

SHRNUTÍ OBSAHU PRÁCE

Práce je věnována tzv. multiplikatívním chybovým modelům (MEM). V první kapitole se řešitel snaží zavést GARCH modely a uvést nějaké jejich vlastnosti. Následně jsou definovány MEM modely, a to jak jednorozměrné, tak i vícerozměrné. Určitá část práce je věnována i modelům s chybami, které nabývají nulové hodnoty s kladnou pravděpodobností. V části 2.4 se autor věnuje vícerozměrným MEM modelům, v části 2.5 pak tzv. semi-parametrickému přístupu. V kapitole 3 je následně provedena analýza reálných dat, a to konkrétně časových řad nahlášených škod.

CELKOVÉ HODNOCENÍ PRÁCE

Práce je kompilát několika článků, jejichž vybrané části autor přeložil a přepsal z nich jednotlivé matematické formule. Žádné vysvětlující komentáře, rozpisy nebo zdůvodnění rovností neposkytuje. Text je proto obtížně čitelný a na mnoha místech působí velmi zmateným dojmem. Navíc se z tohoto důvodu práce pohybuje, podle mého názoru, na hranici plagiátu.

Téma práce je podle mého názoru pro diplomovou práci přiměřené a přirozeným způsobem navazuje na látku vyučovanou v rámci povinného předmětu Časové řady.

Vlastní příspěvek. Vlastní příspěvek autora v teoretické části práce (kapitoly 1 a 2) považuji za minimální. Neprovedl téměř žádné logické propojení jednotlivých článků, nedoplnil žádné rozpisy ani komentáře. Značení také není zcela sjednocené. Vlastní je pouze praktická část, kde popsání modely odhadl vždy pro jeden datový soubor.

Matematická úroveň. Autor žádné vlastní matematické výpočty neprovádí. Pokud uvádí nějakou větu, pak bez důkazu a jedná se o přesný překlad z předlohy. Na několika místech se řešitel pokusil o vlastní definice, tam ovšem pak požaduje poněkud nesmyslné předpoklady (viz připomínka 2 níže).

Práce se zdroji: Autor v práci vždy řádně uvádí, ze kterého zdroje čerpá. Následně jeho text dodržuje strukturu předlohového téměř dokonale. U dvou ze článků uvedených v seznamu použité literatury chybí jakékoliv informace o vydání. Uvedeni jsou jen autoři, název a rok, jedná se o Brownless a kol. (2012) a Cipollini a Gallo (2022). Program R ani využití knihovny řádně citovány nejsou.

Formální úroveň. Po vzhledové stránce text vypadá rozumně, až na to, že obsahuje velmi často jednopísmenné předložky na konci řádku a zalamování řádků je občas ad hoc a text není pak dobře zarovnaný. Použité matematické značení není vždy zcela vysvětleno a lze se jen domnívat, co jím autor myslí. Navíc není značení ani mezi jednotlivými částmi (čerpajícími z jiných zdrojů) dostatečně sjednoceno. Dále pak text obsahuje určité množství dalších drobnějších formálních nedostatků týkajících se zápisu matematického textu.

Jazyková úroveň. Po jazykové stránce je text na rozumné úrovni, zvláště mi přijde pouze používání a skloňování slova kopula namísto kopule v části 2.4.2.

Podle mého názoru si na sebe autor uložil příliš velký úkol a snažil se jít příliš do šířky, aniž by dostatečně pochopil a popsal zcela základní modely. Jako největší nedostatek práce

vidím absenci vlastní iniciativy při prezentaci dané látky a formu textu na hranici plagiátu. Níže uvedené připomínky jsou proto pouze doplňující.

Reakce diplomanta: Účelem práce, který je popsán i v samotném zadání, nebylo dokazovat věty a vlastnosti, ale spíše představit různé typy modelů MEM, ukázat možné přístupy ke statistické inferenci a některé z těchto modelů poté zkusit aplikovat na reálná data, případně provést simulační studii. Z mého pohledu jsem tyto cíle naplnil.

