľné, ak však zájdeme do fyziky, tak tam nachádzame inšpiráciu na riešenie nášho problému. Vhodným príkladom je satelit nachádzajúci sa na orbite, resp. obežnej dráhe okolo planéty Zem, ktorý uvádzame na obrázku 4. V tomto prípade orbita predstavuje cieľové rozdelenie, ktorým chceme prechádzať tak, ako na satelit putuje po orbite a gravitačné pole predstavuje vektorové pole znázorňujúce smer sily (gradient), ktorou je satelit priťahovaný k Zemi. Ak umiestnime satelit na orbitu a necháme na neho pôsobiť gravitačnú silu, tak satelit nabúra do Zeme, čo predstavuje analógiu s gradientom vektorového poľa smerujúcom k oblasti s najväčšou hustotou. Pohľad fyzika nám však poskytuje riešenie, ktoré spočíva v pridaní hybnosti satelitu

tak, aby zostal na orbite. Je dôležité docieliť to, že hybnosti nebude málo a ani priveľa, pretože v takom prípade by mohol satelit z orbity pomaly zísť a nabúrať do Zeme, poprípade vyletieť z orbity von do vesmíru.



i) ii) iii)

Obr. 4: i) Analogický príklad k vektorovému poľu a pohybe po cieľovom rozdelení predstavuje gravitačné pole Zeme (čierne čípky) a satelit na orbite (oranžová). ii) Pri nízkej hybnosti satelitu na orbite by satelit z orbity zišiel a nabúral do Zeme. iii) Pri scenári s vysokou hybnosťou by satelit z orbity vyletel do vesmíru. (*zdroj*: vlastné spracovanie podľa [2])

Analógiou v príklade so satelitom vo vesmíre sme získali návod na to, ako zostrojiť trajektóriu v cieľovom rozdelení v pravdepodobnostnom priestore tak, aby týmto rozdelením prechádzala, ktorý spočíva v hybnosti. Parametrický priestor teda potrebujeme rozšíriť o nový parameter hybnosti, pričom musíme dbať na to, aby sme pridaním nových parametrov zachovali pôvodný pravdepodobnostný model, a teda aby sme vzorky získavali z požadovaného cieľového rozdelenia.

5.3.2 Hamiltonova Monte Carlo metóda v modeloch stochastickej volatility

Po vzore príkladu s pohybom satelitu po družici sa dostávame ku fyzikálnym zákonom, ktoré potrebujeme pre zachovanie stabilnej dynamiky. V konzervatívnej dynamike fyzikálnych systémov je dôležité, aby sa zachovával objem priestoru, ktorý je definovaný v priestore pozície a hybnosti. Ak nastane zmena, napríklad expanzia v priestore pozície, tak hybnosť sa musí zmenšiť tak, aby výsledný objem zostal nezmenený a tým pádom požadovaná dynamika nášho systému stabilná. Z tohto dôvodu

od parametrov hybnosti požadujeme opačné správanie v porovnaní so zmenou hodnôt parametrov tak, aby objem zostal nezmenený, čo dosiahneme zavedením parametera hybnosti $\boldsymbol{p}$ pre každú dimenziu cieľového rozdelenia. Teda pre vzorky $\boldsymbol{q} = (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n)$, kde $\boldsymbol{q} = (\boldsymbol{\theta}, h_1, \ldots, h_n)$ predstavuje vektor parametrov, budeme potrebovať parametre hybnosti $\boldsymbol{p} = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$, čím dôjde k rozšíreniu dimenzie z $k + n$ na dimenziu $2\,(k + n)$ nasledovne [2]:

$$\boldsymbol{q} \to (\boldsymbol{q}, \boldsymbol{p})\,.$$

Ďalej pre rozšírený parametrický priestor potrebujeme rozšíriť aj cieľové rozdelenie, teda nájsť pravdepodobnostnú štruktúru pre parametre hybnosti spolu s hľadanými parametrami

$$\pi\,(\boldsymbol{q}) \to \pi\,(\boldsymbol{q}, \boldsymbol{p})\,,$$

pričom súčinom cieľového rozdelenia $\pi\,(\boldsymbol{q})$ získaného z Bayesovho vzorca (17) a podmienenej hustoty parametrov hybnosti $\pi\,(\boldsymbol{p} \,|\, \boldsymbol{q})$ dostávam združenú hustotu, ktorá sa nazýva ako kanonické rozdelenie [2]:

$$\pi\,(\boldsymbol{q}, \boldsymbol{p}) = \pi\,(\boldsymbol{p} \,|\, \boldsymbol{q})\,\pi\,(\boldsymbol{q})\,. \tag{18}$$

Dôležitým faktom je, že parametre hybnosti $\boldsymbol{p}$ „dvíhajú“ trajektórie pôvodného systém do $2\,(k + n)$-dimenzionálneho priestoru, pričom ak vo vzťahu (18) „vyhodíme“ parametre hybnosti, tak sa opäť dostaneme k pôvodnému cieľovému rozdeleniu, čím teda zabezpečíme, že trajektórie prechádzajú kanonickým rozdelením, tak tieto trajektórie vieme projektovať naspäť do požadovaného cieľového rozdelenia.

Vzhľadom k tomu, že parametre $\boldsymbol{p}$ a $\boldsymbol{q}$ sa správajú navzájom opačne v rámci konceptu zachovania konzervatívnej dynamiky, tak aj ich hustoty sa budú správať navzájom opačne, čo spôsobí, že výsledné kanonické rozdelenie nezávisí od parametrizácie a môžeme ju vyjadriť pomocou hamiltonovskej funkcie $H\,(\boldsymbol{q}, \boldsymbol{p})$ [2]:

$$\pi\,(\boldsymbol{q}, \boldsymbol{p}) = \mathrm{e}^{-H(\boldsymbol{q}, \boldsymbol{p})}, \tag{19}$$

ktorá je využívaná najmä vo fyzike pre svoju vlastnosť invariantnosti. Využitím terminológie z oblasti fyziky označujeme hodnotu hamiltoniánu v ľubovoľnom bode ako energiu. Zo vzťahu (19) si vyjadríme Hamiltonovu funkciu, čím dostávame:

$$H\,(\boldsymbol{q}, \boldsymbol{p}) = -\ln \pi\,(\boldsymbol{q}, \boldsymbol{p})\,,$$

čo vieme ďalej zjednodušiť pomocou vzťahu pre kanonické rozdelenie (18) a vlastnosti logaritmu na tvar [2]

$$\begin{aligned} H(\boldsymbol{q}, \boldsymbol{p}) &= -\ln \pi(\boldsymbol{p} \,|\, \boldsymbol{q})\, \pi(\boldsymbol{q}), \\ &= -\ln \pi(\boldsymbol{p} \,|\, \boldsymbol{q}) - \ln \pi(\boldsymbol{q}), \\ &\equiv K(\boldsymbol{p}, \boldsymbol{q}) + V(\boldsymbol{q}), \end{aligned}$$

pričom $V(\boldsymbol{q}) = -\ln \pi(\boldsymbol{q})$ predstavuje **potenciálnu energiu**, ktorú tvorí transformácia cieľového rozdelenia, ktorú sa snažíme modelovať a získať z nej odhady, a preto podoba potenciálnej energie závisí od vlastností cieľového rozdelenia, a teda jeho parametrov. Člen $K(\boldsymbol{p}, \boldsymbol{q}) = -\ln \pi(\boldsymbol{p} \,|\, \boldsymbol{q})$ je označovaný ako **kinetická energia** a je nezávislý od cieľového rozdelenia, a preto jej tvar musí byť explicitne špecifikovaný, pričom platí, že pre ľubovoľnú voľbu kinetickej energie získavame vhodný algoritmus na skúmanie cieľového rozdelenia.

Na generovanie vektorového poľa, ktoré je „zosúladené“ s požadovanou množinou, sa využívajú Hamiltonove funkcie, ktoré popisujú systém a jeho dynamiku v priestore pozície a hybnosti využívaný najmä v mechanike. V našom prípade je priestor pozície a hybnosti prezentovaný ako priestor hľadaných parametrov $\boldsymbol{q}$ a hybnosti $\boldsymbol{p}$, pre ktoré majú Hamiltonove rovnice tvar [2]:

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}\boldsymbol{t}} &= +\frac{\partial H}{\partial \boldsymbol{p}} = \frac{\partial K}{\partial \boldsymbol{p}}, \\ \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}\boldsymbol{t}} &= -\frac{\partial H}{\partial \boldsymbol{q}} = -\frac{\partial K}{\partial \boldsymbol{q}} - \frac{\partial V}{\partial \boldsymbol{q}}. \end{aligned}$$

Rovnice sú formulované tak, aby smer (logaritmicky transformovaného) gradientu cieľového rozdelenia $\frac{\partial V}{\partial \boldsymbol{q}}$ bol riadený hybnosťou a aby trajektórie viedol v časoch $\boldsymbol{t}$ rýchlo a efektívne v rámci cieľového rozdelenia smerom k miestam s najväčšou pravdepodobnosťou. V prípade, kedy sa mení potenciálna energia v priestore parametrov, mení sa aj hybnosť systému tak, aby bola zachovaná celková energia systému, a teda aj princíp zachovania konzervatívnej dynamiky. Unikátnu funkciu zobrazenia bodu $(\boldsymbol{q}_0, \boldsymbol{p}_0)$ do bodu $(\boldsymbol{q}_t, \boldsymbol{p}_t)$ označíme ako $S_t : (\boldsymbol{q}_0, \boldsymbol{p}_0) \to (\boldsymbol{q}_t, \boldsymbol{p}_t)$, pričom S_t spĺňa Hamiltonove rovnice pre $\forall\, t \in T$.

5.3.3 Hamiltonovská Markovova tranzícia

Na vytvorenie trajektórie Markovovho reťazca HMC metódou potrebujeme hamiltonovskú transformáciu, ktorá „zdvihne“ parametre $\boldsymbol{q}$ do parametrického priestoru

$$\boldsymbol{q} \to (\boldsymbol{q}, \boldsymbol{p}),$$

čo umožňuje jej pravdepodobnostná štruktúra, ktorú sme definovali pomocou podmieneného rozdelenia $\pi(\boldsymbol{p} \,|\, \boldsymbol{q})$. Počiatočnému bodu $\boldsymbol{q}_1$, pre ktorý predpokladáme, že sa nachádza v cieľovom rozdelení, pridáme hybnosť, ktorú získame ako realizáciu z podmienenej hustoty pri danom $\boldsymbol{q}_1$, teda:

$$\boldsymbol{p}_1 \sim \pi(\boldsymbol{p} \,|\, \boldsymbol{q}_1),$$

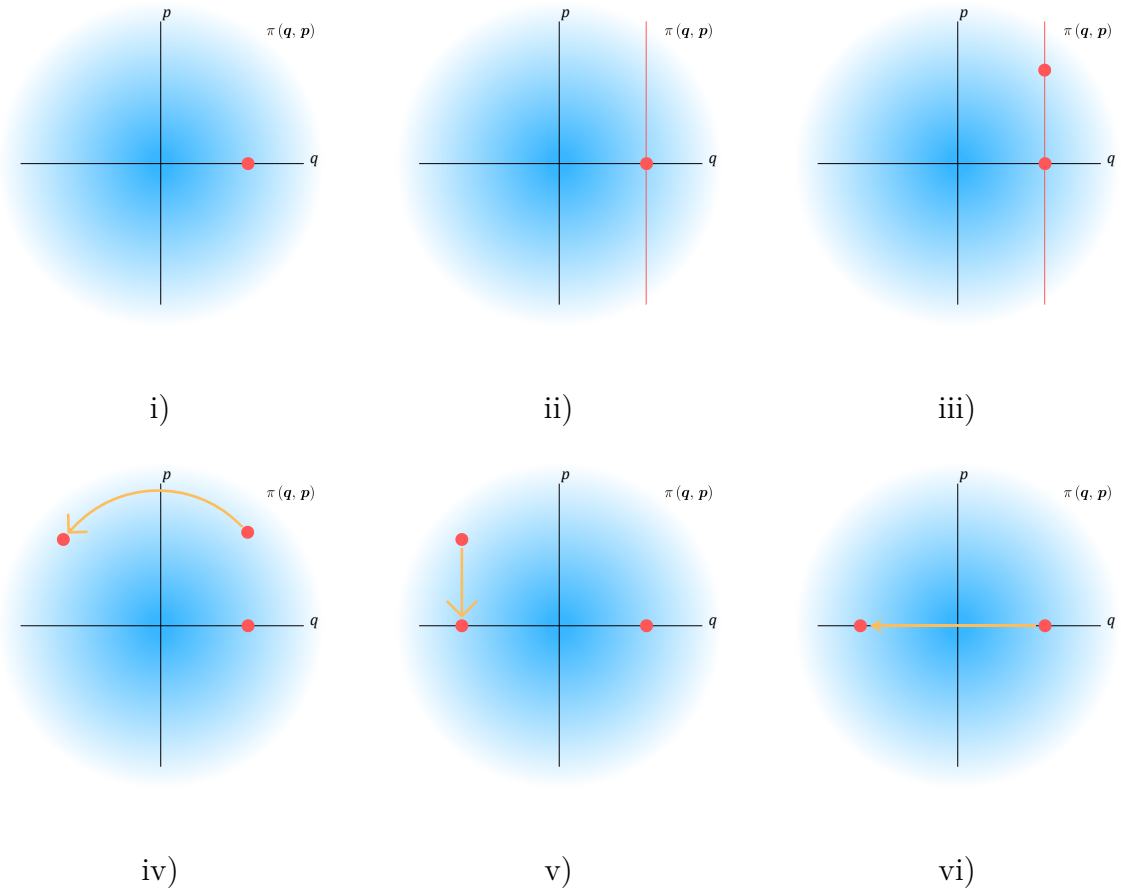
čím sa systém rozšíri do $2(k+n)$-dimenzionálneho priestoru trajektórií, pričom pre bod $(\boldsymbol{q}_1, \boldsymbol{p}_1)$ získame jeho trajektóriu z riešenia Hamiltonových rovníc (skr. $\mathrm{H}(\cdot)$), podľa ktorej sa vo vektorovom poli presunie do nového bodu $(\boldsymbol{q}'_1, \boldsymbol{p}'_1)$

$$(\boldsymbol{q}_1, \boldsymbol{p}_1) \to \mathrm{H}(\boldsymbol{q}_1, \boldsymbol{p}_1) = \left(\boldsymbol{q}'_1, \boldsymbol{p}'_1\right)$$

Presun po trajektórii je efektívny, nakoľko sa nový bod $(\boldsymbol{q}'_1, \boldsymbol{p}'_1)$ nenachádza v blízkosti pôvodného bodu, ktorý sme už preskúmali, no zároveň sa stále nachádza v požadovanom cieľovom rozdelení, hoci zatiaľ v jej „zdvihnutej“ podobe. Spätnou projekciou novozískaného bodu $(\boldsymbol{q}'_1, \boldsymbol{p}'_1)$ do $(k+n)$-dimenzionálneho priestoru parametrov $\boldsymbol{q}$

$$(\boldsymbol{q}, \boldsymbol{p}) \to \boldsymbol{q}$$

získame požadovaný bod Markovovho reťazca, ktorý sa nachádza v požadovanom rozdelení v priestore parametrov a navyše sa nachádza mimo okolia pôvodného bodu $\boldsymbol{q}_1$. Presunom v pôvodnom priestore z bodu $\boldsymbol{q}_1$ do $\boldsymbol{q}'_1$ teda získame tranzíciu v priestore parametrov v rámci cieľového rozdelenia. Spätnú transformáciou nám umožňuje vykonať tvar kanonického rozdelenia (19), ktoré po „vyhodení“ parametra hybnosti predstavuje cieľové rozdelenie $\pi(\boldsymbol{q})$. Celú hamiltonovskú Markovovu tranzíciu sme znázornili na obrázku 5. Hamiltonova Monte Carlo metóda umožňuje efektívne a rýchle skúmanie cieľového rozdelenia a to taktiež aj v chvostoch cieľového rozdelenia, kde sa pravdepodobnosť výskytu vzoriek znižuje, čím efektívne prispieva k presnejším odhadom parametrov v porovnaní s inými metódami.



Obr. 5: i) Priestor parametrov $\boldsymbol{q}$ rozšírime o parameter hybnosti, čím sa budeme pohybovať v priestore $\boldsymbol{q} \times \boldsymbol{p}$. Hamiltonovskú Markovovu tranzíciu začíname s počiatočným bodom $\boldsymbol{q}_1$.

ii) Pre počiatočný bod $\boldsymbol{q}_1$ (červený bod) vygenerujeme z množiny (červená priamka) hybnosť $\boldsymbol{p}_1$.

iii) Získavame teda bod $(\boldsymbol{q}_1, \boldsymbol{p}_1)$.

iv) Riešením Hamiltonových rovníc získavame trajektóriu, ktorá nám poskytne nový bod $(\boldsymbol{q}'_1, \boldsymbol{p}'_1)$ v priestore $\boldsymbol{q} \times \boldsymbol{p}$.

v) Spätnou transformáciou do priestoru parametrov $\boldsymbol{q}$ získavame bod v rámci požadovaného cieľového rozdelenia $\pi(\boldsymbol{q})$.

vi) Posun z bodu $(\boldsymbol{q}_1, \boldsymbol{p}_1)$ do bodu $(\boldsymbol{q}'_1, \boldsymbol{p}'_1)$ predstavuje efektívnu trajektóriu v rámci cieľového rozdelenia.

(*zdroj*: vlastné spracovanie podľa [2])

5.3.4 Voľba kinetickej energie

Fázový portrét v priestore $\boldsymbol{q} \times \boldsymbol{p}$ závisí od cieľového rozdelenia $\pi(\boldsymbol{q})$ a kinetickej energie $K(\boldsymbol{p}, \boldsymbol{q}) = -\ln \pi(\boldsymbol{p} \,|\, \boldsymbol{q})$, čo predstavuje širokú škálu možností pre nastavenie podmienenej hustoty pre parameter hybnosti $\boldsymbol{p} \sim \pi(\boldsymbol{p} \,|\, \boldsymbol{q})$. V praxi sa zvyčajne stretávame s voľbou viacrozmerného normálneho rozdelenia nezávislého od voľby parametra $\boldsymbol{q}$ a nezávislého medzi sebou, s nulovou strednou hodnotou a s Euklidovskou maticou vzdialeností M ako kovariančnou maticou

$$\boldsymbol{p} \sim \mathcal{N}_{2(k+n)}(\mathbf{0},\, M)\,,$$

čo vedie ku kinetickej energii v tvare [2]:

$$K(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^{\top} M^{-1} \boldsymbol{p}\,.$$

Voľbou hustoty parametrov hybnosti nezávislej od hodnoty parametra $\boldsymbol{q}$, teda s podmienenou hustotou v tvare $\pi(\boldsymbol{p} \,|\, \mathbf{0}) = \pi(\boldsymbol{p})$, sa zjednoduší tvar Hamiltonovej funkcie

$$H(\boldsymbol{q},\, \boldsymbol{p}) = K(\boldsymbol{p}) + V(\boldsymbol{q})\,,$$

a v konečnom dôsledku aj tvar Hamiltonových rovníc [2]:

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}\boldsymbol{t}} &= +\frac{\partial H}{\partial \boldsymbol{p}} = \frac{\partial K}{\partial \boldsymbol{p}}, \\ \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}\boldsymbol{t}} &= -\frac{\partial H}{\partial \boldsymbol{q}} = -\frac{\partial V}{\partial \boldsymbol{q}}\,. \end{aligned} \tag{20}$$

5.3.5 Vlastnosti dynamiky Hamiltonových rovníc

To, že môžeme použiť HMC metódu na získanie vzoriek z cieľového rozdelenia, vychádza z vlastnosti Hamiltonovej dynamiky, ktoré si predstavíme v nasledujúcej časti podľa [23], [39].

i) **Reverzibilita** Hamiltonovej dynamiky predstavuje „schopnosť“ systému posúvať sa naspäť v čase, čo znamená, že ak poznáme stav systému v čase $\boldsymbol{t} + \boldsymbol{s}$, tak sa vieme vrátiť v systéme do času $\boldsymbol{t}$. Táto vlastnosť je zabezpečená existenciou inverznej matice $S_{-\mathbf{s}}$ [23], [39] a je dôležitým prvkom pri modelovaní MCMC technikou, pretože nám pomáha overiť, že v priebehu času a iterácií sa modelované cieľové rozdelenie nezmenilo. Jednou z možností, ako to vieme overiť je dokázať, že

Markovov reťazec je reverzibilný, teda že platí rovnosť pravdepodobnosti prechodov medzi jednotlivými stavmi. Na dokázanie reverzibility Markovovho reťazca však potrebujeme reverzibilitu dynamiky generujúcej nové stavy – Hamiltonovu dynamiku, čím zabezpečíme správne fungovanie MCMC metód.

ii) **Zachovanie hamiltoniánu** opisuje vlastnosť Hamiltonovho systému, ktorý prispieva k stabilite dynamického systému. Tým, že energia v čase nepribúda a ani sa nestráca, máme stabilný a konzistentný dynamický systém. Využitím rovníc (20) dostávame [23], [39]:

$$\frac{\partial \mathrm{H}}{\partial \boldsymbol{t}} = \sum_{i=1}^{2(k+n)} \left[\frac{\partial \mathrm{H}}{\partial q_i} \frac{\mathrm{d}q_i}{\mathrm{d}t} + \frac{\partial \mathrm{H}}{\partial p_i} \frac{\mathrm{d}p_i}{\mathrm{d}t} \right] = \sum_{i=1}^{2(k+n)} \left[\frac{\partial \mathrm{H}}{\partial q_i} \frac{\partial \mathrm{H}}{\partial p_i} - \frac{\partial \mathrm{H}}{\partial p_i} \frac{\partial \mathrm{H}}{\partial q_i} \right] = 0.$$

iii) **Zachovávanie objemu** predstavuje požadovanú vlastnosť Hamiltonovho systému, ktorým je princíp zachovania konzervatívnej dynamiky a vychádza z Liouvilleovej vety [23], [39], ktorá hovorí o tom, že v priestore stavov sa objemy nemenia v čase, ak je dynamika systému deterministická a časová.

5.3.6 Výpočet Hamiltonových rovníc

Metóda HMC predstavuje silný nástroj na vytváranie tranzícií Markovovho reťazca, avšak problémom je, že sila tejto metódy spočíva v teórii. V praxi sa ukazuje, že systém Hamiltonových rovníc (20) je vo svojej podstate zložitou sústavou diferenciálnych rovníc, pre ktorú nevieme nájsť analytické riešenie [2], [23], [39]. Zložitosť toho procesu vyplýva ako zo zložitej funkcie hustoty aposteriórneho rozdelenia, tak aj z vysokej dimenzie parametrov. Na výpočet trajektórie sa preto v praxi využívajú numerické metódy, ktoré aproximujú analytické riešenie s určitou presnosťou. Bežné numerické metódy vykazujú vysokú chybovosť, ktorá je spojená s driftom, ktorý aproximovanú trajektóriu odkláňa od tej skutočnej, pričom drift ešte viac narastá vo zvyšujúcej sa dimenzii. Robustnou triedou numerických metód voči driftom sú symplektické metódy, ktoré generujú numerické trajektórie s vlastnosťou zachovávania objemu vo fázovom priestore podobne ako HMC metóda. Tým, že sa objem počas iterácií nemení sa zabezpečí, že numerická aproximácia bude okolo skutočného riešenia oscilovať [2].

V mnohých štatistických softvéroch, vrátane Stan [27], sa pri Euklidovo-Gaussovej voľbe kinetickej energie používa leapfrog metóda, resp. metóda preskakovania. Leapforg

metóda je numerickou metódou druhého rádu, ktorá zachováva tzv. symplektickú štruktúru fázového priestoru, čiže platí, že zachováva energiu a objem v priestore fázových bodov. Pre počiatočný stav $(\boldsymbol{q}(\boldsymbol{t}), \boldsymbol{p}(\boldsymbol{t}))$ v čase $\boldsymbol{t}$ s krokom $\boldsymbol{\epsilon}$ a časovou diskretizáciou $\mathbf{0} \leq \boldsymbol{t} < \llcorner \boldsymbol{T}/\boldsymbol{\epsilon} \lrcorner$, kde $\llcorner \boldsymbol{T}/\boldsymbol{\epsilon} \lrcorner = \boldsymbol{L}$ predstavuje počet leapfrog krokov aproximujeme Hamiltonove rovnice nasledujúcou schémou [2], [28], [39]:

$$\begin{aligned}
\boldsymbol{p}\left(\boldsymbol{t} + \frac{\boldsymbol{\epsilon}}{2}\right) &= \boldsymbol{p}(\boldsymbol{t}) - \frac{\boldsymbol{\epsilon}}{2}\frac{\partial V}{\partial \boldsymbol{q}}(\hat{\boldsymbol{q}}(\boldsymbol{t})), \\
\boldsymbol{q}(\boldsymbol{t} + \boldsymbol{\epsilon}) &= \boldsymbol{q}(\boldsymbol{t}) + \boldsymbol{\epsilon} M^{-1}\boldsymbol{p}\left(\boldsymbol{t} + \frac{\boldsymbol{\epsilon}}{2}\right), \\
\boldsymbol{p}(\boldsymbol{t} + \boldsymbol{\epsilon}) &= \boldsymbol{p}\left(\boldsymbol{t} + \frac{\boldsymbol{\epsilon}}{2}\right) - \frac{\boldsymbol{\epsilon}}{2}\frac{\partial V}{\partial \boldsymbol{q}}(\hat{\boldsymbol{q}}(\boldsymbol{t} + \boldsymbol{\epsilon}))\,.
\end{aligned}$$

Vzhľadom k tomu, že sa jedná o aproximáciu riešenia sústavy diferenciálnych rovníc a nie o presné analytické riešenie, tak potrebujeme toto aproximované riešenie chápať ako návrh možnej trajektórie, a teda v širšom obraze aj návrh Markovovej tranzície, ktorý môžeme schváliť alebo aj zamietnuť pomocou akceptačného kritéria.

5.3.7 Metropolisovo-Hastingsovo akceptačné kritérium

V MCMC metódach sa pri posudzovaní kandidáta najčastejšie používa Metropolisovo-Hastingsovo (skr. MH) akceptačné kritérium. Na odvodenie MH akceptačného kritéria si zavedieme podľa [2] označenie $\mathbb{Q}(\boldsymbol{q}', \boldsymbol{p}' \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0)$ ako funkciu hustoty predstavujúcu pravdepodobnosť návrhu nového stavu bodov $(\boldsymbol{q}', \boldsymbol{p}')$ pri danom $(\boldsymbol{q}_0, \boldsymbol{p}_0)$. Pre poslednú $\boldsymbol{L}$-tú iteráciu numerickej trajektórie, ktorá začína integrovaním počiatočného bodu $(\boldsymbol{q}_0, \boldsymbol{p}_0)$ dopredu, získavame kandidáta $(\boldsymbol{q}_L, \boldsymbol{p}_L)$, pričom je zrejmé, že pravdepodobnosť nadobudnutia stavu $(\boldsymbol{q}_L, \boldsymbol{p}_L)$ vzhľadom na počiatočný stav $(\boldsymbol{q}_0, \boldsymbol{p}_0)$ je rovný 1. Problém však nastáva pri kvantifikovaní pravdepodobnosti nadobudnutia stavu $(\boldsymbol{q}_0, \boldsymbol{p}_0)$, ak sme vychádzali z bodu $(\boldsymbol{q}_L, \boldsymbol{p}_L)$, ktorá je zrejme rovná nule. Ak by sme teda dali do pomeru distribúcie návrhov, tak by sme dosiahli neželaný nulový výsledok vzhľadom k tomu, že môžeme navrhovať len stavy, ktoré smerujú dopredu

$$\frac{\mathbb{Q}(\boldsymbol{q}_0, \boldsymbol{p}_0 \,|\, \boldsymbol{q}_L, \boldsymbol{p}_L)}{\mathbb{Q}(\boldsymbol{q}_L, \boldsymbol{p}_L \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0)} = \frac{0}{1} = 0.$$

Tento problém spočíva v nereverzibilite Hamiltonovej dynamiky, ktorú treba „upraviť“ tak, aby bola reverzibilná. Zároveň, táto reverzibilita sa vyžaduje aj v MH akceptačnom kritériu, ktoré musí byť symetrické, aby akceptácia nového stavu bola rovnaká bez ohľadu na to, či sa hýbeme dopredu alebo dozadu.

Na to, aby sme zabezpečili symetriu a teda numerickú reverzibilitu Hamiltonovej dynamiky, potrebujeme upraviť túto schému pomocou „negácie“ parametra hybnosti na konci trajektórie, teda pre súradnice $(\boldsymbol{q}, \boldsymbol{p})$ získame po $\boldsymbol{L}$ leapfrog krokoch návrh $(\boldsymbol{q}, -\boldsymbol{p})$. Táto úprava nespôsobí žiadne zmeny vo fyzikálnom správaní systému, teda nezmení sa objem a ani energia, jediná zmena nastane v smere trajektórie, ktorá je opačná, nakoľko zmena znamienka nastane pri parametre hybnosti.

V prípade miernej úpravy schémy pomocou negácie vieme vyjadriť nenulovú pravdepodobnosť nadobudnutia stavu návrhu $(\boldsymbol{q}_L, -\boldsymbol{p}_L)$, ak vychádzame zo stavu $(\boldsymbol{q}_0, \boldsymbol{p}_0)$, teda

$$\mathbb{Q}\left(\boldsymbol{q}_L, -\boldsymbol{p}_L \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0\right) = 1,$$

a rovnako aj v prípade nadobudnutia stavu $(\boldsymbol{q}_0, \boldsymbol{p}_0)$ pri počiatočnom $(\boldsymbol{q}_L, -\boldsymbol{p}_L)$, pre ktoré platí:

$$\mathbb{Q}\left(\boldsymbol{q}_0, \boldsymbol{p}_0 \,|\, \boldsymbol{q}_L, -\boldsymbol{p}_L\right) = 1.$$

V takomto prípade získavame návrh, pre ktorý platí zjednodušený tvar MH akceptačného kritéria

$$\alpha\left(\boldsymbol{q}_L, -\boldsymbol{p}_L \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0\right) = \min\left(1; \frac{\mathbb{Q}\left(\boldsymbol{q}_0, \boldsymbol{p}_0 \,|\, \boldsymbol{q}_L, -\boldsymbol{p}_L\right) \pi\left(\boldsymbol{q}_L, -\boldsymbol{p}_L\right)}{\mathbb{Q}\left(\boldsymbol{q}_L, -\boldsymbol{p}_L \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0\right) \pi\left(\boldsymbol{q}_0, \boldsymbol{p}_0\right)}\right),$$

čo po zjednodušení dáva

$$\alpha\left(\boldsymbol{q}_L, -\boldsymbol{p}_L \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0\right) = \min\left(1; \frac{\pi\left(\boldsymbol{q}_L, -\boldsymbol{p}_L\right)}{\pi\left(\boldsymbol{q}_0, \boldsymbol{p}_0\right)}\right). \tag{21}$$

Po dosadení kanonického tvaru vyjadreného pomocou Hamiltonovej funkcie (19) vieme MH akceptačné kritérium ďalej zjednodušiť na tvar

$$\begin{aligned}\alpha\left(\boldsymbol{q}_L, -\boldsymbol{p}_L \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0\right) &= \min\left(1; \frac{\mathrm{e}^{-H(\boldsymbol{q}_L, -\boldsymbol{p}_L)}}{\mathrm{e}^{-H(\boldsymbol{q}_0, \boldsymbol{p}_0)}}\right), \\ &= \min\left(1; \mathrm{e}^{-H(\boldsymbol{q}_L, -\boldsymbol{p}_L) + H(\boldsymbol{q}_0, \boldsymbol{p}_0)}\right),\end{aligned}$$

čím konečne dostávame tvar MH akceptačného kritéria, ktoré vyjadruje pravdepodobnosť prijatia nového stavu $(\boldsymbol{q}_L, \boldsymbol{p}_L)$. Nový stav akceptujeme v prípade, keď akceptačné kritérium nadobúda hodnotu s pravdepodobnosťou rovnou 1.

V prípade, keď je kinetická energia v tvare symetrického rozdelenia, podobne ako v našom prípade viacrozmerného normálneho rozdelenia, môžeme krok s negáciou

parametra hybnosti vynechať. Pre symetrický tvar kinetickej energie platí $K(\boldsymbol{p}) = K(-\boldsymbol{p})$, čo po dosadení do MH akceptačného kritéria dáva

$$\begin{aligned}\alpha(\boldsymbol{q}_L, -\boldsymbol{p}_L \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0) &= \min\left(1; \mathrm{e}^{-H(\boldsymbol{q}_L, -\boldsymbol{p}_L)+H(\boldsymbol{q}_0, \boldsymbol{p}_0)}\right) \\ &= \min\left(1; \mathrm{e}^{-H(\boldsymbol{q}_L, \boldsymbol{p}_L)+H(\boldsymbol{q}_0, \boldsymbol{p}_0)}\right) \\ &= \alpha(\boldsymbol{q}_L, \boldsymbol{p}_L \,|\, \boldsymbol{q}_0, \boldsymbol{p}_0)\,.\end{aligned}$$

5.3.8 Algoritmus HMC metódy

V predošlých častiach, počnúc 5.3, sme si predstavili ideu a základ princípy metódy HMC. Poslednou chýbajúcou časťou skladačky je predstavenie postupu krokov, ktorým vieme získať vzorky z aposteriórneho rozdelenia využitím HMC metódy [2], [28].

1. Pre vektor parametrov $\boldsymbol{q} = (\boldsymbol{\theta}_k, h_1, \ldots, h_n)$ vygenerujeme novú hodnotu parametra hybnosti $\boldsymbol{p}$ z viacrozmerného normálneho rozdelenia $\boldsymbol{p} \sim \mathcal{N}_{2(k+n)}(\boldsymbol{0}, M)$, ktorá je nezávislá od predošlých hodnôt parametra hybnosti a zároveň je nezávislá aj od vektora parametrov.

2. Trajektórie v „zdvihnutom“ priestore aproximujeme numerickým riešením Hamiltonových rovníc využitím leapfrog metódy s počtom krokom $\boldsymbol{L}$ a dĺžkou kroku $\boldsymbol{\epsilon}$.

3. Riešenie získané numerickou aproximáciou trajektórie posudzujeme ako kandidáta v MH akceptačnom kritériu, pričom nový stav $(\boldsymbol{q}_L, \boldsymbol{p}_L)$ buď akceptujeme, alebo zamietneme. Ak navrhovaný stav zamietneme, tak nasledujúci stav je rovnaký ako predošlý.

4. Následne opakujeme kroky 1 až 3, pričom v prvom kroku opäť začíname generovaním novej, nezávislej, hodnoty parametra hybnosti.

Na fungovanie algoritmu sa vyžadujú vstupné parametre, medzi ktoré patrí dĺžka iterácií $\boldsymbol{L}$, veľkosť kroku $\boldsymbol{\epsilon}$ a kovariančná matica M. Pre veľkosť kroku $\boldsymbol{\epsilon}$ platí, že ak je jej hodnota príliš veľká, tak aproximácia trajektórie bude nepresná, čo môže viesť k odkloneniu sa od cieľovej trajektórie a v konečnom dôsledku aj k neakceptovaniu návrhov nových stavov, a teda nízkej efektivite algoritmu. Pre nízku hodnotu naopak

máme veľmi dlhé simulácie, a teda aj dlhý výpočtový čas. Pre počet iterácií $\boldsymbol{L}$ sa pri voľbe malej hodnoty stretávame s problémom nedostatočného „priestoru“ na aproximáciu cieľovej trajektórie. V takomto prípade nedochádza k efektívnemu posunu v rámci cieľového rozdelenia, čím sa skúmajú iba oblasti blízke aktuálneho bodu $(\boldsymbol{q}, \boldsymbol{p})$. Pri veľkej hodnote na druhe strane dochádzka ku zbytočnému výpočtovému preťaženiu a časovej neefektivite. Problém, ktorý sa vyskytuje pri veľkom počte iterácií rieši tzv. vylepšená HMC metóda, ktorá sa nazýva No-U-Turn sampler (skr. NUTS), ktorá slúži ako zastavovacie kritérium vylepšeného HMC algoritmu v prípade, kedy sa trajektória otočí a začne sa pohybovať späť smerom k počiatočnému $\boldsymbol{q}_0$.

5.3.9 No-U-Turn sampler

Dynamickou modifikáciou HMC metódy je jej multiplikatívna NUTS verzia, ktorá rovnako, ako HMC metóda, aproximuje trajektórie pomocou numerickej leapfrog metódy, pričom dynamicky a adaptívne určuje dĺžku trajektórie, a teda počet krokov v každej iterácii tak, aby dosiahla lepšiu výkonnosť v porovnaní s HMC metódou. NUTS teda v priebehu výpočtu automaticky prispôsobuje svoj proces generovania návrhov na základe informácií, ktoré získava počas výpočtu. Algoritmus HMC metódy opísaný v časti 5.3.8 je v kroku 2 zmenený v prípade NUTS tak, že aproximované trajektórie sa hýbu oboma smermi, dopredu aj dozadu, násobne sa zvyšujúcimi krokmi, čím vytvárajú akýsi vyvážený binárny strom [2], [28]. V každej iterácii sa numerická aproximácia trajektórie postupne posunie o 2^z leapfrog krokov dopredu a aj dozadu, teda dva kroky dopredu aj dozadu, štyri, osem a tak ďalej. Týmto spôsobom sa každou jednou násobnou operáciou 2^z zvyšuje hĺbka binárneho stromu a zdvojnásobujú sa počet leapfrog krokov v danej hĺbke stromu oproti tej predchádzajúcej. Efektivita NUTS spočíva v kritériu zastavenia, ktoré ďalšiu iteráciu zastaví

i) ak v priebehu iterácií dôjde k otočeniu trajektórie, teda trajektória sa začne vracať naspäť k počiatočnému bodu, čo je želaný efekt,

ii) ak strom dosiahne maximálnu hĺbku, resp. maximálny počet iterácií – v takomto prípade je potrebné modelovanie opakovať, nakoľko cieľové rozdelenie nebolo ešte dostatočne preskúmané.

Zastavovacie kritérium spočíva na princípe zastavenia numerickej aproximácie v prípade, kedy už viac nedochádza v čase $\boldsymbol{t}$ ku zvyšovaniu vzdialenosti medzi počiatočným bodom $\boldsymbol{q}_0$ a návrhom (v aktuálnej iterácii) $\boldsymbol{q}'$. V prípade multiplikatívneho NUTS algoritmu, ak si označíme $\boldsymbol{q}_-$, $\boldsymbol{p}_-$ ako záporný krajný bod binárneho stromu získaný posunom smerom dozadu a rovnako kladný krajný bod $\boldsymbol{q}_+$, $\boldsymbol{p}_+$ získaný posunom dopredu, tak vieme odvodiť zastavovacie, resp. kritérium otočenia smeru trajektórie [2]

$$(\boldsymbol{q}_+ - \boldsymbol{q}_-) \cdot \boldsymbol{p}_+ < 0 \tag{22}$$

a súčasne

$$(\boldsymbol{q}_- - \boldsymbol{q}_+) \cdot \boldsymbol{p}_- < 0.$$

Záporný súčin vo vzťahu (22) znázorňuje situáciu, kedy je hybnosť smerujúca dopredu $\boldsymbol{p}_+$ a vzdialenosť $(\boldsymbol{q}_+ - \boldsymbol{q}_-)$ záporne orientovaná, čo znamená, že trajektória sa otočila a začína sa vracať späť smerom k bodu $\boldsymbol{q}_-$, čo platí analogicky aj pre opačný smer.

Konkrétna adaptácia jednotlivých metód, akými sú HMC či NUTS závisí od vývojárov balíkov či knižníc. V prípade knižnice Stan [27], je pre NUTS algoritmus po novom MH akceptačné kritérium nahradené multinomickým rozdelením s ohľadom na druhú polovicu krokov v trajektórií, pričom viac informácií je možné nájsť v publikácií [3]. Stan vylaďuje parametre NUTS algoritmu automaticky, počas warm up fázy vzorkovania. Stan volí takú hodnotu leapfrog kroku $\boldsymbol{\epsilon}$, aby bola dosiahnutá cieľová hodnota akceptačného kritéria, ktoré je nastavené na hodnotu $\alpha = 0{,}8$. V prípade problémov s konvergenciou sa odporúča zvyšovať spodnú hranicu akceptačného kritéria α, čo má za dôsledok kratší krok, a teda presnejší výsledok aproximovanej trajektórie [28]. Kovariančná matica M je odhadnutá zo vzoriek získaných počas warm up iterácií a nakoniec, počet krokov $\boldsymbol{L}$ je dynamicky nastavený.

6 Analýza odhadov parametrov z aposteriórneho rozdelenia

Kľúčovou súčasťou pre získanie robustných odhadov je oddelenie vzoriek Markovovho reťazca, ktoré vznikli ešte predtým, ako Markovov reťazec dosiahol cieľove rozdelenie od vzoriek, ktoré vznikli potom, ako Markovov reťazce dosiahol cieľové rozdelenie a začal ním prechádzať [2]. Fáza pred dosiahnutím cieľového reťazca sa nazýva **burn in** alebo aj **warm up**, pričom odhady získané zo vzoriek v tejto fáze sú veľmi vychýlené, teda rozdiel v absolútnych číslach medzi získaným odhadom a skutočným odhadom je veľký. Po dosiahnutí cieľového rozdelenia sa rozdiel medzi skutočným očakávaním hodnoty parametra a odhadnutým výrazne znižuje. Pre dostatočne veľký počet transakcií sa odhady zdokonaľujú a rozdiel je už iba minimálny, asymptoticky smerujúci k nule.

MCMC metódy však predstavujú aj dve potenciálne riziká spojené so simulovaním z aposteriórneho rozdelenia. Dôležitou súčasťou simulovania vzoriek pomocou MCMC techník je schopnosť modelára určiť

i) kedy Markovov reťazec dosiahol cieľové rozdelenie, teda konvergencia do rovnovážneho stavu (*equilibrium*),

ii) či vzorky získané z MCMC nie sú korelované, pretože v takomto prípade sa na takýto Markovov reťazec neplatí centrálna limitná veta.

6.1 Konvergencia Markovovho reťazca k cieľovému rozdeleniu

V praxi sa často využíva simulácia viacerých Markovových reťazcov s cieľom dosiahnutia presnejších a dôveryhodnejších výsledkov, najmä v prípadoch, kde dáta majú zložitú štruktúru alebo veľkú variabilitu. Simulácia viacerých reťazcov umožňuje robustnejšiu analýzu a lepšiu identifikáciu parametrov modelu na základe distribúcie výsledkov, pričom nám poskytuje možnosť porovnania jednotlivých reťazcov či už sa to týka odhadov parametrov alebo samotnej konvergencie. Predstavme si situáciu, kedy v prípade použitia (bez ujmy na všeobecnosti) dvoch Markovových reťazcov $M = 2$ s náhodnými počiatočnými stavmi pre každý Markovov reťazec $\boldsymbol{\theta}_1^1$ a $\boldsymbol{\theta}_2^1$ „dokonvergujú“ po N tranzíciách tieto reťazce do štatisticky významne odlišných rozdelení pravdepo-

dobnosti, čím sa odlišnými stanú aj ich odhady parametrov. Ak by sme pracovali len s jedným reťazcom, tak by sme sa mohli vystaviť riziku vychýlených odhadov v podobe napríklad použitia malého počtu tranzícií, kedy sa reťazcu nepodarilo preskúmať celé cieľové rozdelenie, poprípade nesprávnych očakávaní od apriórnych rozdelení parametrov.