V první kapitole se nachází vlastní komentáře, které vysvětlují, proč mají modely právě tento tvar, a definice jsou též formulovány vlastními slovy a nikoliv doslova převzaté z literatury.

Úvod druhé kapitoly se sice poměrně přesně drží článku „New Frontiers For ARCH Models“, avšak článek je zde správně ocitován a do textu jsem přidal i několik vlastních komentářů. Účelem bylo spíše ukázat myšlenku, která stála za zavedením modelů MEM. Definice jsou poté opět formulovány vlastními slovy, ačkoliv plně uznávám, že mají své nedostatky, na které paní doktorka v posudku správně upozorňuje a které se pokusím v odpovědích uvést na pravou míru. Následné odvození věrohodnosti obsahuje oproti článku hustotu, ze které věrohodnost vychází a na konci podkapitoly je též uveden komentář, který tuto část logicky propojuje s další částí ohledně modelů ZA-MEM.

Podkapitoly 2.2 a 2.3 vychází ze článku „Capturing the zero: A new class of zero-augmented distributions and multiplicative error processes“, ale článek je i tentokrát správně citován. Rovnice věrohodnosti jsou ze článku převzaté, ale jejich odvození jsem plně rozuměl a později jsem je v praktické části rozepsal, okomentoval a implementoval v softwaru. Definice modelů ZA-MEM a DZA-MEM znovu převzaté nejsou, i když mají podobné nedostatky jako definice základních modelů MEM (viz odpovědi na otázky v posudku).

Co se týče části 2.4, řídím se skutečně citovaným článkem (až na doplnění vlastností kopul) a uznávám, že jsem měl zřejmě projevit větší vlastní iniciativu při bližším rozpisu uvedených rovnic, ale odvození daných věrohodností (včetně odhadu korelační matice, který se zmiňuje v posudku), rozumím a věřím, že to svými odpovědmi ukážu. Též si myslím, že tvar rovnice střední hodnoty jsem v práci alespoň částečně vysvětlil, i přesto, že např. použití výnosu v rovnici by se jistě dalo popsat lépe. V podkapitolách o semi-parametrických modelech jsem se měl zřejmě pokusit modely formálně definovat a též se od článku více odklonit a pokusit se rovnosti podrobněji rozepsat. Znovu si ale myslím, že nejde pouze o bezmyšlenkovitý přepis původního textu. Je zde několik vysvětlujících komentářů (např. k rovnici pro krátkodobou složku ξ_t), v jádrovém odhadu jsem si v původním článku všiml chyby a soustavu rovnic u zobecněné momentové metody jsem později podrobně rozepsal a implementoval v softwaru.

Přiznávám, že práce obsahuje některé chyby a nedostatky. Na některých místech jsem se jistě měl od článků více odklonit a též jsem chybně definoval některé koncepty. Rád bych ale dodal, že práci jsem celou dobu psal dle svého nejlepšího vědomí a svědomí a rozhodně ne s úmyslem vytvořit jen bezmyšlenkovitý překlad článků, či dokonce plagiát, jak je v posudku naznačeno. Veškeré kroky a jednotlivé části práce jsem vždy podrobně konzultoval s vedoucím, který k finální verzi již neměl další připomínky a plně schválil

její odevzdání. Myslím si proto, že práce i přes zmiňované nedostatky splňuje zadání.