Pre M Markovových reťazcov s dĺžkou N označme odhad vektora parametrov m-tého Markovovho reťazca v n-tom kroku ako $\boldsymbol{\theta}_m^n$. Pri takto zadefinovanom počte a dĺžke jednotlivých reťazcov získavame vzorku s veľkosťou $M \times N$. S myšlienkou vzájomného porovnávania viacerých reťazcov, ktoré boli náhodne inicializované prišli Andrew Gelman a Donald B. Rubin [30], ktorí vytvorili štatistiku potenciálnej škálovacej redukcie $\hat{\mathrm{R}}$, ktorú pre jednoduchosť označujeme a nazývame ako $\hat{\mathrm{R}}$ štatistika. Zadefinujme si označenia $\bar{\theta}_m^{(\bullet)}$ pre priemernú hodnotu parametra v m-tom reťazci a $\bar{\theta}_\bullet^{(\bullet)}$ ako priemernú hodnotu parametra všetkých reťazcov. Princíp $\hat{\mathrm{R}}$ štatistiky spočíva vo výpočte variancie v rámci každého z M reťazcov [30]:

$$W = \frac{1}{M} \sum_{m=1}^{M} s_m^2, \tag{23}$$

kde

$$s_m^2 = \frac{1}{N-1} \sum_{n=1}^{N} \left(\theta_m^{(n)} - \bar{\theta}_m^{(\bullet)} \right)^2,$$

a variancie medzi jednotlivými reťazcami [30]:

$$B = \frac{N}{M-1} \sum_{m=1}^{M} \left(\bar{\theta}_m^{(\bullet)} - \bar{\theta}_\bullet^{(\bullet)} \right)^2, \tag{24}$$

kde

$$\bar{\theta}_m^{(\bullet)} = \frac{1}{N} \sum_{n=1}^{N} \theta_m^{(n)}$$

a

$$\bar{\theta}_\bullet^{(\bullet)} = \frac{1}{M} \sum_{m=1}^{M} \bar{\theta}_m^{(\bullet)}.$$

Ďalej získame odhad variancie $M \times N$ vzorky váženým súčtom (23) a (24)

$$\hat{\mathrm{var}}(\theta \mid y) = \frac{N-1}{N} W + \frac{1}{N} B, \tag{25}$$

pričom $\hat{\mathrm{R}}$ štatistiku získame prostredníctvom nasledujúcej transformácie [30]:

$$\hat{\mathrm{R}} = \sqrt{\frac{\hat{\mathrm{var}}(\theta \mid y)}{W}}. \tag{26}$$

Intuícia stojaca za týmto kritériom spočíva v overení konvergencie jednotlivých reťazcov tak, že ak dosiahli všetky reťazce rovnovážny stav, tak rozdiely medzi jednotlivými reťazcami budú blízke nule, čo vyplýva z faktu, že variancia v rámci jednotlivých reťazcov a variancia medzi reťazcami je pre $N \to \infty$ rovnaká. V takom prípade má $\hat{\mathrm{R}}$ štatistika hodnotu 1, čo vyplýva pre $B = W$ z (26):

$$\hat{\mathrm{R}} = \sqrt{\frac{\hat{\mathrm{var}}(\theta \mid y)}{W}} = \sqrt{\frac{\frac{N-1}{N}W + \frac{1}{N}B}{W}} = \sqrt{\frac{\frac{N-1}{N}W + \frac{1}{N}W}{W}} = 1.$$

Vzhľadom ku výpočtovým obmedzeniam transakcií N je podľa Gelmana a Rubina vhodným modelom taký model, ktorý pre náhodne inicializované reťazce spĺňa hraničné kritérium $\hat{\mathrm{R}} < 1{,}1$. Na druhej strane je dôležite poznamenať dva dôležité fakty. Medzi fázou, kedy Markovov reťazec dosahuje cieľové rozdelenia a fázou, kedy ním prechádza, a dokonca v ideálnom prípade, kedy ho dôkladne preskúma, nie je presne stanovená hranica. Odstránením *burn in* fázy zo vzorkovania nevieme presne zabezpečiť, že na naše odhady nevplývajú ešte odhady z počiatočnej fázy. Druhým faktom spočíva v silných predpokladoch $\hat{\mathrm{R}}$ štatistiky, medzi ktoré patrí predpoklad, predpoklad o cieľovom rozdelení pravdepodobnosti parametrov, ktoré by malo byť Gaussovo alebo mať najviac dva parametre. V skutočnosti je odhadované rozdelenie oveľa komplexnejšie, a preto je v prípade vysokých hodnôt $\hat{\mathrm{R}}$ štatistiky potrebná ďalšia analýza.

6.2 Efektívna veľkosť vzorky

Ako sme už spomínali, veľkým rizikom pri MCMC metódach je autokorelácia, či antikorelácia vzoriek získaných simulovaním z aposteriórneho rozdelenia pravdepodobnosti. Pre takto získané vzorky však neplatí centrálna limitná veta, čo v konečnom dôsledku nevedie ku získaniu nevychýlených odhadov jednotlivých parametrov. Dôležitou súčasťou modelovania sa preto stáva analýza veľkosti efektívnej vzorky (skr. ESS z angl. *effective sample size*), ktorá predstavuje počet nezávislých vzoriek pre daný parameter [30]. Čím je hodnota ESS pre odhady parametrov nižšia, teda počet nezávislých vzoriek je znížený, tak tým je zvýšená neistota v odhadoch parametrov. V ideálnom

prípade, keď nedochádza k autokorelácii medzi vzorkami je hodnota ESS na úrovni dĺžky reťazca, teda ESS $= N$. Pri autokorelácii medzi vzorkami sa efektivita získaných vzoriek znižuje a platí ESS $\in \langle 0, N)$, pričom v prípade rýchlejšej konvergencie v cieľovom rozdelení môže ESS presiahnuť hodnotu N. Autori balíčka `rstan` [2], [27] odporúčajú, aby veľkosť efektívnej vzorky bola nad hranicou 1 % z počtu post-warm up iterácií pre zabezpečenie stabilných a nevychýlených odhadov [31]. V nasledujúcej časti načrtneme teóriu kritéria výpočtu efektívnej vzorky podľa [30].

Autokorelácia pre $t \in \mathbb{N}_0$ posunov (angl. *lag*) v rámci reťazca s aposteriórnou funkciou hustoty $f(\theta)$ so strednou hodnotou μ a varianciou σ^2 je daná vzťahom:

$$\rho_t = \frac{1}{\sigma^2} \int_\Theta \left(\theta^{(n)} - \mu\right) \left(\theta^{(n+t)} - \mu\right) f(\theta) \, \mathrm{d}\theta,$$

pričom ρ_t predstavuje hodnotu autokorelácie medzi n-tou a $(n+t)$-tou iteráciou v rámci Markovovho reťazca, postupne, cez n. Veľkosť efektívnej vzorky potom získame nasledovným spôsobom:

$$N_{\text{eff}} = \frac{N}{1 + 2\sum_{t=1}^{\infty} \rho_t},$$

z ktorého intuitívne vyplýva, že veľkosť efektívnej vzorky sa znižuje so zvyšujúcou sa koreláciou cez všetky posuny $t \in \mathbb{N}_0$.

Vzhľadom k zložitému tvaru aposteriórnej hustoty však nie sme schopní pre modely stochastickej volatility vypočítať autokoreláciu parametrov z aposteriórnej hustoty a tým pádom ani hodnotu efektívnej veľkosti vzorky. Na odhad ESS je preto potrebné použiť teda aj odhad autokorelácie. V praktickej časti práce pracujeme s balíkom `rstan` [27] zabudovanom v softvéri R [24], ktorý na odhad autokorelácií využíva rýchlu Fourierovu transformáciu (z angl. *Fast Fourier Transform*) zabudovanú v balíčku `fft`, ktorá transformuje časové dáta do frekvenčného priestoru, ktorý umožňuje identifikovať periodické vzory a cyklické závislosti v časových radových dátach, pričom viac informácií je možné nájsť v [30]. Takto získané odhady autokorelácie $\hat{\rho}_{t,m}$ posunu t z jednotlivých reťazcov $m \in \{1, \ldots, M\}$ sú ďalej podľa [30] kombinované s varianciou v rámci každého z reťazcov (23) a odhadom variancie medzi jednotlivými reťazcami (25) vo vzťahu, formálne ho označíme ako $\hat{\text{var}}$, nasledovným spôsobom:

$$\hat{\rho}_t = 1 - \frac{W - \frac{1}{M}\sum_{m=1}^{M} s_m^2 \hat{\rho}_{t,m}}{\hat{\text{var}}}.$$

V prípade, keď reťazce nekonvergujú do cieľového rozdelenia, je odhad variancie medzi jednotlivými reťazcami vâr veľký, čo spôsobí vysokú hodnotu odhadu autokorelácie a v konečnom dôsledku nižšie skóre pre ESS.

Odhad pre ESS už vieme vypočítať z odhadu $\hat{\rho}_t$, avšak treba si dať pozor na dve skutočnosti. Medzi autokoreláciami sa môžu vyskytovať aj také, ktoré sú záporné, ktoré by mohli byť v súčte $\sum_{t=1}^{\infty} \hat{\rho}_t$ nesprávne interpretované, čím by výsledný odhad ESS štatistiky nebol správny. Výskyt zápornej autokorelácie je iba pri nepárnych posunoch, pričom podľa Charles J. Geyera a [30] je užitočnou pomôckou na vytvorenie kladných autokorelácií ich párovanie začínajúc postupne od posunu $t' = 0$. Druhou súčasťou počítania odhadu ESS je určenie hornej hranice pre súčet autokorelácií. Pre rastúci počet posunov sa zvyšuje aj šum v odhadoch autokorelácie, preto sa pre ďalšie výpočty zvykne používať sprava useknutá suma hodnôt $\hat{\rho}_t$. Odhad ESS potom vypočítame ako [30]:

$$\hat{N}_{\text{eff}} = \frac{M \times N}{\hat{\tau}},$$

kde

$$\hat{\tau} = 1 + 2 \sum_{t=1}^{2m+1} \hat{\rho}_t = -1 + 2 \sum_{t'=0}^{m} \hat{P}_{t'} = -1 + 2 \sum_{t'=0}^{m} \hat{\rho}_{2t'} + \hat{\rho}_{2t'+1},$$

pričom $\hat{P}_{t'} = \hat{\rho}_{2t'} + \hat{\rho}_{2t'+1}$ a horná hranica sumy je zvolená tak, aby m bolo najväčšie číslo, pre ktoré platí, že páry autokorelácií sú kladné $\hat{P}_{t'} > 0$ pre všetky $t' = 1, \ldots, m$. Postupnosť $\hat{P}_{t'}$ je navyše zoraďovaná tak, aby bola monotónne klesajúca a autokorelácia v jednotlivých iteráciách postupne klesala, pretože monotónna postupnosť autokorelácií zvyšuje presnosť odhadu ESS a poskytuje lepšiu predstavu o tom, koľko nezávislých vzoriek je dostupných v danom stave, v ktorom sa systém nachádza.

V prípadoch, kedy je autokorelácia medzi vzorkami vysoká, stojí za zváženie použiť techniku zrieďovania (angl. *thinning*), ktorá na výpočet odhadov ponechá len každú k-tu vzorku. Predstavme si dve modelové situácie, v ktorých generujeme po 1 000 vzoriek, pričom prvá situácia predstavuje techniku modelovania $N = 1\,000$ vzoriek so zrieďovaním $k = 1$, teda každá vzorka získaná v MCMC je uložená a zahrnutá vo výpočte odhadov parametrov. Druhý prípad predstavuje situáciu, kedy generujeme 10 000 vzoriek s koeficientom zrieďovania $k = 10$, čím dostávame výslednú vzorku s počtom 1 000. Oba modelové príklady poskytujú rovnaký počet výsledných vzoriek,

avšak druhá situácia má potenciál poskytnúť vyššiu hodnotu ESS vzhľadom k faktu, že vzorky získané zrieďovaním majú autokoreláciu na úrovni ρ_t pre posun t, ktorá je ekvivalentná autokorelácii ρ_{10t} v reťazci bez zrieďovania [30]. Zrieďovanie distribúciu parametrov nemení, iba zachováva každú k-tu vzorku, no na druhej strane je to spôsob „hádzania vzoriek do koša“, čo predstavuje istú stratu informácie.

6.3 Odhad štandardnej chyby pre MCMC

Ďalším dôležitým ukazovateľom pre aposteriórnu analýzu výsledkov je štandardná chyba Markov Chain Monte Carlo metódy (skr. MCMC-SE), ktorou meriame ako (ne)presné odhady parametrov sme získali v modelovacej fáze. Čím nižšiu chybovosť MCMC-SE dosiahneme, tým presnejšie a spoľahlivejšie odhady získame. MCMC-SE je pre parameter θ_i z k - prvkového vektora parametrov $\boldsymbol{\theta}$ pre reťazce $1, \ldots, m$ definovaný ako

$$\text{MCMC-SE}(\theta_i) = \frac{\hat{\sigma}_i}{\sqrt{\hat{N}_{\text{eff}}}},$$

kde na získanie odhadov pre strednú hodnotu $\bar{\theta}_i = \sum_{l=1}^{m} \theta_i^{(l)}$ a štandardnú odchýlku $\hat{\sigma}_i = \sqrt{\frac{1}{m-1} \sum_{l=1}^{m} (\theta_i^{(l)} - \bar{\theta}_i)^2}$ použijeme vzorky získané zo simulovania z aposteriórneho rozdelenia. MCMC-SE závisí od veľkosti efektívnej vzorky, teda je čím vyššia hodnota ESS a tým pádom čím viac kvalitnejšie a nezávislé sú naše vzorky, tým dosahujú odhady parametrov MCMC-SE nižšiu chybovosť.

7 Porovnanie a testovanie modelov stochastickej volatility

Kľúčovou súčasťou tvorby jednotlivých modelov, či analýzou ich výsledkov je ich testovanie a vzájomné porovnanie, no najmä určenie modelu, ktorý dokáže správne a zároveň aj najlepšie popísať dynamiku a štruktúru dát. Pri testovaní sa zameriavame na splnenie predpokladov rozdelení a prvých dvoch momentov chybových členov ϵ_t a η_t a tým pádom na správnosť modelu. Z modelov A, B, C a D budeme vyberať taký, ktorý najlepšie dokáže zachytiť správanie dát, a zároveň je „modelársky šetrný“ vzhľadom na počet parametrov.

7.1 Testovanie

Nevyhnutnou súčasťou techník matematického modelára je testovanie výsledkov a predpokladov, ktoré jednotlivé modely obsahujú. V našej práci sme testovali splnenie predpokladov šumu ϵ_t, ktorý by sa vzhľadom na modely A, B, C a D mal riadiť podľa normálneho, zošikmeného normálneho, Studentovho t a GED rozdelenia, v uvedenom poradí. Pri testovaní využívame predpoklad normality šumu η_t, ktorý vygenerujeme z normálneho rozdelenia s parametrami $\tilde{\eta}_t \sim \mathcal{N}(0; \sigma_\eta^2)$. Takisto používame odhady parametrov z jednotlivých modelov $\left(\bar{\mu}, \bar{\phi}, \bar{\sigma}_\eta^2\right)$. Pre počiatočnú, náhodne vygenerovanú hodnotu procesu volatility $h_1 \sim \mathcal{N}(-10; 5)$ získame odhad procesu volatility h_t využitím cyklu s pevným počtom opakovaním daným vzťahom:

$$\tilde{h}_t = \bar{\mu} + \bar{\phi}\left(\tilde{h}_{t-1} - \bar{\mu}\right) + \tilde{\eta}_t,$$

kde $t = 2, \ldots, N$, pričom odhad šumu ϵ_t získame transformáciou

$$\tilde{\epsilon}_t = \frac{y_t}{\exp\left(\frac{\tilde{h}_t}{2}\right)}.$$

Takto získanú premennú $\tilde{\epsilon}_t$ sme následne testovali využitím nasledujúcich testov:

i) **testovanie voči normálnemu rozdeleniu** - distribúciu šumu $\tilde{\epsilon}_t$ sme testovali voči normálnu rozdeleniu pomocou A-D testu (skr. z *Andersonov-Darlingov test*) pričom parametre $\tilde{\mu}$ a $\tilde{\sigma}$ sme odhadli z dát $\tilde{\epsilon}_t$; v prípade splnenia podmienky

normality rozdelenia sme ďalej testovali hypotézy o nulovej strednej hodnote, pričom sme využili t-test a na testovanie hypotézy o jednotkovej variancii sme použili χ^2-test,

Testy o normalite rozdelenia šumu $\tilde{\epsilon}_t$ využili pre všetky modely A, B, C a D. Pre modely B, C a D sme šum testovali voči danému predpokladanému rozdeleniu (označíme ako *voči špecifickému rozdeleniu*), teda voči zošikmenému normálnemu, Studentovmu t, či GED rozdeleniu, v uvedenom poradí:

i) **testovanie voči špecifickému rozdeleniu** - na testovanie voči špecifickému rozdeleniu sme rovnako, ako v predošlom prípade, využili A-D test, pričom parametre sme získali odhadom z dát pomocou funkcie `fitdist` zabudovanej v balíčku `fitdistrplus` [9]; na testovanie nulovej strednej hodnoty sme využili Wilcoxonov test a na testovanie jednotkovej variancie sme použili Leveneov dvojsúborový test o variancii, tieto testy sú robustné aj voči dátam, ktoré sa neriadia podľa normálneho rozdelenia.

Vzhľadom k faktu, že testovanie jednotkovej variancie pre dáta, ktoré sa neriadia podľa normálneho rozdelenia si vyžaduje aspoň dva súbory, tak na testovanie pomocou Leveneovho testu sme druhý súbor vygenerovali zo štandardizovaného normálneho rozdelenia. Generovanie druhého súboru a následne testovanie sme kvôli náhodnosti vykonali jedenásťkrát, pričom ako výslednú hodnotu Leveneovho testu sme uvažovali ich mediánovú hodnotu.

Vzhľadom pre istú náhodnosť, ktoré testovanie s generovanými náhodnými premennými prináša, sme pre získanie robustných výsledkov vykonali vyššie opísané testovanie cyklom s pevným počtom opakovaní, rovným 5 000-krát. Sumárne štatistiky pozostávali z percentuálnej „úspešnosti“ testov, ktorých p-hodnota prekročila hladinu významnosti $\alpha = 0{,}05$, teda nás zaujímal podiel „úspešných“ testov voči 5 000 pokusom. Druhým ukazovateľom bola mediánová p-hodnota získaná naprieč iteráciami.

7.2 Informačné kritériá

Správnym výberom modelu sa vieme v praxi priblížiť ku presnejším budúcim hodnotám volatility, ktorej odhady sú žiadané najmä vo finančníctve pri modelovaní

cien. Pri jednoduchších modeloch môže nastať problém, kedy nedokážu presne zachytiť a popísať správanie sa dát, no na druhej strane, komplexné modely môžu so sebou prinášať problém tzv. preparametrizácie (*overfitting*), kedy sa model dobre prispôsobí dátam, na ktorých je trénovaný, avšak pri jeho použití na nových dátach či v predikciách jeho kvalita a presnosť výrazne klesá [39]. V skutočnosti neexistuje žiadne jednoznačné pravidlo, ktorým vieme s istotou povedať ktorý z modelov je pre naše dáta najvhodnejší, avšak máme nástroje, tzv. **informačné kritériá**, ktoré nás vedia k tomuto modelu nasmerovať.

7.2.1 Informačné kritérium DIC

Medzi najjednoduchšie a najčastejšie používané informačné kritérium v oblasti bayesovskej štatistiky a najmä MCMC modelov patrí informačné kritérium založené na deviancii (skr. DIC z angl. *deviance information criterion*), teda miery toho, ako dobre sa model prispôsobil dátam a jeho komplexnosti. Pre vyššiu hodnotu deviancie máme vyššiu odchýlku medzi pozorovanými hodnotami a výsledkami modelu, čo vedie k vyššej chybovosti a vyššej nekvalite modelu. Pre nízku odchýlku presnejšie a kvalitnejšie odhady modelu, ktoré sa od pozorovaných hodnôt líšia len veľmi málo [10]. DIC je definované ako súčet očakávanej (prípadne optimálnej) hodnoty mínus dvojnásobku logaritmickej funkcie vierohodnosti (*minus twice the optimal log-likelihood function*) (ozn. $\bar{D}$), ktorý popisuje ako dobre model odpovedá skutočným pozorovaniam a tzv. penalizačného člena (ozn. p_D), ktorý meria zložitosť modelu na základe počtu parametrov [10], [39]:

$$\mathrm{DIC} = \bar{D} + p_D.$$

Prvý člen vyjadrujúci mieru správnosti popísania našich dát je možné vyjadriť ako [10], [39]:

$$\bar{D} = \mathrm{E}_{\boldsymbol{\theta}\,|\,\boldsymbol{y},\boldsymbol{h}}\left[-2\ln\left(\boldsymbol{y}\,|\boldsymbol{\theta}, h_t\right)\right],$$

čo vieme odhadnúť pre jednotlivé pozorovania $t = 1, \ldots N$ a vzorky $i = 1, \ldots, I$ ako

$$\hat{\bar{D}} = -2\sum_{t=1}^{N}\frac{1}{I}\sum_{i=1}^{I}\ln f\left(y_t\,|\,h_t^{(i)}\right), \tag{27}$$

kde $f\left(y_t\,|\,h_t^{(i)}\right)$ označuje funkciu vierohodnosti pre t-te pozorovanie dát a i-tu vzorku z celkového počtu I simulácií. Čím väčšia hodnota funkcie vierohodnosti je dosiahnutá,

tým je model kvalitnejší, keďže táto hodnota rastie spolu s kvalitou modelu. Transformácia hodnoty funkcie vierohodnosti výrazom $-2 \times \ln(\cdot)$ „preklopí“ naše očakávania o čo najvyšších hodnotách do hodnôt, ktoré sú najnižšie možné, a teda od člena $\hat{\bar{D}}$ budeme požadovať najmenšiu možnú hodnotu. Pre penalizačný člen p_D máme tvar:

$$p_D = \bar{D} - D\left(\bar{\boldsymbol{\theta}}\right),$$

ktorý vyjadruje rozdiel medzi skutočnou a odhadovanou informáciou obsiahnutou v dátach. Člen $\bar{D}$ poznáme z (27) a odhad pre $D\left(\bar{\boldsymbol{\theta}}\right)$ získame ako [10], [39]:

$$\hat{D}\left(\bar{\boldsymbol{\theta}}\right) = -2\sum_{t=1}^{N} \ln f\left(y_t \,|\, \bar{h}_t\right),$$

kde $\bar{h}_t$ predstavuje aposteriórny priemer $\bar{h}_t = \sum_{i=1}^{I} h_t^{(i)}$. DIC predstavuje výpočtovo jednoduché kritérium na porovnanie modelov stochastickej volatility, avšak táto metóda predpokladá, že aposteriórne rozdelenie sa riadi podľa viacrozmerného normálneho rozdelenia, čo v mnohých prípadoch nevieme zaručiť [39] a na porovnanie modelov sa odporúča použitie aj iných kritérií.