Doplňující otázky

1. **Otázka:** *Kapitola 1 si klade za cíl představení GARCH modelů. Bohužel hned úvodní část je poněkud nesrozumitelná. V celé práci se pracuje se sigma algebrou $\Psi_t = \sigma\{Y_s, s \leq t\}$. V Definicí 1 se specifikuje podmíněná střední hodnota $E[Y_t | \Psi_t] = \mu_t$, ale v Poznámce za defínicí je uvažováno μ_t jako lineární funkce nějakého \mathbf{X}_t , přičemž vztah \mathbf{X}_t k $\{Y_t\}$ (a tedy Ψ_t) je nejasný. Totéž se vyskytuje za Definicí 2. Dále je dosti omezující, že se Defínice 1,2 a 3 omezují na podmíněné normální rozdělení. Toto pojetí také do značné míry znemožňuje využívat vztah mezi MEM a GARCH modely.*

Odpověď: V poznámce pod defínicí k modelům ARCH a GARCH jsem uvedl možnou specifikaci střední hodnoty odezvy v závislosti na náhodných veličinách sdružených do vektoru \mathbf{X}_t a vektoru neznámých parametrů γ . Pokud se jedná o exogenní veličiny, střední hodnota závisí při daném $\mathbf{X}_t = \mathbf{x}_t$ právě na realizaci těchto náhodných veličin, které pozorujeme společně s odezvou $\{Y_t, t \in \mathbb{Z}\}$, nikoliv na Ψ_{t-1} . Uznávám, že poznámka je v práci zbytečná a bylo by lepší ji neuvádět. Častější je v tomto případě defínice μ_t za pomoci např. $AR(p)$ procesu, kde je již závislost na Ψ_{t-1} zřejmá. Předpoklad normálního rozdělení je převzat z původního článku, kde byly modely zavedeny, a jistě by se dal zobecnit, avšak v následujících částech, které se věnují modelům MEM, jsou již vždy uvažovány vlastní předpoklady o distribuci odchylek.

2. **Otázka:** *V práci se v řadě defínic vyskytuje předpoklad, že veličiny $\{\varepsilon_t, t \in \mathbb{Z}\}$ jsou podmíněně nezávislé (viz např. str. 18, kde se pro to zavádí i matematické značení). Prosím o přesné matematické vyjádření toho, co je tímto myšleno.*

Odpověď: Je pravda, že toto je v práci špatně formulováno, protože odchylky do času $(t-1)$ jsou též obsaženy v sigma algebře Ψ_{t-1} , a tedy nemá smysl uvažovat jejich podmíněnou nezávislost, kde bychom touto sigma-algebrou podmiňovali (při podmínění touto sigma algebrou jsou $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ konstanty, tedy jsou nezávislé automaticky). Např. v defínici modelu MEM by tedy mělo být správně uvedeno, že $\{\varepsilon_t, t \in \mathbb{Z}\}$ jsou nezávislé.

3. **Otázka:** *Na str. 8 je zavedena podmíněná logaritmická věrohodnost, používá se značení $l_n(\boldsymbol{\theta} | \Psi_{n-1})$. Prosím o přesné vysvětlení, čím je přesně daná věrohodnost podmíněná a jak byla odvozena.*

Odpověď: Uznávám, že i toto by bylo vhodné formulovat jinak. Zatímco odchylky ε_t jsou nezávislé, Y_t nezávislé nejsou. Za předpokladu exponenciálního rozdělení odchylek platí, že podmíněná hustota Y_t je

$$f_{Y_t | \Psi_{t-1}}(y) = \frac{e^{-y/\mu_t}}{\mu_t}, \quad y \geq 0.$$

V praxi však nemáme k dispozici pozorování Y_t pro $t \leq 0$, je tedy nutné logaritmickou věrohodnost podmínit sigma algebrou Ψ_0 . Potom pro sdruženou podmíněnou hustotu vektoru $(Y_1, \dots, Y_n)^\top$ platí:

$$f_{Y_1, \dots, Y_n | \Psi_0}(y_1, \dots, y_n) = \prod_{t=1}^n f_{Y_t | \Psi_{t-1}}(y_t), \quad y_i \geq 0 \quad \forall i = 1, \dots, n.$$

Oproti formulaci uvedené v práci by tedy bylo vhodnější psát

$$\ell_n(\boldsymbol{\theta}, y_1, \dots, y_n | \Psi_0) = \sum_{t=1}^n \log(f_{Y_t | \Psi_{t-1}}(y_t)),$$

kde $Y_1 = y_1, \dots, Y_n = y_n$ jsou pozorovaná data. Tvar pro exponenciální rozdělení odchylek, který je uveden v práci, se již získá jednoduchým dosazením výše uvedené podmíněné hustoty:

$$\ell_n(\boldsymbol{\theta}, y_1, \dots, y_n | \Psi_0) = \sum_{t=1}^n \left[-\log(\mu_t(\boldsymbol{\theta})) - \frac{y_t}{\mu_t(\boldsymbol{\theta})} \right].$$

4. **Otázka:** Na str. 9 je uvedeno tvrzení týkající se vlastností odhadů lineárního MEM(1,1) modelu. Nesprávně je zde uvedeno, že specifikace μ_t odpovídá GARCH(1,1) modelu (podobná poznámka je i na str. 31). Mezi GARCH a MEM modely jistě vztah je, ale ne takový jako autor uvádí a navíc ne přímo za předpokladů uvažovaných v práci (viz i poznámka 1).

Odpověď: Zde uvedená poznámka o GARCH modelech je myšlena tak, že střední hodnota je modelována způsobem analogickým modelování rozptylu v modelech GARCH, tedy že závisí na zpožděných hodnotách pozorování a též na vlastních zpožděných hodnotách. Podobná formulace ohledně rovnice střední hodnoty se v literatuře nachází velmi často.

5. **Otázka:** Str. 10: Co přesně je myšleno tvrzením, že rozdělení je nespojitě v nule a o jakou hustotu f_X se jedná (vzhledem k čemu)?

Odpověď: Jedná se o směs diskrétního a spojitého rozdělení, které má určitou pravděpodobnostní masu soustředěnou v nule. Hustota je vzhledem k míře $\delta(0) + \lambda^+$. Tedy pracujeme s Diracovou mírou v 0 a s Lebesgueovou mírou na intervalu $(0, \infty)$.

6. **Otázka:** Jak přesně je odvozena log věrohodnost na str. 10? Jedná se o podmíněnou nebo nepodmíněnou věrohodnost? Dále pak v definici 6 se požaduje nějaké podmíněné rozdělení ε_t při daném Ψ_{t-1} . Jakou závislost ε_t a Ψ_{t-1} tento předpoklad implikuje?

Odpověď: Hned na začátku odstavce je uvedeno, že budeme uvažovat náhodnou veličinu X a její nezávislá pozorování. Logaritmická věrohodnost je odvozena z nepodmíněné hustoty f_X , která je uvedena o několik řádků výše. V tomto případě je věrohodnost nepodmíněná, protože ještě neuvažujeme žádný speciální model, pouze posloupnost nezávislých a stejně rozdělených náhodných veličin. Věrohodnost pro samotný model je poté blíže rozepsána v praktické části, což je asi chyba a bylo by vhodnější ji uvést už v části teoretické.

Co se týče odvození věrohodnosti, uvažujme nyní, že odchylky v modelu jsou nezávislé a stejně rozdělené s hustotou

$$f_\varepsilon(x) = p_0 \delta(x) + (1 - p_0) g_\varepsilon(x) \mathbb{1}(x > 0), \quad x \geq 0.$$

Potom analogicky k bodu 3 má Y_t podmíněnou hustotu

$$f_{Y_t | \Psi_{t-1}}(y) = f_\varepsilon\left(\frac{y}{\mu_t}\right) \frac{1}{\mu_t}, \quad y \geq 0.$$

Nyní se opět využije vztahu pro podmíněnou sdruženou hustotu vektoru $(Y_1, \dots, Y_n)^\top$:

$$\begin{aligned} f_{Y_1, \dots, Y_n | \Psi_0}(y_1, \dots, y_n) &= \prod_{t=1}^n f_{Y_t | \Psi_{t-1}}(y_t) \\ &= \prod_{t=1}^n \left[p_0 \delta(y_t) + (1 - p_0) g_\varepsilon \left(\frac{y_t}{\mu_t} \right) \mathbb{1}(y_t > 0) \right] \frac{1}{\mu_t}. \end{aligned}$$