7.2.2 Informačné kritérium WAIC

Medzi ďalšie známe informačné kritériá sa zaraďuje aj WAIC (z angl. *widely applicable information criterion*), čo v preklade znamená „široko uplatniteľné informačné kritérium“, ktoré sa často označuje po menách autorov ako Watanabeho-Akaikeho informačné kritérium (skr. WAIC). Na rozdiel od DIC, informačná kritérium WAIC nie je založené na Fisherovej konvergencii, a teda konvergencii k nejakému bodu, čím sa stáva širšie uplatniteľným kritériom v praxi [10]. Vzorec pozostáva rovnako z člena dobrej zhody a penalizačného člena merajúceho komplexnosť modelu a má tvar [10], [39]:

$$\mathrm{WAIC} = -2\sum_{t=1}^{N} \ln \mathrm{E}_{\boldsymbol{\theta}\,|\,\boldsymbol{y}}\left[f\left(y_t \,|\, \boldsymbol{\theta}, h_t\right)\right] + 2p_{\mathrm{WAIC}},$$

kde člen dobrej zhody sa dá odhadnúť ako

$$-2\sum_{t=1}^{N} \ln\left[\frac{1}{I}\sum_{i=1}^{I} f\left(y_t \,|\, \boldsymbol{\theta}^{(i)}, h_t^{(i)}\right)\right]$$

a penalizačný člen p_{WAIC} má tvar

$$p_{\mathrm{WAIC}} = \sum_{t=1}^{N} \mathrm{var}\left[\ln f\left(y_t \,|\, \boldsymbol{\theta}\right)\right].$$

Odhad penalizačného člena získame pre každé y_t, kde $t = 1, \ldots, N$ ako [10], [39]:

$$\hat{p}_{\text{WAIC}} = \sum_{t=1}^{N} \text{var}_{i=1}^{I} \ln f\left(y_t \mid \boldsymbol{\theta}^{(i)}, h_t^{(i)}\right),$$

kde $\text{var}_{i=1}^{I}$ predstavuje výberový rozptyl funkciíln $f\left(y_t \mid \boldsymbol{\theta}^{(1)}, h_t^{(1)}\right), \ldots, \ln f\left(y_t \mid \boldsymbol{\theta}^{(I)}, h_t^{(I)}\right)$ pre všetky $t = 1. \ldots, N$. WAIC kritérium, podobne ako DIC kritérium, predstavuje meria ako dobre model opisuje dané dáta, pričom zohľadňuje aj zložitosť modelu v podobe penalizačného člena p_{WAIC}. Čím vyššiu hodnotu dosahuje dosahuje penalizačný člen, tým je model komplexnejší a WAIC kritérium má neželanú vyššiu hodnotu.

WAIC predstavuje robustné kritérium voči zložitejším modelom, avšak jeho nevýhoda spočíva v tom, že jeho teoretická správnosť pre modely s latentnými premennými ešte nebola dostatočne preskúmaná [39].

8 Modelovanie stochastickej volatility

V nasledujúcej kapitole prakticky spracujeme modelovanie stochastickej volatility pre výmenné kurzy EUR/USD a HUF/EUR v štatistickom softvéri R [24] pomocou rozhrania RStan [27], ktorý ponúka praktické modelovanie pomocou bayesovského prístupu k štatistike na základe Hamiltonovej Monte Carlo algoritmu a jeho vylepšenej NUTS verzie.

8.1 Rozhranie RStan

RStan funguje ako rozhranie medzi štatistickým softvérom R a knižnicou Stan, ktorá je implementovaná v programovacom jazyku C++. RStan umožňuje vytváranie matematických modelov priamo v softvéri R, pričom samotné modelovanie je vykonávané pomocou jazyka C++. V nasledujúcej časti si predstavíme definíciu modelu A (ozn. `model_A`) v jazyku rozhrania RStan implementovaného v softvéri R [24] využitím nasledujúcich blokov

- **dátový blok** definuje vstupné dáta modelu, teda vektor log-výnosov výmenného kurzu $\boldsymbol{y}$ a jeho dĺžku T. Pre parameter dĺžky zavedieme aj jeho dolné ohraničenie, ktoré definujeme pomocou príkazu `lower`

```
data {
  int<lower = 0> T;       // počet pozorovaní
  vector[T] y;            // vektor pozorovaní
}
```

- **blok parametrov** obsahuje definíciu odhadovaných parametrov modelu a požiadavky na ich ohraničenie, takisto v tomto bloku zadefinujeme premennú pre odhadovaný proces volatility h_t ako vektor dĺžky T

```
parameters {
  real mu;                                  // hladina procesu
  real<lower = 0, upper = 1> phi_hat;       // perzistencia procesu
  real<lower = 0.01> s2;                    // variancia procesu
  vector[T] h;                              // latentná premenná
}
```

- **blok transformácie parametrov** umožňuje transformovať odhadované parametre modelu tak, ako sme to potrebovali v prípade parametra perzistencie ϕ, od ktorého pre zachovanie stacionarity procesu požadujeme, aby nadobúdal hodnoty z intervalu $(-1; 1)$. Pomocný parameter ϕ^* modelujeme pomocou Beta rozdelenia z intervalu $(0; 1)$, pričom transformáciou, $\phi = 2\phi^* - 1$ dosiahneme požadované ohraničenie. Rovnakým spôsobom vieme transformovať parameter variancie procesu volatility σ^2 na hodnotu σ, ktorú používame vo výpočtoch.

```
transformed parameters{
 real<lower = -1, upper = 1> phi;
 phi = (2*phi_hat - 1);     // transformácia parametra perzistencie
 real<lower = 0> sigma;
 sigma = sqrt(s2);          // transformácia parametra volatility
}
```

- **blok modelu** obsahuje základnú definíciu modelu, v ktorej zadefinujeme apriórne rozdelenia parametrov a rovnako aj združenú hustotu parametrov a dát. Parametre generujeme z apriórnych rozdelení dát, ktoré použijeme na výpočet odhadovaného procesu volatility. Vzorky pre náhodnú premennú y_t vzhľadom na odhad procesu volatility získame využitím cyklu s pevným počtom opakovaní pre $\forall t \in T$.

```
model {
  // apriórne rozdelenia parametrov
  mu ~ normal(-10, 5);
  phi_hat ~ beta(20, 1.5);
  s2 ~ inv_gamma(2.5, 0.025);
  h[1] ~ normal(mu, sigma/sqrt(1 - phi*phi));

  // odhadovanie latentnej volatility
  for(t in 2:T)
     {
      h[t] ~ normal(mu + phi*(h[t - 1] - mu), sigma);
     }

  // funkcia vierohodnosti
  for(t in 1:T)
     {
```

```
        y[t] ~ normal(0, exp(h[t]/2));
      }
}
```

- **blok generujúci veličiny modelu** je nepovinný, resp. doplňujúci blok, ktorý umožňuje generovanie ďalších veličín z odhadnutých parametrov a modelu. V takomto prípade sa generuje vektor `log_lik`, ktorý obsahuje logaritmické podmienené hustoty pozorovaní, ktoré sa používajú na výpočet informačných kritérií. Funkcia `normal_lpdf` vracia logaritmickú hodnotu hustoty pravdepodobnosti pre každú hodnotu vstupného vektora $\boldsymbol{y}$, pri daných parametroch. Stan používa na výpočty logaritmy v pravdepodobnostných výpočtoch, čím poskytuje presnejšie výpočty, ktoré sú robustnejšie vzhľadom na zaokrúhľovanie a znižuje výpočtovú náročnosť.

```
generated quantities {
      vector[T] log_lik;
      for (t in 1:T){
       log_lik[t] = normal_lpdf(y[t]| 0, exp (h[t]/2));
      }
}
```

Jadrom balíčka RStan je funkcia `stan`, ktorá vykonáva proces vzorkovania z aposteriórnej hustoty, pomocou vylepšenej – NUTS verzie Hamiltonovej Monte Carlo metódy. Funkcia `stan` má tvar

```
fit <- stan(model_code = model_A, data = stan_data, chains = 2,
            warmup = 10000, iter = 20000, thin = 1)
```

kde `model_code` predstavuje definíciu modelu v blokoch, ktorú sme vyššie uviedli pre model A, `chains` predstavuje počet reťazcov, `iter` predstavuje počet iterácii, pričom počet warm up iterácii je štandardne nastavený na polovicu z počtu iterácii, čo sa dá prispôsobiť voľbou `warmup`. Parameter `thin` predstavuje periódu ukladania pozorovaní, tzv. vyššie spomínaný *thinning* a nakoniec vstupne dáta, ktoré definujeme pomocou `data`, ktorý definujeme pomocou funkcie `list`.

```
stan_data <- list(T = length(ret), y = ret)      // definícia vstupu
```

Funkcia `stan` ponúka množstvo ďalších nastavení, medzi ktoré patrí aj špecifikácia algoritmu, kde si užívateľ vie zmeniť štandardne nastavenú metódu NUTS za metódu HMC.

8.2 Techniky matematického modelovania

Techniky matematického modelára vyžadujú niekoľko krokov, ktoré zahŕňajú získavanie dát, tvorbu modelu, jeho testovanie a nakoniec porovnávanie v rámci skupiny modelov. Tento postup je do istej miery algoritmický, avšak vyžaduje si aj isté empirické zásahy, ktoré zahŕňajú napríklad opakované merania s cieľom získania čo najkonzistentnejších a najstabilnejších výsledkov, či dodatočné úpravy v modelovaní, pre zaistenie konvergencie algoritmu.

1. **Získanie a spracovanie dát** zahŕňalo načítanie dát vybraných výmenných kurzov z portálu Yahoo Finance [41], na ktoré sme využívali funkciu `getSymbols` z knižnice `quantmod` zabudovanú v softvéri R [24]. Získané výnosy sme následne upravili o strednú hodnotu a logaritmovali pomocou funkcie `logret` zabudovanej v knižnici `stochvol`[1]. Upravené log-výnosy sme následne rozdelili na jednoročné časové okná, ktoré sme následne použili ako vstupný vektor dát.

2. **Modelovanie** sme uskutočnili pomocou funkcie `stan`, pričom pre každé časové okno a pre každý z modelov sme modelovanie uskutočnili päťkrát a prezentované výsledky v tejto diplomovej práci predstavujú najlepší výsledok dosiahnutý pri opakovanom modelovaní daného časového okna a modelu. Opakovaným meraním sme posudzovali konzistenciu výsledkov, no najmä sme analyzovali správnosť konvergencie, ktorá nebola vždy pri modelovaní zabezpečená a vyžadovala si dodatočné empirické zásahy do modelovania. Pri modelovaní s nižším počtom iterácií sme sa stretli s problémami konvergencie, na ktoré nás Stan upozornil, no tieto problémy naznačovala, okrem varovaní, aj aposteriórna analýza výsledkov modelovania vo funkcii `summary`. Pre potreby zabezpečenia konvergencie a súčasne aj pre zachovanie konzistencie v modelovaní sme zvýšili počet iterácii, pričom sme každé časové okno modelovali pre 2 reťazce, kde každý reťazec zahŕňal 10 000 warm up iterácii a 30 000 post-warm up iterácii, čo prestavuje spolu 40 000 iterácií pre každý reťazec. Pre odstránenie autokorelácie medzi susediacimi členmi

v reťazcoch využijeme príkaz `thin = 10`, ktorý ponechá každú desiatu vzorku, čím prispejeme k zvýšeniu nezávislosti medzi vzorkami, a tým pádom k zlepšeniu presnosti odhadov parametrov. Výsledná vzorka pre každý reťazec bude teda obsahovať každú desiatu vzorku z 30 000 vzoriek, čo predstavuje 3 000 vzoriek. Výsledné štatistiky sú získavané z dvoch reťazcov priemerovaním výsledkov medzi reťazcami, čiže v našom prípade z celkového počtu 6 000 vzoriek. Vzorkovanie pomocou funkcie `stan` sme spustili pomocou príkazu:

```
fit <- stan(model_code = model_A, data = stan_data, chains = 2,
            warmup = 10000, iter = 40000, thin = 10).
```

3. **Validácia modelov** Z pätice kandidátov (piatich opakovaní modelovania) pre konkrétne časové okno a konkrétny model, sme na základe aposteriórnej analýzy parametrov vybrali najlepšieho kandidáta, ktorý reprezentuje dané časové okno. Pozerali sme sa pri tom na kritéria ako sú efektívna veľkosť vzorky, $\hat{R}$ štatistika či vývoj odhadov parametrov v čase (tzv. traceplot), ktorý zobrazuje či a ako dobre simulácie v jednotlivých reťazcoch splynuli. Takýmto spôsobom sme pre každý model získali výsledky, ktoré reprezentujú dané časové okno pre každý z modelov A, B, C a D.

4. **Testovanie a porovnanie modelov** Pre dané časové okno sme štvoricu modelov testovali, pričom sme skúmali základné predpoklady chybového člena ϵ_t, teda jeho rozdelenie pravdepodobnosti, jeho strednú hodnotu a varianciu, čím sme potvrdili, alebo aj vyvrátili, korektnosť použitia modelu na danom časovom okne. Modely sme taktiež aj porovnávali medzi sebou z hľadiska komplexnosti daného modelu a mierou dobrej zhody pomocou informačných kritérií. Výsledky sme nakoniec vecne zhrnuli v tabuľkách a grafoch, ktoré sme následne vyhodnotili.

8.3 Postup aposteriórnej analýzy výsledkov

Pre každé časové okno a pre každý model sme mali k dispozícii päticu kandidátov, ktoré sme vo fáze validácie a testovania analyzovali a vzájomne porovnávali.

Prvou metódou overovania bolo hrubé porovnanie výsledkov modelovania pre

[1] Balíček `stochvol` [19] taktiež ponúka možnosti modelovania stochastickej volatility.

určité časové okno a model. Konkrétne sme pre časové okno 2023 a model A porovnali sumárne štatistiky a výsledky modelovania. Skúmali sme, či tieto výsledky vykazujú podobné charakteristiky a či neexistujú výrazné rozdiely, ktoré by naznačovali odchýlku alebo divergenciu.

Následne sme hodnotili modely na základe efektívnej veľkosti vzorky $\hat{N}_{\text{eff}}$, pričom sme sa zamerali na identifikáciu modelu s najvyššou hodnotou $\hat{N}_{\text{eff}}$. Čím model dosahoval vyššiu hodnotu $\hat{N}_{\text{eff}}$, tým disponoval s väčším počtom nezávislých vzoriek, čo viedlo k najpresnejším odhadom parametrov modelu. Zároveň sme analyzovali odhad štandardnej chyby pre MCMC metódu MCMC-SE(θ), ktorá je na veľkosť efektívnej vzorky naviazaná.

Pre každý z piatich kandidátov sme súčasne skúmali konvergenciu reťazcov, na základe porovnávania distribúcie jednotlivých parametrov, ktorú sme analyzovali graficky a aj analyticky pomocou $\hat{R}$ štatistiky, kde hodnota $\hat{R}$ štatistiky blízka 1 naznačovala dobrú konvergenciu reťazcov.

V neposlednom rade sme sa zaujímali o to, ako jednotliví kandidáti obstáli v hodnotení na základe informačných kritérií, ktoré posudzovali mieru, ako dobre sa vie model prispôsobiť dátam, pričom sa berie do úvahy aj zložitosť tohto modelu.

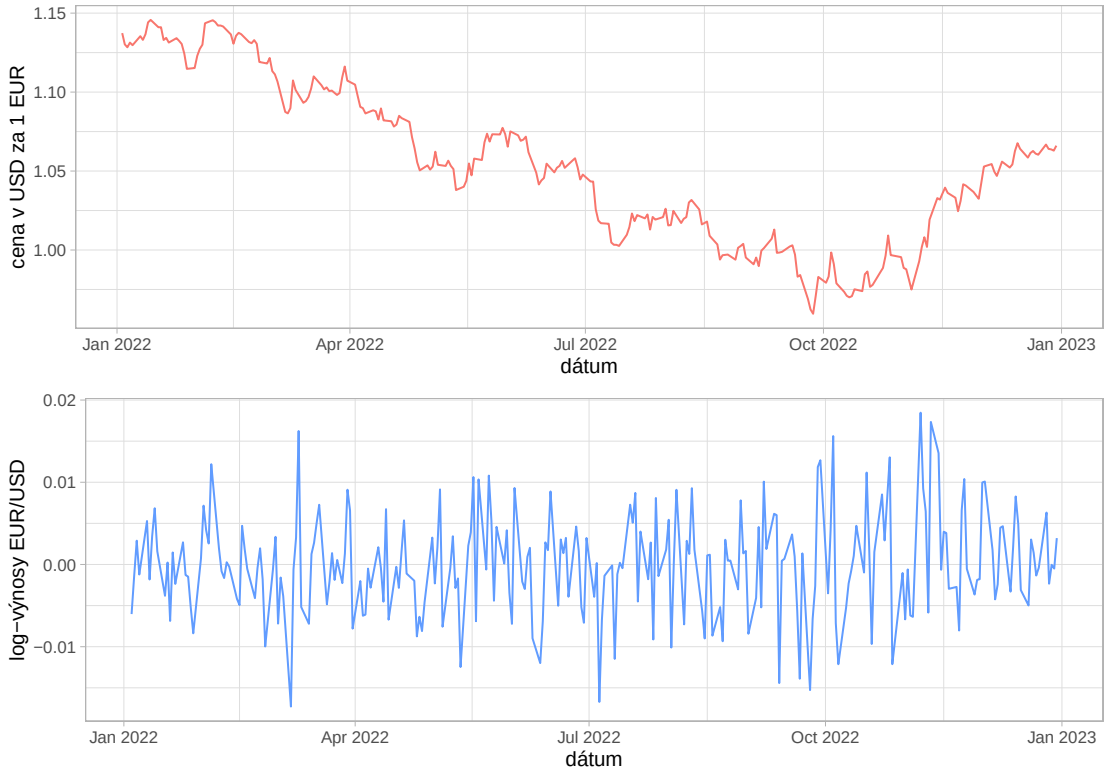
Aposteriórnou analýzou sme vybrali „víťaza“ spomedzi kandidátov, ktorý najlepšie obstál v testovaniach, pričom výsledky uvádzame v tejto diplomovej práci.

8.4 Modelovanie výmenného kurzu EUR/USD v roku 2022

Kurz EUR/USD je jedným z najdôležitejších výmenných kurzov na trhu, ktorý určuje hodnotu jedného eura v amerických dolároch. Tento kurz zohráva významnú úlohu na finančných trhoch a má vplyv na medzinárodný obchod, pretože Spojené štáty sú kľúčovým obchodným partnerom pre eurozónu. Okrem toho, keďže mnohé významné aktíva, ako napríklad akcie amerických spoločností, sú obchodované v dolároch, tak má kurz EUR/USD tiež istý vplyv na dynamiku finančných trhov.

V našej analýze sme sa zamerali na štyri jednoročné časové okna od januára 2020 do decembra 2023, ktoré teda obsahovali celkovo štyri modelovacie časové okná, každé s približne 260 dátovými bodmi.

Rok 2022 predstavoval pomerne „búrlivé“ obdobie pre vývoj výmenného kurzu EUR/USD, kedy držiteľ eura dokázal za jedno euro získať stále menej a menej tovaru v amerických dolárov. Výmenný kurz sa v tomto časovom okne pohyboval v rozmedzí 0,96 až 1,15 amerického doláru za jedno euro, no v priebehu roka sme zaznamenali viacero výraznejších zmien vo vývoji ceny, ktoré sa pretavili aj do citlivosti log-výnosov.



Obr. 6: Vývoj ceny a log-výnosov výmenného kurzu EUR/USD v roku 2022 (*zdroj*: vlastné spracovanie podľa [41])

Na modelovanie časového okna sme použili dva nezávislé reťazce, každý po 40 000 iterácii, tak, ako sme uviedli v podkapitole 8.2. Pre každý z modelov sme päťkrát zopakovali meranie a na základe aposteriórnej analýzy výsledkov sme vybrali model, ktorý dosahoval najlepšie hodnotenia na základe kritérií opísaných v podkapitole 8.3.

Tabuľka 1 zahŕňa strednú hodnotu a štandardnú odchýlku (v zátvorke) apriórnych nastavení parametrov jednotlivých modelov a rovnako aj ich aposteriórne odhady, ktoré predstavujú priemer a štandardnú odchýlku. Pri apriórnych rozdeleniach pravdepodobnosti, ktorých hyperparametre predstavovali parametere tvaru, sme prvý a druhý moment vypočítali na základe teoretických vlastností daného rozdelenia pravdepodobnosti.

Tabuľka 1: Prehľad nastavení parametrov apriórnych rozdelení a ich priemerný aposteriórny odhad pre výmenný kurz EUR/USD za rok 2022. Pre každý parameter a rozdelenie predstavujú hodnoty v prvom riadku priemer a v zátvorke štandardnú odchýlku odhadu.

výmenný kurz EUR/USD v roku 2022					
parameter	rozdelenie	model A	model B	model C	model D
μ	apriórne	$-10{,}00$	$-10{,}00$	$-10{,}00$	$-10{,}00$
		(5,00)	(5,00)	(5,00)	(5,00)
	aposteriórne	$-10{,}310$	$-10{,}301$	$-10{,}218$	$-10{,}261$
		(0,241)	(0,245)	(0,260)	(0,239)
ϕ^*	apriórne	0,86	0,86	0,86	0,86
		(0,11)	(0,11)	(0,11)	(0,11)
	aposteriórne	0,939	0,935	0,935	0,935
		(0,045)	(0,049)	(0,051)	(0,047)
σ_η^2	apriórne	0,0125	0,0125	0,0125	0,0125
		(0,001)	(0,001)	(0,001)	(0,001)
	aposteriórne	0,051	0,055	0,042	0,043
		(0,047)	(0,051)	(0,042)	(0,038)
ν	apriórne	–	0	3	0,06
		–	(2)	(3)	(0,001)
	aposteriórne	–	$-0{,}009$	7,922	1,655
		–	(0,080)	(3,790)	(0,276)

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Z výsledkov v tabuľke 1 môžeme pozorovať, že odhady získané modelovaním sa približujú apriórnym nastaveniam hyperparametrov. Bez ujmy na všeobecnosti, uvažujme pri analýze výsledkov modelovania výmenného kurzu EUR/USD za rok 2022 model D. Priemerná odhadovaná úroveň procesu volatility h_t je $\bar{\mu} = -10{,}261$, čo sa môže zdať ako nezmyselná hodnota, keďže volatilita nenadobúda záporné hodnoty. Avšak, keďže μ predstavuje hladinu procesu volatility h_t, tak samotnú hladinu volatility získame transformáciou vzťahu (1), čo formálne znamená:

$$\bar{h} = \ln \bar{\sigma}^2. \tag{28}$$

Vyjadrením volatility vo vzťahu (28) získame hladinu, resp. úroveň volatility v absolútnych číslach, čo je v prípade modelu D rovné:

$$\bar{\sigma} = \exp\left(\frac{\bar{h}}{2}\right) = \exp\left(\frac{\bar{\mu}}{2}\right) = \exp\left(\frac{-10{,}261}{2}\right) = 0{,}0059\,.$$

Vzhľadom k faktu, že volatilita sa zvyčajne uvádza v percentách, tak úroveň volatility v modeli D sa pohybovala okolo 0,59%, čo považujeme pri stabilnom kurze, akým je EUR/USD, ako prijateľnú hodnotu.

Odhad parametera perzistencie ϕ získame spätnou transformáciou parametra ϕ^*, ktorý sme generovali z beta rozdelenia. V prípade modelu D vieme povedať, že priemerná perzistencia na predchádzajúcich hodnotách procesu volatility je rovná

$$\phi = 2 \cdot \phi^* - 1 = 2 \cdot 0{,}935 - 1 = 0,87,$$

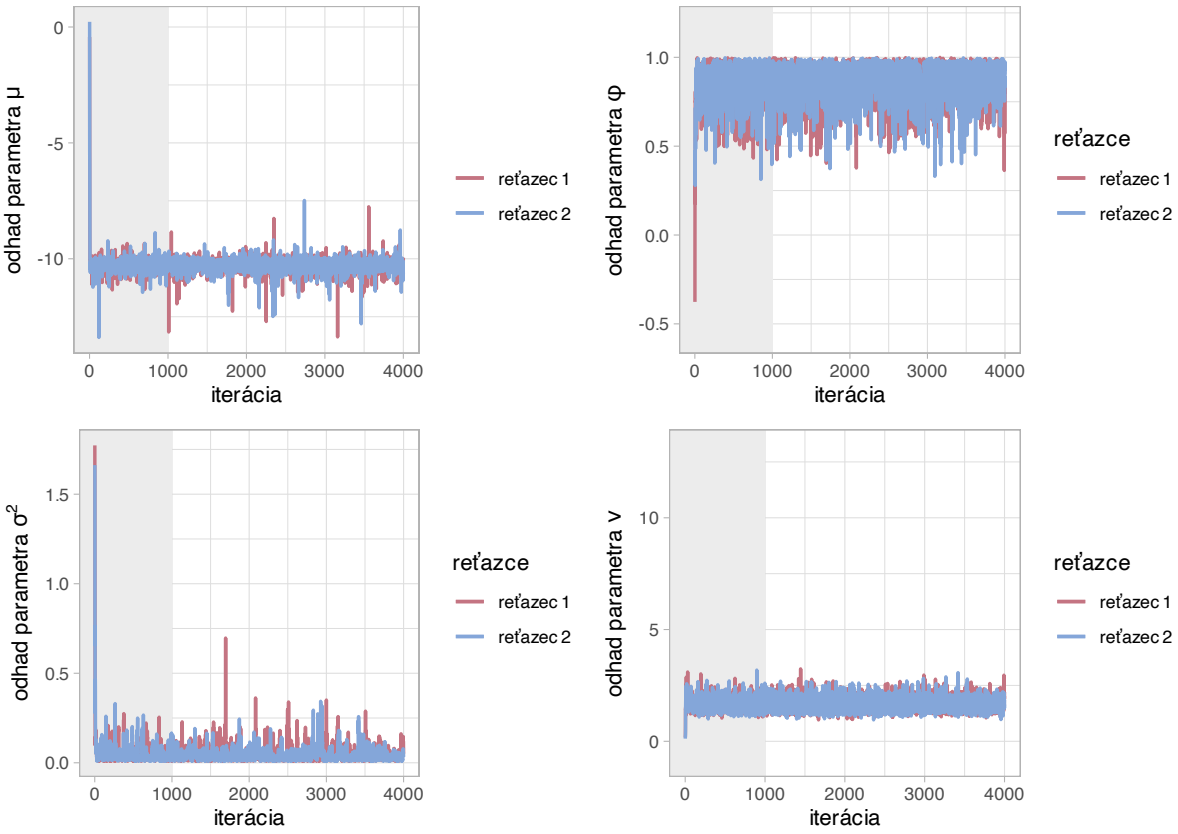
čo predstavuje pozitívnu závislosť na autoregresnom člene h_t a takisto to naznačuje splnenie vlastnosti pretrvávania volatility opísanej v podkapitole 1.2.

Odhady parametra volatility procesu σ_η^2 pozitívne naznačujú nízku hladinu šumu, a to najmä v modeloch C a D, ktorý sa pretaví do odhadu latentnej premennej h_t, čím sa odhady procesu volatility stávajú presnejšími. Vzhľadom k tomu, že kritérium na modelovanie parametra σ_η^2 bolo konzervatívne, šum, ktorý nebol zachytený v procese log-volatility, sa premieta do šumu ϵ_t, ktorý je súčasťou modelovania log-výnosu y_t. Tento fakt podporuje aj skutočnosť, že vlastnosť ťažkých chvostov bola užitočná pri modelovaní šumu v modeloch C a D. Vzhľadom k tomu, že v modeli C modelujeme počet stupňov voľnosti pomocou transformovaného exponenciálneho rozdelenia $\nu - 2 \sim \mathcal{E}$, tak spätnou transformáciou danou:

$$\nu = \bar{\nu} + 2$$

získame finálnu hodnotu počtu stupňov voľnosti, čo je v našom prípade rovné $\nu = 9{,}922$. V modeli D bola priemerná hladina parametra tvaru ν rovná 1,655, pričom v oboch prípadoch tieto hodnoty naznačujú dôležitosť modelovania log-výnosov pomocou rozdelení pravdepodobnosti s ťažkými chvostami, v modeli D uvažujeme dokonca rozdelenie pravdepodobnosti s hrubými chvostami. V prípade modelu B, hodnota parametra tvaru ν sa nelíši významne od nuly, čo naznačuje, že dáta neprejavujú žiadne výrazné zošikmenie, čím sa modely A a B navzájom takmer vôbec nelíšia.

Správnosť konvergencie, nezávislosť získaných vzoriek a takisto aj nepresnosť odhadov MCMC-SE(θ) sme analyticky zhrnuli v tabuľke 2, pričom konvergenciu sme overovali aj vizualizáciou distribúcie reťazcov jednotlivých parametrov na obrázku 7, pričom bez ujmy na všeobecnosti publikujeme výsledky modelu D.



Obr. 7: Vizualizácia distribúcie odhadnutých parametrov modelu D pre výmenný kurz EUR/USD v roku 2022. Prvých tisíc iterácií (sivá časť) predstavuje warm up fázu, zvyšných tritisíc iterácií predstavuje post-warm up fázu.

(*zdroj*: vlastné spracovanie)

Ak by sa distribúcia parametrov v jednotlivých reťazcoch neprekrývala, mohlo by

to znamenať, že Markovove reťazce nedosiahli cieľové rozdelenie, a teda neskonvergovali do rovnovážneho stavu. V takomto prípade by sa distribúcia parametrov výrazne líšila, čo by viedlo k nepresným a vychýleným odhadom. Takýto stav by analyticky zachytilo kritérium $\hat{R}$ štatistiky.

Tabuľka 2: Analýza konvergencie a správnosti odhadov parametrov z aposteriórneho rozdelenia pre výmenný kurz EUR/USD za rok 2022.

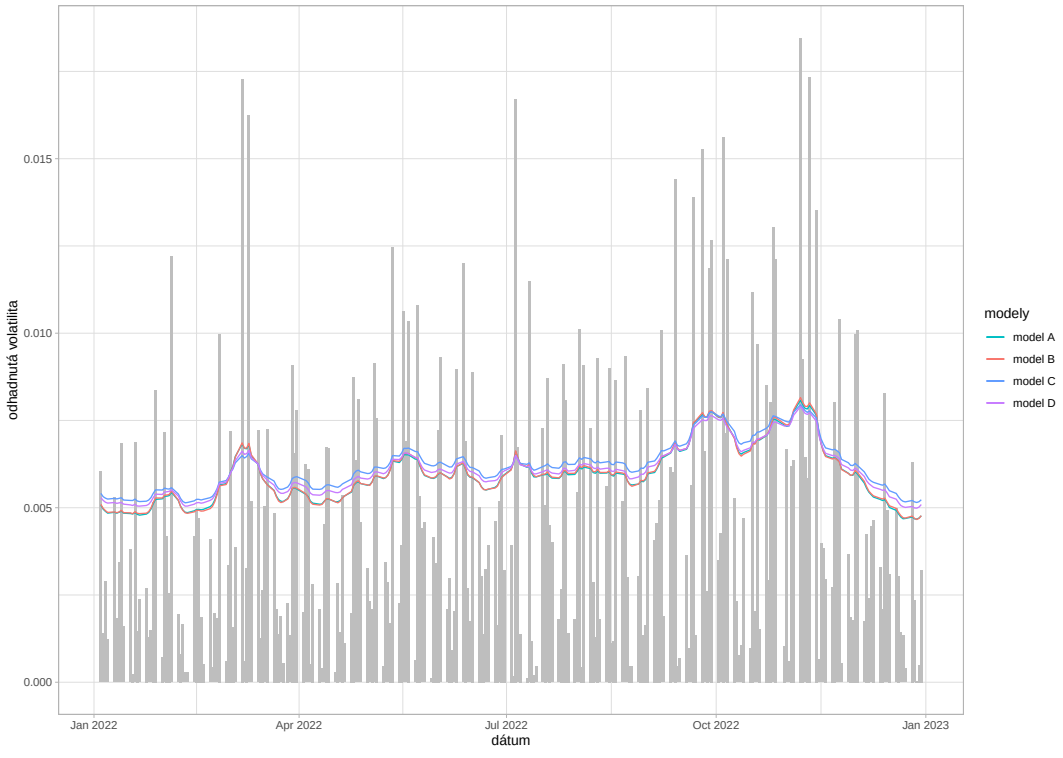
výmenný kurz EUR/USD v roku 2022				
model	parameter	$\hat{R}$	$\hat{N}_{\text{eff}}$	MCMC-SE
model A	μ	0,9998	5668	0,0032
	ϕ^*	1,0031	807	0,0016
	σ_η^2	1,0037	496	0,0021
model B	μ	0,9999	5560	0,0033
	ϕ^*	0,9998	828	0,0017
	σ_η^2	1,0001	475	0,0024
	ν	1,0005	6343	0,0010
model C	μ	0,9998	5374	0,0036
	ϕ^*	1,0006	1171	0,0015
	σ_η^2	1,0027	482	0,0020
	ν	0,9998	5071	0,0532
model D	μ	0,9999	5294	0,0033
	ϕ^*	0,9999	1170	0,0014
	σ_η^2	1,0056	451	0,0018
	ν	1,0007	4779	0,0040

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Výsledky v tabuľke 2 naznačujú stabilné a robustné odhady parametrov. Čo sa týka $\hat{R}$ štatistiky, tak jej hodnoty sú pre jednotlivé modely a ich parametre veľmi

blízke hodnote 1, čo poukazuje na správnu a stabilnú konvergenciu nezávislých reťazcov. Pre veľkosť efektívnej vzorky $\hat{N}_{\text{eff}}$ platí, že dosahuje pomerne vysoké hodnoty, až na parameter σ_η^2, ktorého $\hat{N}_{\text{eff}}$ sa nám nepodarilo zvýšiť ani pri reparametrizácii modelov pomocou inverzného gama rozdelenia opísaného v podkapitole 3.2, ani pri použití techniky zrieďovania, tzv. thinningu. Veľkosť efektívnej vzorky pre σ_η^2 je nad kritickou hranicou, takže odhady môžeme označiť za hodnoverné. Chybovosť odhadov MCMC-SE(θ) dosahovala nízke hodnoty, čo nasvedčuje, že odhady získané MCMC môžu byť spoľahlivé. Celkovo, z aposteriórnej štatistiky teda usudzujeme, že modely poskytujú stabilné a nevychýlené odhady parametrov.

Na obrázku 8 pozorujeme odhadnutú volatilitu v absolútnych číslach v rámci jednotlivých modelov na pozadí log-výnosov transformovaných absolútnou hodnotou. Tento postup zabezpečuje prehľadnejšiu vizualizáciu, pretože absolútna hodnota transformuje záporné log-výnosy na kladné čísla. Táto vizualizácia umožňuje zreteľnejšie sledovať zmeny vo volatilite, keďže volatilita je citlivá na akékoľvek zmeny log-výnosov, či už ide o ich rast alebo pokles.



Obr. 8: Vizualizácia odhadnutej volatility s absolútnou hodnotou logaritmických výnosov pre výmenný kurz EUR/USD v roku 2022
(*zdroj*: vlastné spracovanie podľa [41])

Z obrázku 8 je zrejmé, že jednotlivé modely správne reagujú na zmeny v log-výnosoch, nakoľko kopírujú zmeny v log-výnosoch, čo môžeme najlepšie pozorovať v období s výraznejšími zmenami, pričom táto charakteristika súvisí s vlastnosťou zhlukovania volatility. Trajektórie odhadnutej volatility sa kopírovali a líšili sa na úrovni približne jednej tisíciny, čo v percentách, v ktorých sa volatilita zvyčajne udáva, predstavuje menej než jedno percento.

Modely posudzujeme z kvalitatívneho hľadiska na základe testovania ich predpokladov a rovnako aj navzájom, na základe kvality zachytenia dát a ich komplexnosti pomocou informačných kritérií. Tabuľka 3 obsahuje výsledky testovania štatistických vlastností šumu ϵ_t, ktoré zahŕňajú mieru „úspešnosti“ testov pri 5 000 simuláciách, a rovnako aj mediánovú hodnotu jednotlivých testov[2].

Tabuľka 3: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz EUR/USD v roku 2022

EUR/USD, 2022	**percentuálny podiel p-hodnôt väčších než hladina α = 0,05**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	93,20%	100,00%	43,36%	–	–	–
model B	94,28%	100,00%	43,18%	94,60%	100,00%	98,02%
model C	98,42%	100,00%	76,22%	98,34%	100,00%	97,02%
model D	98,24%	100,00%	63,38%	100,00%	98,96%	98,96%

EUR/USD, 2022	**mediánová p-hodnota**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	0,1608	0,8780	0,0263	–	–	–
model B	0,1628	0,8775	0,0264	0,1665	0,8693	0,5136
model C	0,2039	0,8904	0,2666	0,2499	0,8716	0,4989
model D	0,1994	0,8932	0,1341	0,9763	0,8713	0,5369

(*zdroj*: vlastné spracovanie)

Výsledky testov dosahovali pomerne veľkú percentuálnu úspešnosť až na chí-kvadrát test, ktorého výsledky v prípade modelov A a B dosahovali nízku percentuálnu úspešnosť, a zároveň ich mediánová p-hodnota bola nižšia, ako kritická hladina 0,05.

[2]Hviezdička (*) pri A-D teste symbolizuje rozdelenie v závislosti od modelu; t. j. model B - zošikmené normálne rozdelenie, model C - Studentovo t-rozdelenie, model D - GED rozdelenie.

Testovanie šumu voči jednotlivým rozdeleniam dopadlo v prospech modelu D, ktoré so 100 % úspešnosťou a najvyššou mediánovou p-hodnotou rovnou 0,9763 jednoznačne viedlo v porovnaní s inými modelmi. Výsledky testovania pre model C síce nevyvrátili splnenie predpokladov, avšak nevylúčili ani normálne rozdelenie.

V rámci vzájomného porovnania modelov nás zaujíma, ako uspeli modely C a D na základe informačných kritérií, nakoľko modely A a B nesplnili predpoklad o jednotkovej štandardnej odchýlke. Výsledky, ktoré uvádzame v tabuľke 4, po vylúčení modelov A a B, uprednostnili s malým náskokom model D, ktorý podľa našej analýzy najlepšie obstál v modelovaní volatility pre výmenný kurz EUR/USD v roku 2022.

Tabuľka 4: Prehľad výsledkov informačných kritérií DIC a WAIC pre výmenný kurz EUR/USD za rok 2022.

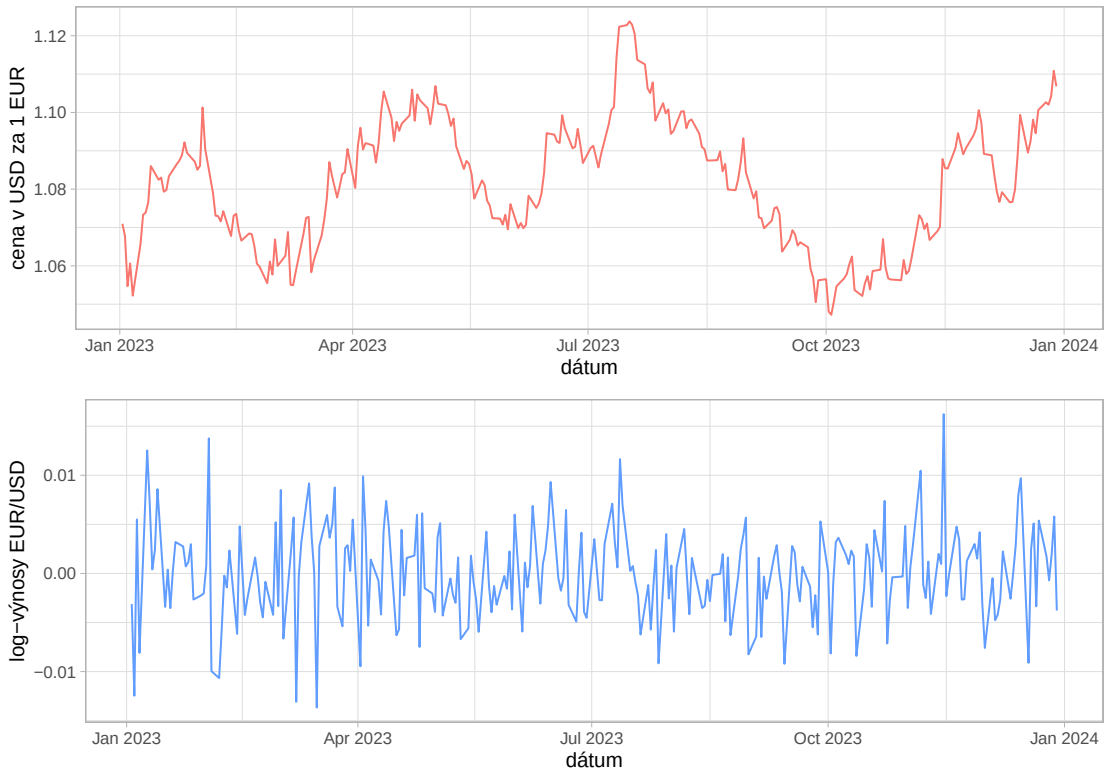
výmenný kurz EUR/USD v roku 2022				
model	DIC	poradie	WAIC	poradie
model A	−1907,660	1	−1908,120	1
model B	−1905,569	2	−1905,945	2
model C	−1899,179	4	−1903,461	4
model D	−1901,788	3	−1905,174	3

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Všeobecne môžeme skonštatovať, že aj v prípade stabilného výmenného kurzu, akým je EUR/USD, bolo užitočné nasadenie modelu D a jeho GED rozdelenia, ktoré v našej analýze obstálo najlepšie.

8.5 Modelovanie výmenného kurzu EUR/USD v roku 2023

Na časovom okne 2023 pre výmenný kurz by sme chceli demonštrovať fakt, že nie v každom prípade je na modelovanie log-výnosov nutné použiť rozdelenia pravdepodobnosti s ťažkými chvostami. Na obrázku 9 pozorujeme časový vývoj ceny amerického dolára za jedno euro v roku 2023. Síce modelovacie okno obsahovalo pomerne strmé vzo-



Obr. 9: Vývoj ceny a log-výnosov výmenného kurzu EUR/USD v roku 2023 (*zdroj*: vlastné spracovanie podľa [41])

stupy a pády, ktoré by mohli indikovať extrémnejšie hodnoty, no napriek tomu všetky modely obstáli ako v modelovaní samotnej volatility, tak aj v následnej validácií, pričom výsledky testovania uvádzame v tabuľke 5.

Testy nevyvrátili hypotézy o splnení predpokladov pre výmenný kurz EUR/USD v modelovacom roku 2023. Percentuálna miera úspešnosti všetkých testov bola nad úrovňou 80 %, čím testy zhodne potvrdili relevantnosť modelov pri danom časovom okne. Mediánová hodnota v prípade A-D testu bola najvyššia pri modeli D, a to na úrovni 0,9886, pri ktorom testy presvedčivo nevyvrátili hypotézu, že šum ϵ_t sa riadi podľa GED rozdelenia. Prekvapením sa stal model C, ktorého mediánová p-hodnota v prípade Studentovho t-rozdelenia nepotvrdila ťažké chvosty šumu tak jednoznačne, ako tomu nasvedčujú výsledky testov v prípade modelu D a jeho GED rozdelenia. Čo sa týka testov o nulovej strednej a jednotkovej štandardnej odchýlke, tak nulové hypotézy nevyvrátili.

Vzhľadom na skutočnosť, že testy nezamietli vhodnosť a relevantnosť modelov, tak nás zaujímalo ich porovnanie na základe informačných kritérií, ktoré sme zhrnuli

Tabuľka 5: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz EUR/USD v roku 2023

EUR/USD, 2023	percentuálny podiel p-hodnôt väčších než hladina α = 0,05					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	99,92%	100,00%	82,42%	–	–	–
model B	99,92%	100,00%	82,28%	99,92%	100,00%	99,74%
model C	100,00%	100,00%	93,20%	99,32%	100,00%	94,26%
model D	99,96%	100,00%	93,38%	100,00%	98,64%	98,64%

EUR/USD, 2023	mediánová p-hodnota					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	0,4621	0,8853	0,3158	–	–	–
model B	0,4686	0,8884	0,3140	0,4793	0,9185	0,5578
model C	0,5931	0,9052	0,4315	0,2764	0,9145	0,2745
model D	0,5147	0,8940	0,4813	0,9886	0,9185	0,4936

(*zdroj*: vlastné spracovanie)

v tabuľke 6. Z výsledkov vyplýva tesné víťazstvo modelov A a B, ktoré v prípade DIC aj WAIC dosiahli lepšie skóre v porovnaní s modelmi C a D. Všeobecne môžeme skonštatovať, že pre výmenný kurz EUR/USD v modelovacom časovom okne 2022 jednotlivé modely splnili svoje predpoklady a sú vhodnými kandidátmi na odhadovanie volatility.

Na tomto príklade sme chceli ukázať, že „menej je niekedy viac“, pričom toto porekadlo má svoje uplatnenie aj v matematickom modelovaní, kde aj normálne rozdelenie môže byť v niektorých prípadoch užitočnejšie, než rozdelenia s ťažkými chvostami.

8.6 Zhrnutie výsledkov modelovania pre výmenný kurz EUR/-USD

V prílohe A sme zhrnuli výsledky modelovania pre výmenný kurz EUR/USD pre štyri jednoročné obdobia 2020, 2021, 2022 a 2023. Tabuľka 11 obsahuje odhady strednej hodnoty a štandardnej odchýlky parametrov pre jednotlivé modely A, B, C a D. Tabuľky 12, 13 a 14 obsahujú sumárne štatistiky ukazovateľov pre $\hat{R}$ štatistiku,

Tabuľka 6: Prehľad výsledkov informačných kritérií DIC a WAIC pre výmenný kurz EUR/USD za rok 2023.

výmenný kurz EUR/USD v roku 2023				
model	DIC	poradie	WAIC	poradie
model A	−2040,093	1	−2040,495	1
model B	−2038,418	2	−2038,701	2
model C	−2034,023	4	−2037,018	4
model D	−2035,468	3	−2038,060	3

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

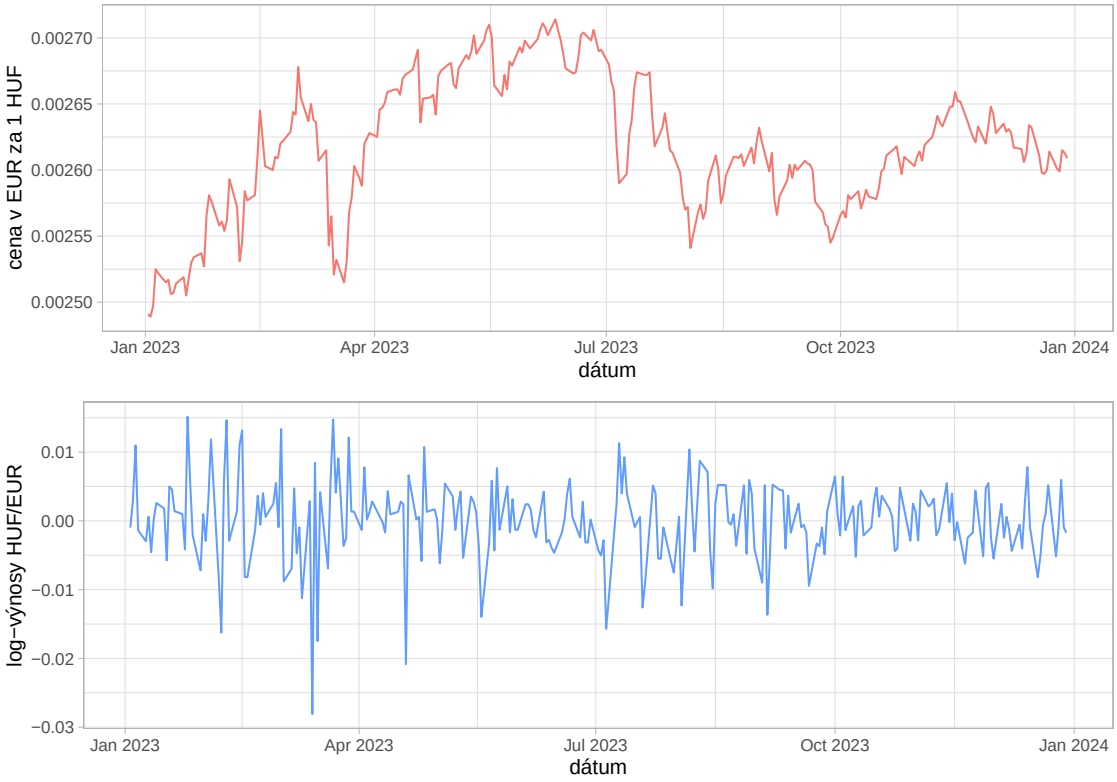
veľkosť efektívnej vzorky $\hat{N}_{\text{eff}}$ a štandardnú chybu pre MCMC-SE, v uvedenom poradí. Výsledky testovania a porovnanie jednotlivých modelov sme zhrnuli v tabuľkách 15, 16, 17, 18 a 19, pričom pre úplnosť a lepšie porovnanie sme do príloh doplnili aj prezentované výsledky z tejto kapitoly.

Odhady parametrov pre jednotlivé časové okná ukázali, že modely A a B sa nelíšili, nakoľko odhady parametera tvaru ν v prípade modelu B sa pohybovali na úrovni blízkej nule. Tento fakt naznačuje, že v log-výnosoch nebol výrazný sklon, ktorý by si vyžiadal na modelovanie šumu zošikmené rozdelenie pravdepodobnosti. Z odhadov taktiež pozorujeme vysokú pozitívnu závislosť na autoregresnom člene h_t, ktorú charakterizuje parameter ϕ^* blízky jednotke. Odhady parametra σ_η^2 naznačujú pozitívne nízku mieru šumu pre odhady procesu volatility. Čo sa týka odhadov parametra ν v modeloch C a D, tie sa pohybovali v intervaloch, kedy parí Studentovo t a GED rozdelenie do triedy rozdelení s ťažkými chvostami, pričom práve GED rozdelenie suverénne obstálo v testoch o splnení predpokladov. Prekvapením bol model C, ktorého odhad parametra počtu stupňov voľnosti chvosty nasvedčoval, že rozdelenie má ťažké chvosty, no v testovaní sa nevylúčilo normálne rozdelenie šumu ϵ_t. Výsledky aposteriórnej analýzy potvrdili konvergenciu reťazcov a aj nevychýlenosť či nízku chybovosť odhadov. Testovanie dopadlo v závislosti od modelovaného časového okna aj v prospech modelov A a B, či modelov C a D, z čoho usudzujeme, že neexistuje jednoznačné

kritérium, na základe ktorého by sme ešte pred uskutočnením modelovania vedeli posúdiť, ktorý typ modelov bude v prípade modelovaných dát najvhodnejším. Odporúčame teda zostavenie takého portfólia modelov, ktoré obsahuje modely s ťažkými a ľahkými chvostami, či modely na zachytenie špecifického tvaru dát.

8.7 Modelovanie výmenného kurzu HUF/EUR v roku 2023

Oproti stabilnému výmennému kurzu EUR/USD, ktorý je významným hráčom na svetovom trhu sme sa zamerali aj na analýzu malej, otvorenej ekonomiky Maďarska, ktorá zaznamenala viaceré prudké pády. Maďarský forint patrí medzi volatilnú menu, z dôsledku viacerých faktorov, medzi ktoré patria napríklad medzinárodné obchodné dohody, domáce hospodárstvo či politické rozhodnutia, preto od modelovania maďarského forintu očakávame silnú potrebu nasadenia modelov C a D na zachytenie odľahlých denných log-výnosov. Vývoj výmenného kurzu a jeho log-výnosov HUF/EUR pre rok 2023 sme znázornili na obrázku 10.



Obr. 10: Vývoj ceny a log-výnosov výmenného kurzu HUF/EUR v roku 2023 (*zdroj*: vlastné spracovanie podľa [41])

Nastavenia apriórnych parametrov, ktoré sme uviedli v tabuľke 1 sa nezmenili, nakoľko na modelovanie volatility výmenného kurzu HUF/EUR používame rovnaké modely a rovnaké parametrizácie, ako to bolo v prípade výmenného kurzu EUR/USD. Z tohto dôvodu uvedieme výsledky odhadov parametrov v skrátenej tabuľke, t. j. bez hodnôt nastavení parametrov apriórnych rozdelení.

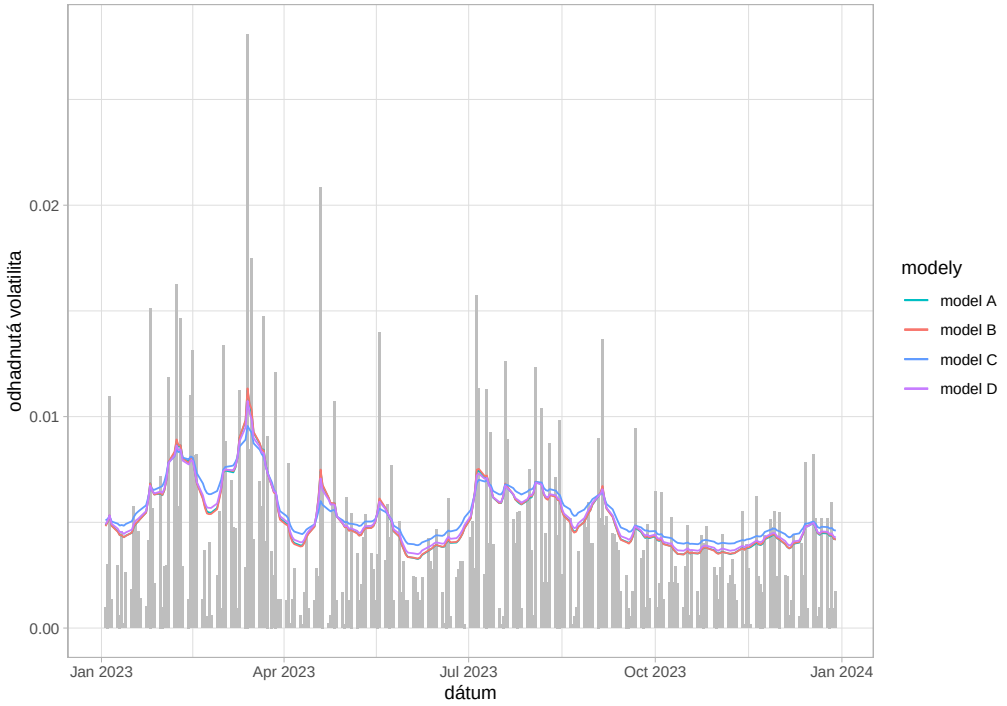
Tabuľka 7: Prehľad aposteriórneho odhadu parametrov pre výmenný kurz HUF/EUR za rok 2023. Pre každý parameter predstavujú hodnoty v prvom riadku priemer a v zátvorke štandardnú odchýlku odhadu.

výmenný kurz HUF/EUR v roku 2023				
parameter	model A	model B	model C	model D
μ	$-10{,}588$	$-10{,}587$	$-10{,}472$	$-10{,}557$
	(0,285)	(0,283)	(0,293)	(0,324)
ϕ^*	0,939	0,937	0,953	0,942
	(0,035)	(0,037)	(0,031)	(0,36)
σ_η^2	0,143	0,151	0,083	0,131
	(0,087)	(0,097)	(0,065)	(0,098)
ν	–	0,029	6,703	1,828
	–	(0,080)	(3,418)	(0,372)

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

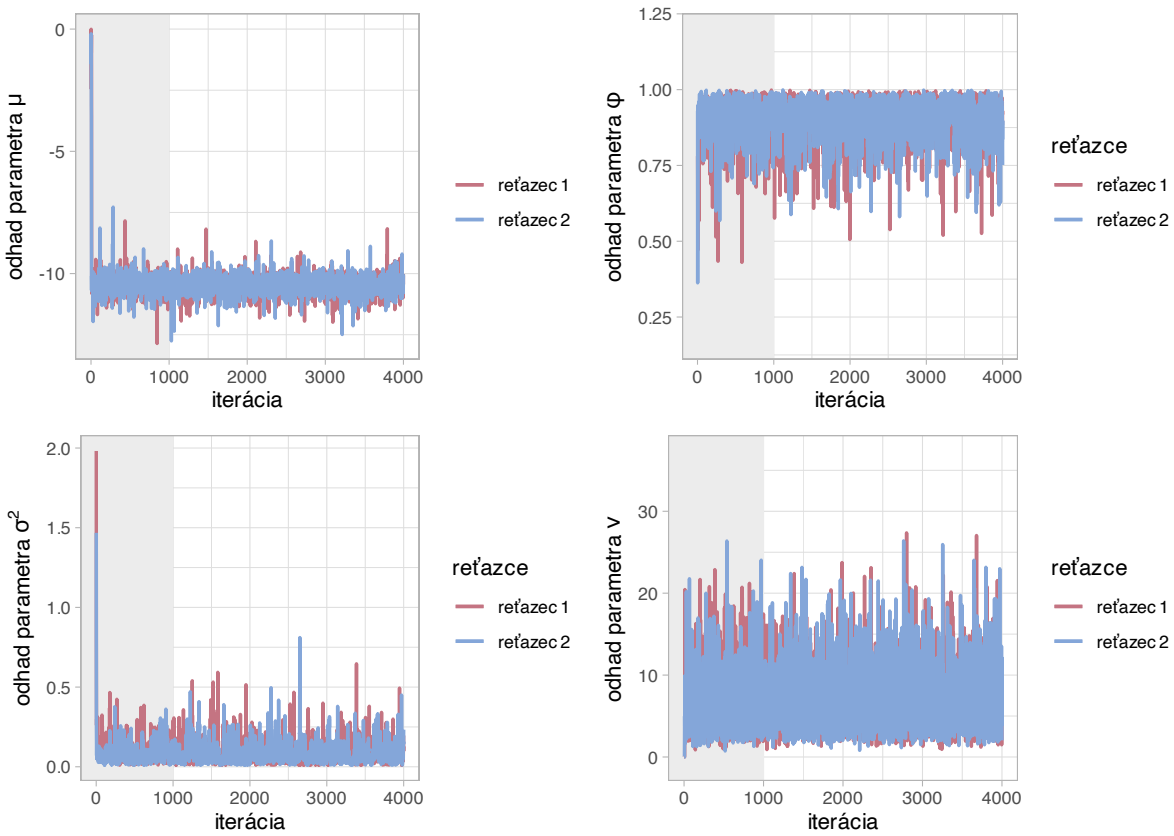
Grafickú vizualizáciu odhadov volatility môžeme pozorovať na obrázku 11. Na grafe je zreteľných niekoľko výrazných zmien v absolútnych hodnotách log-výnosov, ktoré predstavujú extrémnejšie hodnoty v denných log-výnosoch. Takisto môžeme pozorovať, že percentuálna hodnota odhadu volatility sa pohybovala na úrovni okolo 1 %, čo predstavuje vyššiu hodnotu v porovnaní s prezentovaným výmenným kurzom EUR/USD v roku 2022. Jednotlivé modely zachytili toto správanie a prispôsobili mu odhad volatility. Výraznejším modelom bol model C, ktorý v porovnaní s ostatnými modelmi odhadol volatilitu s nižšou mierou citlivosti. V tabuľke 7 s odhadmi pre parametre môžeme pri modeli C sledovať najnižšiu mieru volatility šumu σ_η^2, ktorá sa prenáša do nižšej miery šumu v odhadoch procesu volatility z čoho usudzujeme, že odhady volatility získané z modelu C by mohli byť najpresnejšie.

Čo sa týka odhadov jednotlivých parametrov, bez ujmy na všeobecnosti uvažujme model C, ktorého distribúciu parametrov môžeme pozorovať na obrázku 12. Distribúcie



Obr. 11: Vizualizácia odhadnutej volatility s absolútnou hodnotou logaritmických výnosov pre výmenný kurz HUF/EUR v roku 2023
(*zdroj*: vlastné spracovanie podľa [41])

sa prekrývali pri všetkých parametroch, čo naznačuje stabilnú konvergenciu v aposteriórnom rozdelení. Zároveň, pre parameter počtu stupňov voľnosti ν, ktorý modelujeme pomocou transformácie $\nu - 2$ s exponenciálnym rozdelením, pozorujeme pomerne vysokú mieru volatility odhadu, na úrovni 3,418. Po vzore mnohých autorov [10], [18], [40], ktorí takisto modelovali transformáciu parametera ν pomocou exponenciálneho rozdelenia, aj my pokladáme túto voľbu rozdelenia ako vhodnú, nakoľko toto rozdelenie umožňuje modelovať šum medzi rozdelením s ťažkými chvostami a normálnym rozdelením, pre $\nu \to \infty$.



Obr. 12: Vizualizácia distribúcie odhadnutých parametrov modelu C pre výmenný kurz HUF/EUR v roku 2023. Prvých tisíc iterácií (sivá časť) predstavuje warm up fázu, zvyšných tritisíc iterácií predstavuje post-warm up fázu.
(*zdroj*: vlastné spracovanie)

Výsledky analýzy konvergencie a správnosti odhadov uvádzame v tabuľke 8. Analyticky sme správnosť konvergencie overili pomocou $\hat{R}$ štatistiky, ktorá dosahovala hodnoty okolo 1, ktoré predstavujú takmer prekrývajúcu sa distribúciu dvoch, nezávislých reťazcov. Veľkosť efektívnej vzorky dosahovala hodnoty, ktoré poskytujú nevychýlené odhady, a to aj v prípade parametra σ_η^2, ktorý podobne, ako pri výmennom kurze EUR/USD, dosahoval v porovnaní s ostatnými parametrami najnižšiu, no zároveň akceptovateľnú hodnotu. Nízka chybovosť odhadov MCMC-SE(θ) naznačuje správnosť a nevychýlenosť odhadov.

Tabuľka 8: Analýza konvergencie a správnosti odhadov parametrov z aposteriórneho rozdelenia pre výmenný kurz HUF/EUR za rok 2023.

výmenný kurz HUF/EUR v roku 2023				
model	parameter	$\hat{R}$	$\hat{N}_{\text{eff}}$	MCMC-SE
model A	μ	1,0001	5339	0,0039
	ϕ^*	1,0010	942	0,0011
	σ_η^2	1,0030	718	0,0033
model B	μ	0,9998	5951	0,0037
	ϕ^*	1,0003	981	0,0012
	σ_η^2	1,0010	791	0,0035
	ν	1,0003	5557	0,0011
model C	μ	0,9998	5735	0,0039
	ϕ^*	1,0010	950	0,0010
	σ_η^2	1,0016	627	0,0026
	ν	0,9997	5649	0,0455
model D	μ	1,0004	4765	0,0047
	ϕ^*	0,9997	718	0,0014
	σ_η^2	1,0002	502	0,0044
	ν	0,9999	1625	0,0092

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Výsledky testov v tabuľke 9 vylúčili modely A a B v prípade volatilného obdobia 2023 pre výmenný kurz HUF/EUR. Modely A a B dosahovali nízku percentuálnu mieru úspešnosti, čomu nasvedčuje aj mediánová p-hodnota, ktorá je nižšia, než kritická hladina $\alpha = 0{,}05$. Najlepšie z testovania vyšiel opäť model D, ktorý dosiahol suverénne výsledky v prípade A-D testu, ktorý testoval šum ϵ_t oproti GED rozdeleniu, takisto vo Wilcoxonovom teste o nulovej strednej hodnote a Leveneovom teste o jednotkovej

štandardnej odchýlke. Veľmi tesne testami prešiel aj model C, ktorého percentuálna úspešnosť a aj mediánová p-hodnota bola blízko hraničných hodnôt, pri ktorých by sme model C vylúčili.

Tabuľka 9: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz HUF/EUR v roku 2023

HUF/EUR, 2023	**percentuálny podiel p-hodnôt väčších než hladina α = 0,05**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	14,14%	100,00%	10,26%	–	–	–
model B	14,60%	100,00%	9,40%	16,84%	100,00%	76,28%
model C	25,04%	100,00%	37,36%	51,88%	100,00%	86,22%
model D	16,54%	100,00%	15,54%	100,00%	82,98%	82,98%

HUF/EUR, 2023	**mediánová p-hodnota**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	0,0078	0,8017	< 0,0001	–	–	–
model B	0,0073	0,7967	< 0,0001	0,0091	0,6536	0,2334
model C	0,0174	0,7854	0,0061	0,0540	0,6698	0,3714
model D	0,0091	0,7971	< 0,0001	0,8896	0,6593	0,3095

(*zdroj*: vlastné spracovanie)

V takomto prípade nás zaujímalo, ako obstáli modely C a D na základe hodnotenia informačných kritérií, nakoľko modely A a B prepadli v testovacej fáze.

Z dvojice modelov C a D obstál model D najlepšie, pričom najnižšiu hodnotu dosahoval v oboch informačných kritériách, DIC a aj WAIC. V prípade volatilného časového okna, akým bol rok 2023 sa teda preukázal model D ako najlepší spomedzi kandidátov na modelovanie volatility výmenného kurzu HUF/EUR.

8.8 Zhrnutie výsledkov modelovania pre výmenný kurz HUF/EUR

Výsledky modelovania volatility pre výmenný kurz HUF/EUR sme zhrnuli v prílohe B, pre štyri jednoročné obdobia v období 2020 až 2023. V tabuľke 20 uvádzame odhady pre parametre modelov. Aposteriórnu štatistiku presnosti odhadov sme zhrnuli

Tabuľka 10: Prehľad výsledkov informačných kritérií DIC a WAIC pre výmenný kurz HUF/EUR za rok 2023.

výmenný kurz HUF/EUR v roku 2023				
model	DIC	poradie	WAIC	poradie
model A	−1975,624	1	−1976,280	1
model B	−1973,678	2	−1974,234	2
model C	−1967,023	4	−1968,647	4
model D	−1969,034	3	−1972,068	3

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

v tabuľkách 21, 22 a 23. Validáciu modelov v podobe testovania predpokladov o šume ϵ_t sme zhrnuli v tabuľkách 24, 25 a 26, 27. Nakoniec, jednotlivé porovnanie modelov z hľadiska informačných kritérií uvádzame v tabuľke 28.

Modely A a B sa ani v prípade HUF/EUR takmer vôbec nelíšili, parameter tvaru ν v prípade modelu B bol na úrovni veľmi blízkej nule, čo naznačuje rovnomerné rozloženie log-výnosov okolo nuly. Model B pokladáme za užitočný v prípade modelovania výnosov aktív, akými sú akcie, akciové indexy, či dokonca kryptomeny, kedy ceny v zlých časoch výrazne klesajú, resp. v tých dobrých rastú, čo sa z dlhodobého hľadiska pretaví do extrémnejších hodnôt výnosov a sklonu dát. Z odhadov parametrov modelov v tabuľke 20 vyplýva silná závislosť na autoregresnom člene h_t, ktorá sa pretavila do parametra ϕ^*, nízky šum pri odhade procesu volatility v podobe variancie σ_η^2 a zároveň aj ťažké chvosty pri modeloch C a D, ktoré reprezentoval parameter ν. Aposteriórna štatistika odhadov naznačuje stabilné a nevychýlené odhady, v podobe tzv. „splynutia“ reťazcov v $\hat{R}$ štatistike, vysokej efektívnej veľkosti vzorky $\hat{N}_{\text{eff}}$, či nízkej miere chybovosti na základe MCMC-SE kritériá. Testovanie vyšlo takmer vždy v prospech modelov C a D, ktoré favorizoval najmä A-D test a rovnako aj Leveneov test o štandardnej odchýlke. Odhadnutý šum ϵ_t v prípade modelov A a B takmer vždy nespĺňal kritérium normálneho rozdelenia, resp. kritérium zošikmeného normálneho rozdelenia, no podobne, ako v prípade výmenného EUR/USD, kde sme v podkapitole

8.5 demonštrovali vhodnosť modelov A a B, tak aj v prípade výmenného kurzu HUF/EUR bol rok 2021 zaujímavým rokom, kedy na modelovanie volatility „postačili“ aj modely s ľahkými chvostami. Zaujímavosťou však bolo, že informačné kritéria DIC a WAIC však v tomto prípade preferovali model D pred modelom A.

Záver

Cieľom predkladanej záverečnej práce bolo skúmanie využiteľnosti rozdelení pravdepodobnosti s ťažkými chvostami pri modelovaní latentného procesu volatility pomocou modelov stochastickej volatility. Predkladaná práca poskytuje čitateľovi ucelené spracovanie témy modelovania časovo závislej volatility, ktorá je vybudovaná na netriviálnych matematických konceptoch stochastických simulačných techník, ktoré si vyžadujú pokročilé znalosti z oblasti štatistiky, Markovových reťazcov, diferenciálnych rovníc, dynamiky fyzikálnych systémov, numerických metód a ďalších matematických princípov. Implementačná časť tejto práce zahŕňa praktické modelovanie volatility pre denné log-výnosy výmenných kurzov EUR/USD a HUF/EUR pre štyri jednoročné časové okná od prvého obchodného dňa v januári 2020 do posledného obchodného dňa v decembri 2023. Modelovanie sme uskutočnili pomocou knižnice Stan [27] implementovanej v softvéri R [24] pomocou balíčka `rstan`.

V prvej kapitole sme si predstavili volatilitu, jej vlastnosti a prístupy, akými sa dá pristupovať k jej modelovaniu. Dôležitou súčasťou prvej kapitoly bola demonštrácia dôležitosti modelovania tejto veličiny, ktorá vplýva na dynamiku finančného systému, ceny aktív a je užitočným nástrojom na odhad budúcej rizikovosti obchodu.

V druhej kapitole sme si zadefinovali základný model stochastickej volatility, ktorý bol predstavený britským štatistikom, Stephenom J. Taylorom v roku 1982, ako alternatíva k (G)ARCH modelom, založený na priestore stavov a pravdepodobnosti. Ukázali sme, že nepodmienené výnosy môžu byť leptokurtické aj v prípade základného, Taylorovho modelu - modelu A, čo viedlo k potrebe zavedenia ďalších modelov, ktoré sú vhodnými alternatívami k základnému modelu. Pri modelovaní sme si zostrojili portfólio štyroch modelov, ktoré pokrývalo všetky prípady, ktoré mohli modelované dáta obnášať, a to: šikmosť pomocou zošikmeného normálneho rozdelenia v modeli B a ťažké chvosty, ktoré poskytovalo Studentovo t a GED rozdelenie v modeloch C a D.

Prístup k odhadovaniu parametrov modelu sme predstavili v tretej kapitole, ktorý zahŕňal definíciu bayesovského prístupu k štatistike, pričom v tomto prístupe sme skompletizovali definíciu apriórnych rozdelení jednotlivých modelov A, B, C a D. Takisto sme čitateľovi poskytli náhľad do histórie a vývoja modelov stochastickej volatility. Ďalej sme čitateľa oboznámili s históriou a aj aktuálnym vývojom v modeloch stochas-

tickej volatility, ktorý zahŕňa publikácie, ktoré prvýkrát zadefinovali modely volatility až po publikácie, ktoré zahŕňajú techniky v odhadovaní modelov, ktoré sa vyvíjajú a zefektívňujú dodnes.

Vo štvrtej kapitole sme čitateľa oboznámili s konceptom Monte Carlo simulačných techník, ktoré patria medzi efektívne techniky na získanie odhadov parametrov a latentnej premennej. Vysvetlili sme, že modely stochastickej volatility patria medzi zložité modely, pri ktorých sa v praxi stretávame s problémom analytického vyjadrenia ich aposteriórnej funkcie, a to najmä pre vysokú dimenziu parametrov, ktorú navyše v prípade modelov stochastickej volatility tvoria aj odhady latentnej premennej. Princíp získania proporcionálneho cieľového rozdelenia sme vysvetlili na základe bayesovského prístupu k štatistike v modeloch stochastickej volatility a predstavili sme metódy odhadov parametrov získaných z Monte Carlo simulácií.

Piata kapitola patrí simulačným, Markov Chain Monte Carlo, technikám, ktoré prostredníctvom veľkého množstva simulácií, reprezentujúcich možné realizácie systému, poskytujú odhady jednotlivých parametrov a latentnej premennej. V súčasnosti existuje niekoľko metód, popisujúcich, ako skonštruovať Markovov reťazec a získať vzorky z aposteriórneho rozdelenia. V predkladanej práci sme využili jednu z najnovších metód, a to Hamiltonovskú Monte Carlo metódu, ktorá pridáva do systému parametrov novú dimenziu pomocou parametra hybnosti tak, aby generovala trajektórie vo vektorovom priestore, ktoré sa pohybujú v cieľovom rozdelení. Čitateľa sme oboznámili so základnými princípmi HMC metódy a aj jej NUTS verzie, ktorá poskytuje nadstavbu základného HMC algoritmu, pomocou adaptívnej dĺžky kroku pri aproximácií trajektórií a kritéria ukončenia, ktoré efektívne ukončí iteráciu po prejdení celej trajektórie.

V šiestej kapitole sme zadefinovali kritériá, na základe ktorých sme posudzovali správnosť odhadov, teda ich konvergenciu, nekorelovanosť vzorky a aj chybovosť odhadov. Pre kritérium $\hat{R}$ štatistiky sme sledovali distribúciu každého z odhadovaných parametrov, pričom hodnota kritéria rovná 1 predstavuje ideálny prípad, kedy dva nezávislé reťazce majú pre daný parameter rovnakú distribúciu, čo naznačuje stabilnú konvergenciu v cieľovom rozdelení. Ukazovateľ $\hat{N}_{\text{eff}}$ predstavuje efektívnu veľkosť vzorky, ktorá meria kvalitu získanej vzorky vzhľadom na autokoreláciu, z ktorej sú následne vypočítavané odhady. V ideálnom prípade by mala byť rovná počtu post-warm

up iteráciám, no môže ich aj presiahnuť, čo predstavuje ešte kvalitnejšiu a stabilnejšiu sadu vzoriek a tým pádom nevychýlený odhad. Posledným kritériom analýzy odhadov parametrov z aposteriórneho rozdelenia bolo skúmanie veľkosti chyby MCMC-SE odhadov, ktorá závisí od veľkosti efektívnej vzorky tak, že čím vyššia je kvalita získanej vzorky, tým má parameter nižšiu chybovosť.

Metódy validácie a techniky vzájomného porovnania modelov sme predstavili v siedmej kapitole. Modely stochastickej volatility sú postavené na predpoklade o rozdelení šumu ϵ_t a η_t a tento fakt nás viedol k zostrojeniu testov, kde sme pre vygenerovaný šum η_t a jednotlivé odhady parametrov získané z výsledkov modelovania, spätnou transformáciou získali hodnoty šumu ϵ_t, ktorý sme testovali pomocou A-D testu s rozdelením pravdepodobnosti v závislosti od typu modelu a teda jeho rozdelenia. Súčasne sme sa zamerali aj na testovanie o predpoklade nulovej strednej hodnoty a jednotkovej variancie. Vzhľadom k náhodnosti, ktoré generovanie šumu η_t predstavovalo, sme testy zopakovali 5 000-krát a posudzovali sme sumárne štatistiky, ktoré pozostávali z percentuálnej úspešnosti testov, ktoré dosiahli p-hodnotu vyššiu, než je kritická hladina α; a zároveň, z mediánovej p-hodnoty testov. V závislosti od výsledkov testov, t. j. splnenia predpokladov, sme modely porovnávali na základe informačných kritérií, ktoré posudzovali, ako správne modely popisujú modelované dáta, pričom boli kritické k zložitosti modelov v podobe ich vyššej parametrizácie.

V osmej kapitole sme sa venovali praktickej aplikácií modelovania volatility, pričom sme čitateľovi predstavili syntax a zápis modelu stochastickej volatility do jazyka rozhrania Stan. Zhrnuli sme techniky matematického modelovanie, ktoré zahŕňa získanie a spracovanie dát, modelovanie, validáciu výsledkov a následné testovanie a porovnávanie modelov. Pre zaistenie kvalitných výsledkov sme päťkrát opakovali modelovanie pre každý z modelov, pričom publikované výsledky predstavovali víťaza z pätice kandidátov, ktorého sme vybrali na základe analýzy odhadov parametrov získaných z aposteriórneho rozdelenia. Výsledky modelovania sme vecne zhrnuli do prehľadných tabuliek a grafov, pričom kvôli veľkému rozsahu výsledkov sme do hlavnej časti práce zahrnuli iba modelované časové okná 2022 a 2023 pre výmenný kurz EUR/USD a výsledky pre výmenný kurz HUF/EUR za rok 2023, pričom kompletné výsledky modelovania uvádzame v prílohách.

Naše prvotné očakávania o leptokurtóze log-výnosov sa do značnej miery napl-

nili. V prípade volatilnejšie výmenného kurzu maďarského forintu a eura sme sa stretli v troch zo štyroch modelovaných časových s výsledkami testov, ktoré zamietli modely A a B, ktoré „predstavovali ľahké chvosty“. Najsilnejším z modelov sa ukázal model D, ktorý splnil predpoklady o rozdelení šumu ϵ_t, nulovej strednej hodnote a jednotkovej variancii. Prekvapením bol model C, ktorého počet stupňov voľnosti naznačoval ťažké chvosty, no testy v mnohých prípadoch nevylúčili ani normálne rozdelenie šumu. V prípade výmenného kurzu eura a amerického dolára, ktorý predstavuje dve silné obchodné meny, sme sa v dvoch zo štyroch prípadov stretli so zamietnutím modelov A a B, čo naznačuje, že v prípade stabilnejšieho výmenného kurzu by mohli modely A a B stačiť. Použitie rozdelení pravdepodobnosti s ťažkými chvostami preto nemôžeme jednoznačne potvrdiť, nakoľko rozdelenie šumu závisí od modelovaného výmenného kurzu a aj časového okna, ktoré síce môže zaznamenávať výrazné zmeny v cene, no ak sú postupné, tak sa nepretavia do odľahlých pozorovaní v log-výnosoch.

V prípade výmenného kurzu EUR/USD v časových oknách 2021 a 2023; a výmenného kurzu HUF/EUR v roku 2021, testy v prípade modelov A a B nevyvrátili splnenie predpokladov, preto považujeme za povinnú „výbavu“ matematického modelára portfólio modelov, ktoré je schopné zachytiť akýkoľvek jav v dátach, a to rozdelenia pravdepodobnosti s ľahkými a ťažkými chvostami, či rozdelenia, ktoré dokážu zachytiť špecifický tvar dát.

Z výsledkov môžeme pozorovať, že modely A a B sa takmer nelíšili, čo pripisujeme parametru tvaru ν, ktorého odhady sa pohybovali na hodnotách veľmi blízkych nule. Usudzujeme, že jednotlivé log-výnosy neobsahovali žiadne výrazné zošikmenie, ale naopak, boli koncentrované takmer rovnomerne okolo nuly, čom sa model B pre naše potreby stal zbytočným. Predpokladáme, že model B by mohol byť užitočným nástrojom pri modelovaní volatilnejších aktív, akými sú akcie, akciové indexy či dokonca kryptomeny, ktoré sú známe dobrými a zlými časmi, kedy sú zmeny v cenách výrazné.

Analýza odhadov parametrov aposteriórneho rozdelenia poskytovala hodnoty predstavujúce stabilnú konvergenciu a nezávislú vzorku a tým pádom nevychýlené odhady. Pre parameter perzistencie ϕ^* pozorujeme hodnoty, ktoré sa aj po spätnej transformácii na parameter $\phi = 2\phi^* - 1$ pohybujú v okolí jednotky, čo naznačuje pomerne silnú pozitívnu závislosť na autoregresnom člene h_t, a teda vlastnosť perzistencie a zhlukovania

volatility. Odhady pre parameter σ_η^2 naznačujú nízku mieru šumu, ktorý sa pretaví do odhadu procesu volatility, čo predstavuje želaný efekt, nakoľko naším cieľom je získať najpresnejšie odhady pre proces volatility. Nízka miera šumu v procese volatility sa potom v prípade volatilnejšieho časového okna pretaví do vysokej miery šumu v log-výnosoch, pričom tento šum zachytí člen ϵ_t. Čo sa týka parametrov tvaru ν, ich odhady dosahovali také hodnoty, pri ktorých patria Studentovo t a GED rozdelenia do triedy rozdelení s ťažkými chvostami, v prípade GED rozdelenia a parametra ν z intervalu $(0; 2)$ dokonca do triedy rozdelení s hrubými chvostami.

Hlavný prínos tejto diplomovej práce spočíva v komplexnom a ucelenom spracovaní témy modelov stochastickej volatility, ako metódy, ktorá zahŕňa široké okruhy netriviálnych matematických konceptov. Z teoretického hľadiska autorka priblížila čitateľovi princíp modelovania stochastickej volatility pomocou bayesovského prístupu k štatistike a simulačných MCMC techník, ktoré boli postavené na teórií a vlastnostiach dynamických systémov prevzatých z fyziky prostredníctvom Hamilonovskej Monte Carlo metódy. Z praktického hľadiska na vybraných výmenných kurzoch a časových oknách získala odhady pre volatilitu a úspešne vykonala analýzu využiteľnosti rozdelení pravdepodobnosti s ťažkými chvostami, čím naplnila cieľ diplomového výskumu.

Zoznam použitej literatúry

[1] Baker, R., D.: *A useful family of fat-tailed distributions.* arXiv preprint, 2022. [cit. 20.4.2024] Dostupné na adrese: `https://arxiv.org/pdf/2202.03253.pdf`.

[2] Betancourt, M.: *A conceptual introduction to Hamiltonian Monte Carlo.* arXiv preprint, 2017. [cit. 1.3.2024] Dostupné na adrese: `https://arxiv.org/pdf/1701.02434.pdf`.

[3] Betancourt, M.: *Identifying the Optimal Integration Time in Hamiltonian Monte Carlo.* arXiv preprint, 2016. Dostupné na adrese: `https://arxiv.org/pdf/1601.00225`.

[4] Billio, M., Sartore, D.: Stochastic volatility models: a survey with applications to option pricing and value at risk. *Applied quantitative methods for trading and investment*, 2003, pp. 239–291. [cit. 7.11.2023] Dostupné na adrese: `https://www.greta.it/old/wp/02.10.PDF`.

[5] Black, F., Scholes, M.: The valuation of option contracts and a test of market efficiency. *Journal of Finance*, 1972, vol. 27, no. 2, pp. 416. [cit. 6.11.2023] Dostupné na adrese: `https://www.jstor.org/stable/pdf/2978484.pdf`.

[6] Boháčková, E.: *Metódy a algoritmy pre modely stochastickej volatility.* Diplomová práca, Fakulta matematiky, fyziky a informatiky Univerzity Komenského v Bratislave, 2007. [cit. 10.1.2024] Dostupné na adrese: `http://www.iam.fmph.uniba.sk/studium/efm/diplomovky/2007/bohackova/diplomovka.pdf`.

[7] Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 1986, vol. 31, no. 3, pp. 307–327.

[8] Cirillo, P. [online]: *Risks and Fat Tails.* Conference on Complex Systems, 2020. cit. [21.4.2024] Dostupné na adrese: `https://www.youtube.com/watch?v=iVwAG9q8lk8`.

[9] Delignette-Muller, M. L., Dutang, C.: fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software*, 2015, vol. 64, no. 4, pp. 1–34.

[10] Dias, D., S., Ehlers, R., S.: *Stochastic volatily models using Hamiltonian Monte Carlo methods and Stan.* arXiv preprint, 2017. [cit. 2.1.2024] Dostupné na adrese: `https://arxiv.org/pdf/1712.02326.pdf`.

[11] Embrechts, P., Klüppelberg, C., Mikosch, T.: *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, 1997.

[12] Engle, R. F.: Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica: Journal of the econometric society*, 1982, pp. 987–1007.

[13] Engle, R. F., Patton, A. J.: What good is a volatility model? *Quantitative Finance*, 2001, vol. 1, no. 2. [cit. 6.11.2023] Dostupné na adrese: `https://archive.nyu.edu/bitstream/2451/26881/2/S-DRP-01-03.pdf`.

[14] Frühwirth-Schnatter, S., Wagner, H.: Stochastic Model Specification Search for Gaussian and Partial Non-Gaussian State Space Models. *Journal of Econometrics*2010, vol. 154, no. 1, pp. 85–100.

[15] Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, no. 6, pp. 721–741.

[16] Hastings, W. K.: Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrica*, 1970, vol. 57, no. 1, pp. 97–109.

[17] Hoffman, M., D., Gelman, A.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res*, 2014, vol. 15, no. 1, pp. 1593–1623. [cit. 17.3.2024] Dostupné na adrese: `https://jmlr.org/papers/volume15/hoffman14a/hoffman14a.pdf`.

[18] Hosszejni, D., Kastner, G.: Modeling Univariate and Multivariate Stochastic Volatility in R with stochvol and factorstochvol. *Journal of Statistical Software*, 2021, vol. 100, no. 12, pp. 1-–34.

[19] Kastner, G.: Dealing with Stochastic Volatility in Time Series Using the R Package stochvol. *Journal of Statistical Software*, 2016, vol. 69, no. 5, pp. 1–30.

[cit. 2.1.2024] Dostupné na adrese: https://www.jstatsoft.org/index.php/jss/article/view/v069i05.

[20] Kim, S., Shephard, N., Chib, S.: Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *The review of Economic Studies*, 1998, vol. 65, no. 3, pp. 361–393.

[21] Mentel, G., Brożyna, J.: Historical Data in the Context of Risk Prediction, *International Journal of Business and Social Research*, 2014, vol. 3, pp. 48–60.

[22] Metropolis, N. a kol.: Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 1953, vol. 21, no. 6, pp. 1087–1092.

[23] Neal, R., M.: MCMC using Hamiltonian dynamics, *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC Press, 2011, vol. 2, no. 11. [cit. 13.3.2024] Dostupné na adrese: https://www.mcmchandbook.net/HandbookChapter5.pdf.

[24] R Core Team: *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Viedeň, Rakúsko, 2021. [cit. 11.1.2024] Dostupné na adrese: http://www.R-project.org/.

[25] Serfozo, R.: *Basics of applied stochastic processes*. Springer Science & Business Media, 2009. [cit. 2.4.2024] Dostupné na adrese: http://www.stat.yale.edu/~jtc5/251/readings/Basics%20of%20Applied%20Stochastic%20Processes_Serfozo.pdf#page=109.

[26] Shephard, N., Andersen, T. G.: Stochastic Volatility: Origins and Overview. *Handbook of financial time series*, 2009, pp. 233–254. [cit. 6.11.2023] Dostupné na adrese: https://link.springer.com/chapter/10.1007/978-3-540-71297-8_10.

[27] Stan Development Team: *Stan Modeling Language Users Guide and Reference Manual*, verzia 2.32.3. [cit. 11.1.2024] Dostupné na adrese: https://mc-stan.org.

[28] Stan Development Team [online]: *MCMC Sampling*. [cit. 14.3.2024] Dostupné na adrese: https://mc-stan.org/docs/reference-manual/mcmc.html#ref-RobertsEtAl:1997.

[29] Stan Development Team [online]: *Stochastic volatility models.* [cit. 3.1.2024] Dostupné na adrese: https://mc-stan.org/docs/stan-users-guide/time-series.html#stochastic-volatility-models.

[30] Stan Development Team [online]: *Posterior Analysis.* [cit. 27.2.2024] Dostupné na adrese: https://mc-stan.org/docs/reference-manual/analysis.html.

[31] Stan Development Team [online]: *Have I converged? Convergence checks in Stan.* 2020. [cit. 12.4.2024] Dostupné na adrese: https://www.youtube.com/watch?v=0FdMZwIbJ_4&t=541s.

[32] Stehlíková, B.: *Časové rady.* prednášky, Fakulta matematiky, fyziky a informatiky Univerzity Komenského v Bratislave, Bratislava, 2022.

[33] Stehlíková, B.: *Finančné deriváty.* prednášky, Fakulta matematiky, fyziky a informatiky Univerzity Komenského v Bratislave, Bratislava, 2023.

[34] Suzuki, J.: *WAIC and WBIC with R Stan.* Springer Nature, Singapore, 2023. [cit. 23.2.2024] Dostupné na adrese: https://doi.org/10.1007/978-981-99-3838-4_3.

[35] Szűcs, G.: *Matematické modelovanie v poisťovníctve.* Interné výučbové materiály k predmetu Poistná matematika a Všeobecná teória poistenia. Fakulta matematiky, fyziky a informatiky Univerzity Komenského v Bratislave, Bratislava, 2021.

[36] Tamayo, C., E.: The AR(1) Process. Department of Economics, Rutgers University. [cit. 7.2.2024] Dostupné na adrese: https://econweb.rutgers.edu/ctamayo/teaching/AR(1)_process.pdf.

[37] Taylor, S. J.: Financial returns modelled by the product of two stochastic processes, a study of daily sugar prices 1961-79. *Time Series Analysis: Theory and Practice 1*, 1982, North Holland, Amsterdam, pp. 203–226.

[38] Tierney, L.: Markov chains for exploring posterior distributions. *the Annals of Statistics*, 1994, pp. 1701-1728. [cit. 2.4.2024] Dostupné na adrese: https://projecteuclid.org/journals/annals-of-statistics/volume-22/issue-4/Markov-Chains-for-Exploring-Posterior-Distributions/10.1214/aos/1176325750.full?tab=ArticleLink.

[39] Wang, Y.: *Comparison of stochastic volatility models using integrated information criteria*. Doctoral dissertation, University of Saskatchewan, Canada, 2016. [cit. 10.1.2024] Dostupné na adrese: `https://harvest.usask.ca/server/api/core/bitstreams/c1c8f470-5caf-4d2f-b071-c88c1bfe9a44/content`.

[40] Watanabe, T., Asai, M.: Stochastic volatility models with heavy-tailed distributions: a Bayesian analysis. *Institute for Monetary and Economic Studies*, Bank of Japan, 2001, no. 17. [cit. 4.1.2023] Dostupné na adrese: `https://www.researchgate.net/publication/228568385_Stochastic_volatility_models_with_heavy-tailed_distributions_A_Bayesian_analysis` .

[41] Yahoo, Inc. [online]: Yahoo! Finance. Dostupné na adrese: `https://finance.yahoo.com`.

Príloha A

Výsledky modelovania výmenného kurzu EUR/USD

Tabuľka 11: Odhady pre aposteriórnu strednú hodnotu a štandardnú odchýlku parametrov (v zátvorke) jednotlivých modelov pre výmenný kurz EUR/USD.

výmenný kurz EUR/USD					
časové okno	model	μ	ϕ^*	σ_η^2	ν
rok 2020	model A	−11,047 (0,339)	0,963 (0,023)	0,068 (0,042)	–
	model B	−11,036 (0,373)	0,964 (0,022)	0,066 (0,040)	−0,008 (0,080)
	model C	−10,955 (0,338)	0,964 (0,024)	0,054 (0,040)	8,976 (4,084)
	model D	−11,037 (0,336)	0,962 (0,023)	0,070 (0,045)	2,081 (0,408)
rok 2021	model A	−11,363 (0,148)	0,906 (0,059)	0,052 (0,053)	–
	model B	−11,361 (0,148)	0,901 (0,061)	0,057 (0,061)	0,010 (0,080)
	model C	−11,280 (0,175)	0,912 (0,058)	0,042 (0,041)	7,289 (3,638)
	model D	−11,314 (0,175)	0,908 (0,060)	0,038 (0,039)	1,500 (0,238)
rok 2022	model A	−10,310 (0,241)	0,939 (0,045)	0,051 (0,047)	–
	model B	−10,301 (0,245)	0,935 (0,049)	0,055 (0,051)	−0,009 (0,080)
	model C	−10,218 (0,260)	0,935 (0,051)	0,042 (0,042)	7,922 (3,790)
	model D	−10,261 (0,239)	0,935 (0,047)	0,043 (0,038)	1,655 (0,276)
rok 2023	model A	−10,775 (0,142)	0,886 (0,061)	0,086 (0,075)	–
	model B	−10,775 (0,148)	0,882 (0,062)	0,089 (0,073)	−0,015 (0,080)
	model C	−10,680 (0,163)	0,896 (0,061)	0,051 (0,048)	7,599 (3,715)
	model D	−10,743 (0,155)	0,892 (0,061)	0,068 (0,072)	1,742 (0,330)

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Tabuľka 12: Analýza konvergencie reťazcov pre parametre modelov na základe kritéria $\hat{R}$ štatistiky pre výmenný kurz EUR/USD v priebehu štyroch jednoročných období $2020 - 2023$.

$\hat{R}$ štatistika parametrov modelov pre výmenný kurz EUR/USD					
rok	parameter	model A	model B	model C	model D
2020	μ	0,9999	0,9999	0,9999	0,9999
	ϕ^*	1,0020	0,9999	1,0005	1,0014
	σ_η^2	1,0040	1,0010	1,0018	1,0019
	ν	–	0,9999	0,9999	0,9999
2021	μ	0,9997	0,9999	0,9999	1,0001
	ϕ^*	1,0031	0,9998	1,0007	1,0000
	σ_η^2	1,0077	1,0003	1,0030	1,0003
	ν	–	1,0020	0,9998	0,9997
2022	μ	0,9998	0,9999	0,9998	0,9999
	ϕ^*	1,0031	0,9998	1,0006	0,9999
	σ_η^2	1,0037	1,0001	1,0027	1,0056
	ν	–	1,0005	0,9998	1,0007
2023	μ	0,9997	1,0003	1,0001	1,0007
	ϕ^*	1,0024	1,0018	1,0037	1,0005
	σ_η^2	1,0029	1,0011	1,0066	1,0019
	ν	–	0,9998	1,0003	1,0004

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Tabuľka 13: Analýza konvergencie reťazcov pre parametre modelov na základe efektívnej veľkosti vzorky $\hat{N}_{\text{eff}}$ pre výmenný kurz EUR/USD v priebehu štyroch jednoročných období 2020 – 2023.

$\hat{N}_{\text{eff}}$ pre parametre modelov pre výmenný kurz EUR/USD					
rok	parameter	model A	model B	model C	model D
2020	μ	5901	5926	5825	5685
	ϕ^*	1372	1451	1297	1583
	σ_η^2	931	960	723	882
	ν	–	5954	5501	4846
2021	μ	4849	5001	5835	5294
	ϕ^*	872	932	1299	1466
	σ_η^2	403	424	408	521
	ν	–	5570	5891	5431
2022	μ	5668	5560	5374	5294
	ϕ^*	807	828	1171	1170
	σ_η^2	496	475	482	451
	ν	–	6343	5071	4779
2023	μ	3638	3840	4975	2726
	ϕ^*	892	1007	1105	759
	σ_η^2	490	552	410	334
	ν	–	6116	5655	1227

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Tabuľka 14: Analýza konvergencie reťazcov pre parametre modelov na základe štandardnej chyby MCMC-SE pre výmenný kurz EUR/USD v priebehu štyroch jednoročných období 2020 – 2023.

MCMC-SE pre parametre modelov pre výmenný kurz EUR/USD					
rok	parameter	model A	model B	model C	model D
2020	μ	0,0044	0,0048	0,0044	0,0045
	ϕ^*	0,0006	0,0006	0,0007	0,0006
	σ_η^2	0,0014	0,0013	0,0015	0,0015
	ν	–	0,0010	0,0551	0,0059
2021	μ	0,0021	0,0021	0,0023	0,0024
	ϕ^*	0,0020	0,0020	0,0016	0,0016
	σ_η^2	0,0026	0,0029	0,0020	0,0017
	ν	–	0,0011	0,0474	0,0032
2022	μ	0,0032	0,0033	0,0036	0,0033
	ϕ^*	0,0016	0,0017	0,0015	0,0014
	σ_η^2	0,0021	0,0024	0,0020	0,0018
	ν	–	0,0010	0,0532	0,0040
2023	μ	0,0024	0,0024	0,0023	0,0030
	ϕ^*	0,0020	0,0020	0,0018	0,0022
	σ_η^2	0,0034	0,0031	0,0024	0,0039
	ν	–	0,0010	0,0504	0,0094

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Tabuľka 15: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz EUR/USD v roku 2020

EUR/USD, 2020	percentuálny podiel p-hodnôt väčších než hladina α = 0,05					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	47,64%	100,00%	17,78%	–	–	–
model B	47,32%	100,00%	19,70%	48,54%	100,00%	77,64%
model C	55,98%	100,00%	31,30%	78,28%	100,00%	84,84%
model D	45,02%	100,00%	18,76%	100,00%	76,60%	76,60%

EUR/USD, 2020	mediánová p-hodnota					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	0,0461	0,6807	< 0,0001	–	–	–
model B	0,0452	0,6831	< 0,0001	0,0477	0,6869	0,2707
model C	0,0618	0,6832	0,0013	0,1569	0,6797	0,3467
model D	0,0405	0,6949	< 0,0001	0,7242	0,6972	0,2492

(*zdroj*: vlastné spracovanie)

Tabuľka 16: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz EUR/USD v roku 2021

EUR/USD, 2021	percentuálny podiel p-hodnôt väčších než hladina α = 0,05					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	99,80%	100,00%	92,88%	–	–	–
model B	99,48%	100,00%	91,86%	99,54%	100,00%	98,68%
model C	99,86%	100,00%	92,68%	89,28%	100,00%	85,30%
model D	99,98%	100,00%	96,70%	100,00%	93,36%	93,36%

EUR/USD, 2021	mediánová p-hodnota					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	0,1909	0,9039	0,4682	–	–	–
model B	0,1909	0,8949	0,4421	0,1950	0,8192	0,4790
model C	0,1918	0,9121	0,4213	0,1049	0,8045	0,1776
model D	0,2105	0,9173	0,5279	0,9242	0,8076	0,2576

(*zdroj*: vlastné spracovanie)

Tabuľka 17: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz EUR/USD v roku 2022

EUR/USD, 2022	percentuálny podiel p-hodnôt väčších než hladina α = 0,05					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	93,20%	100,00%	43,36%	–	–	–
model B	94,28%	100,00%	43,18%	94,60%	100,00%	98,02%
model C	98,42%	100,00%	76,22%	98,34%	100,00%	97,02%
model D	98,24%	100,00%	63,38%	100,00%	98,96%	98,96%

EUR/USD, 2022	mediánová p-hodnota					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	0,1608	0,8780	0,0263	–	–	–
model B	0,1628	0,8775	0,0264	0,1665	0,8693	0,5136
model C	0,2039	0,8904	0,2666	0,2499	0,8716	0,4989
model D	0,1994	0,8932	0,1341	0,9763	0,8713	0,5369

(*zdroj*: vlastné spracovanie)

Tabuľka 18: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz EUR/USD v roku 2023

EUR/USD, 2023	percentuálny podiel p-hodnôt väčších než hladina α = 0,05					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	99,92%	100,00%	82,42%	–	–	–
model B	99,92%	100,00%	82,28%	99,92%	100,00%	99,74%
model C	100,00%	100,00%	93,20%	99,32%	100,00%	94,26%
model D	99,96%	100,00%	93,38%	100,00%	98,64%	98,64%

EUR/USD, 2023	mediánová p-hodnota					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	A-D test (norm)	t-test (o str. hodnote)	chí-kvadrát test (o štand. odchýlke)	A-D test (rozdelenie*)	Wikcoxonov test (o str. hodnote)	Leveneov test (o štand.
model A	0,4621	0,8853	0,3158	–	–	–
model B	0,4686	0,8884	0,3140	0,4793	0,9185	0,5578
model C	0,5931	0,9052	0,4315	0,2764	0,9145	0,2745
model D	0,5147	0,8940	0,4813	0,9886	0,9185	0,4936

(*zdroj*: vlastné spracovanie)

Tabuľka 19: Analýza modelov na základe informačných kritérií DIC a WAIC pre výmenný kurz EUR/USD v priebehu štyroch jednoročných období 2020 – 2023.

porovnanie modelov pre výmenný kurz EUR/USD					
rok	model	DIC	poradie	WAIC	poradie
2020	model A	−2103,907	1	−2105,659	1
	model B	−2101,964	3	−2103,651	3
	model C	−2092,705	4	−2098,955	4
	model D	−2103,890	2	−2104,691	2
2021	model A	−2203,850	1	−2203,549	1
	model B	−2201,680	2	−2201,404	3
	model C	−2197,943	4	−2201,129	4
	model D	−2200,538	3	−2202,673	2
2022	model A	−1907,660	1	−1908,120	1
	model B	−1905,569	2	−1905,945	2
	model C	−1899,179	4	−1903,461	4
	model D	−1901,788	3	−1905,174	3
2023	model A	−2040,093	1	−2040,495	1
	model B	−2038,418	2	−2038,701	2
	model C	−2034,023	4	−2037,018	4
	model D	−2035,468	3	−2038,060	3

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Príloha B

Výsledky modelovania výmenného kurzu HUF/EUR

Tabuľka 20: Odhady pre aposteriórnu strednú hodnotu a štandardnú odchýlku parametrov (v zátvorke) jednotlivých modelov pre výmenný kurz HUF/EUR.

výmenný kurz HUF/EUR					
časové okno	model	μ	ϕ^*	σ_η^2	ν
rok 2020	model A	−10,801 (0,253)	0,949 (0,030)	0,077 (0,052)	–
	model B	−10,799 (0,247)	0,949 (0,030)	0,076 (0,052)	0,006 (0,079)
	model C	−10,713 (0,279)	0,949 (0,032)	0,063 (0,049)	9,235 (4,293)
	model D	−10,801 (0,275)	0,948 (0,031)	0,082 (0,060)	2,156 (0,524)
rok 2021	model A	−11,242 (0,158)	0,924 (0,048)	0,036 (0,029)	–
	model B	−11,244 (0,153)	0,922 (0,048)	0,039 (0,030)	0,005 (0,079)
	model C	−11,169 (0,164)	0,918 (0,050)	0,038 (0,031)	11,007 (4,651)
	model D	−11,262 (0,162)	0,925 (0,046)	0,048 (0,039)	2,592 (0,574)
rok 2022	model A	−9,739 (0,194)	0,902 (0,054)	0,158 (0,121)	–
	model B	−9,736 (0,187)	0,898 (0,056)	0,171 (0,128)	0,010 (0,081)
	model C	−9,595 (0,205)	0,914 (0,051)	0,087 (0,081)	5,717 (3,099)
	model D	−9,636 (0,196)	0,912 (0,052)	0,088 (0,078)	1,443 (0,238)
rok 2023	model A	−10,588 (0,285)	0,939 (0,035)	0,143 (0,087)	–
	model B	−10,587 (0,283)	0,937 (0,037)	0,151 (0,097)	0,029 (0,080)
	model C	−10,472 (0,293)	0,953 (0,031)	0,083 (0,065)	6,703 (3,418)
	model D	−10,557 (0,324)	0,942 (0,036)	0,131 (0,098)	1,828 (0,372)

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Tabuľka 21: Analýza konvergencie reťazcov pre parametre modelov na základe kritéria $\hat{R}$ štatistiky pre výmenný kurz HUF/EUR v priebehu štyroch jednoročných období $2020 - 2023$.

$\hat{R}$ štatistika parametrov modelov pre výmenný kurz HUF/EUR					
rok	parameter	model A	model B	model C	model D
2020	μ	0,9999	0,9997	1,0004	0,9997
	ϕ^*	1,0001	1,0007	0,9997	1,0035
	σ_η^2	0,9998	1,0036	1,0003	1,0044
	ν	–	0,9999	0,9999	1,0005
2021	μ	1,0006	0,9998	1,0001	1,9999
	ϕ^*	1,0001	1,0003	1,0002	1,0007
	σ_η^2	1,0086	1,0029	1,0052	1,0002
	ν	–	0,9999	0,9999	0,9999
2022	μ	0,9999	1,0005	1,0002	0,9998
	ϕ^*	1,0012	1,0004	1,0016	1,0023
	σ_η^2	1,0014	1,0008	1,0027	1,0023
	ν	–	0,9999	0,9999	0,9998
2023	μ	1,0001	0,9998	0,9998	1,0004
	ϕ^*	1,0010	1,0003	1,0010	0,9997
	σ_η^2	1,0030	1,0010	1,0016	1,0002
	ν	–	1,0003	0,9997	0,9999

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Tabuľka 22: Analýza konvergencie reťazcov pre parametre modelov na základe efektívnej veľkosti vzorky $\hat{N}_{\text{eff}}$ pre výmenný kurz HUF/EUR v priebehu štyroch jednoročných období 2020 – 2023.

$\hat{N}_{\text{eff}}$ pre parametre modelov pre výmenný kurz HUF/EUR					
rok	parameter	model A	model B	model C	model D
2020	μ	5462	5545	5534	5869
	ϕ^*	1170	1154	1099	1086
	σ_η^2	748	694	623	708
	ν	–	5838	5428	2922
2021	μ	5810	5331	5901	4941
	ϕ^*	1971	2054	2021	1273
	σ_η^2	608	619	532	604
	ν	–	5499	5459	3705
2022	μ	4934	4787	4072	3460
	ϕ^*	869	866	1190	1315
	σ_η^2	583	617	444	487
	ν	–	5894	4029	2430
2023	μ	5339	5951	5735	4765
	ϕ^*	942	981	950	718
	σ_η^2	718	791	627	502
	ν	–	5557	5649	1625

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Tabuľka 23: Analýza konvergencie reťazcov pre parametre modelov na základe štandardnej chyby MCMC-SE pre výmenný kurz HUF/EUR v priebehu štyroch jednoročných období 2020 – 2023.

MCMC-SE pre parametre modelov pre výmenný kurz HUF/EUR					
rok	parameter	model A	model B	model C	model D
2020	μ	0,0035	0,0033	0,0037	0,0036
	ϕ^*	0,0009	0,0009	0,0010	0,0009
	σ_η^2	0,0019	0,0020	0,0020	0,0023
	ν	–	0,0010	0,0583	0,0097
2021	μ	0,0021	0,0021	0,0021	0,0023
	ϕ^*	0,0011	0,0011	0,0011	0,0013
	σ_η^2	0,0012	0,0011	0,0014	0,0016
	ν	–	0,0012	0,0630	0,0094
2022	μ	0,0028	0,0027	0,0032	0,0033
	ϕ^*	0,0018	0,0019	0,0015	0,0014
	σ_η^2	0,0050	0,0052	0,0038	0,0035
	ν	–	0,0011	0,0488	0,0048
2023	μ	0,0039	0,0037	0,0039	0,0047
	ϕ^*	0,0011	0,0012	0,0010	0,0014
	σ_η^2	0,0033	0,0035	0,0026	0,0044
	ν	–	0,0011	0,0455	0,0092

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])

Tabuľka 24: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz HUF/EUR v roku 2020

HUF/EUR, 2020	**percentuálny podiel p-hodnôt väčších než hladina α = 0,05**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	40,76%	100,00%	20,62%	–	–	–
model B	41,36%	100,00%	21,02%	42,22%	100,00%	88,76%
model C	49,82%	100,00%	37,74%	77,58%	100,00%	92,52%
model D	39,76%	100,00%	20,50%	100,00%	87,14%	87,14%

HUF/EUR, 2020	**mediánová p-hodnota**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	0,0366	0,8186	< 0,0001	–	–	–
model B	0,0374	0,8189	< 0,0001	0,0387	0,8776	0,3812
model C	0,0497	0,8306	0,0085	0,1228	0,8841	0,4395
model D	0,0352	0,8197	< 0,0001	0,9191	0,8821	0,3612

(*zdroj*: vlastné spracovanie)

Tabuľka 25: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz HUF/EUR v roku 2021

HUF/EUR, 2021	**percentuálny podiel p-hodnôt väčších než hladina α = 0,05**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	100,00%	100,00%	94,72%	–	–	–
model B	100,00%	100,00%	93,74%	100,00%	100,00%	99,56%
model C	100,00%	100,00%	90,28%	99,94%	100,00%	97,18%
model D	100,00%	100,00%	85,82%	100,00%	99,30%	99,30%

HUF/EUR, 2021	**mediánová p-hodnota**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	0,8908	0,8271	0,4976	–	–	–
model B	0,8874	0,8276	0,4756	0,8918	0,9288	0,5419
model C	0,9121	0,8379	0,4064	0,6905	0,9340	0,4168
model D	0,8339	0,8223	0,3738	0,9741	0,9245	0,5344

(*zdroj*: vlastné spracovanie)

Tabuľka 26: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz HUF/EUR v roku 2022

HUF/EUR, 2022	**percentuálny podiel p-hodnôt väčších než hladina α = 0,05**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	26,86%	100,00%	9,58%	–	–	–
model B	26,06%	100,00%	9,40%	27,16%	100,00%	92,04%
model C	53,80%	100,00%	54,44%	69,24%	100,00%	96,62%
model D	54,24%	100,00%	44,16%	100,00%	98,44%	98,44%

HUF/EUR, 2022	**mediánová p-hodnota**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	0,0261	0,7810	< 0,0001	–	–	–
model B	0,0240	0,7786	< 0,0001	0,0250	0,4856	0,3981
model C	0,0547	0,7906	0,0740	0,0771	0,4892	0,5053
model D	0,0545	0,7910	0,0261	0,8904	0,4903	0,5395

(*zdroj*: vlastné spracovanie)

Tabuľka 27: Prehľad výsledkov štatistických testov pri skúmaní splnenia predpokladov o šume ϵ_t pre výmenný kurz HUF/EUR v roku 2023

HUF/EUR, 2023	**percentuálny podiel p-hodnôt väčších než hladina α = 0,05**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	14,14%	100,00%	10,26%	–	–	–
model B	14,60%	100,00%	9,40%	16,84%	100,00%	76,28%
model C	25,04%	100,00%	37,36%	51,88%	100,00%	86,22%
model D	16,54%	100,00%	15,54%	100,00%	82,98%	82,98%

HUF/EUR, 2023	**mediánová p-hodnota**					
	testy pri nezamietnutej hypotéze o normalite			testy pri zamietnutej hypotéze o normalite		
model	**A-D test (norm)**	**t-test (o str. hodnote)**	**chí-kvadrát test (o štand. odchýlke)**	**A-D test (rozdelenie*)**	**Wikcoxonov test (o str. hodnote)**	**Leveneov test (o štand.**
model A	0,0078	0,8017	< 0,0001	–	–	–
model B	0,0073	0,7967	< 0,0001	0,0091	0,6536	0,2334
model C	0,0174	0,7854	0,0061	0,0540	0,6698	0,3714
model D	0,0091	0,7971	< 0,0001	0,8896	0,6593	0,3095

(*zdroj*: vlastné spracovanie)

Tabuľka 28: Analýza modelov na základe informačných kritérií DIC a WAIC pre výmenný kurz HUF/EUR v priebehu štyroch jednoročných období 2020 – 2023.

porovnanie modelov pre výmenný kurz HUF/EUR					
rok	model	DIC	poradie	WAIC	poradie
2020	model A	−2052,273	2	−2055,424	1
	model B	−2050,626	3	−2053,736	3
	model C	−2039,327	4	−2047,118	4
	model D	−2054,522	1	−2055,200	2
2021	model A	−2174,197	2	−2176,938	2
	model B	−2172,209	3	−2175,015	3
	model C	−2161,234	4	−2170,273	4
	model D	−2186,228	1	−2180,845	1
2022	model A	−1755,671	1	−1754,238	1
	model B	−1753,860	2	−1752,692	2
	model C	−1749,906	3	−1749,136	4
	model D	−1747,654	4	−1749,475	3
2023	model A	−1975,624	1	−1976,280	1
	model B	−1973,678	2	−1974,234	2
	model C	−1967,023	4	−1968,647	4
	model D	−1969,034	3	−1972,068	3

(*zdroj*: vlastné spracovanie pomocou balíčka `rstan` [27] v softvéri R [24])