Všechny členy, kde je součin $\delta(y_t)$ a $\mathbb{1}(y_t > 0)$, jsou nulové. Podmíněná logaritmická věrohodnost je tedy tvaru:

$$\begin{aligned} \ell_n(\boldsymbol{\nu}, y_1, \dots, y_n | \Psi_0) &= \sum_{t=1}^n \log(f_{Y_t | \Psi_{t-1}}(y_t)) \\ &= \sum_{t \in I_{n_0}} \log(p_0) + \sum_{t \in I_{n_+}} \log \left((1 - p_0) g_\varepsilon \left(\frac{y_t}{\mu_t(\boldsymbol{\theta})}, \boldsymbol{\nu}_g \right) \right) \\ &\quad - \sum_{t=1}^n \log(\mu_t(\boldsymbol{\theta})) \\ &= n_0 \log(p_0) + n_+ \log(1 - p_0) \\ &\quad + \sum_{t \in I_{n_+}} \log \left(g_\varepsilon \left(\frac{y_t}{\mu_t(\boldsymbol{\theta})}, \boldsymbol{\nu}_g \right) \right) - \sum_{t=1}^n \log(\mu_t(\boldsymbol{\theta})), \end{aligned}$$

kde n_0 je počet nulových pozorování, n_+ je počet nenulových pozorování, I_{n_0} značí množinu indexů nulových pozorování, I_{n_+} je množina indexů nenulových pozorování, $\boldsymbol{\theta}$ je vektor parametrů rovnice střední hodnoty, $\boldsymbol{\nu}_g$ je vektor parametrů hustoty g_ε uvažované na kladné části nosiče odchylek a $\boldsymbol{\nu} = (p_0, \boldsymbol{\theta}^\top, \boldsymbol{\nu}_g^\top)^\top$ je vektor všech parametrů, které odhadujeme.

Co se týče definice modelů ZA-MEM uvedené v práci, je zde stejný problém, který jsem již popisoval u základních modelů MEM. Odchytky $\{\varepsilon_t, t \in \mathbb{Z}\}$ by měly být nezávislé a stejně rozdělené s rozdělením s hustotou f_ε , která je v definici zbytečně uvedena jako podmíněná, i přesto, že na Ψ_{t-1} nijak nezávisí.

7. **Otázka:** Proč se v části 2.3 uvažuje σ -algebra \mathcal{F}_t namísto Ψ_t ?

Odpověď: Máte pravdu, že to v práci není dostatečně zdůvodněno a působí to nekonzistentně. Bylo by lepší používat stále stejné značení.

8. **Otázka Str. 13:** Popis vícerozměrného gama rozdělení formou rozdělení má složitý tvar obsahující integrál mi přijde poněkud úsměvné. Jedná se samozřejmě o překlad textu z předlohy, kde je ale v příloze toto rozdělení skutečně formálně zavedeno (pomocí vhodné konstrukce).

Odpověď: Na tomto místě se mi nezdálo, že by mělo uvedení hustoty tohoto rozdělení nějaký zvláštní přínos. Úvod o vícerozměrném gamma rozdělení je zde uveden spíše pro příklad, blíže se poté věnuji specifikaci za pomoci kopul. Jak uvádíte, explicitní tvar hustoty je v příloze C.1 citovaného článku.

9. **Otázka:** V Definici 10 se připouští jakékoliv vícerozměrné normální rozdělení? Nebo musí toto rozdělení něco splňovat?

Odpověď: V běžných definicích normální kopuly, které jsem dohledal, se žádný další předpoklad neuplatňuje. Je ale pravda, že v dalších částech práce se již pracuje s hustotou, pro kterou je potřeba regularita korelační matice \mathbf{R} . V kontextu práce by proto bylo vhodnější definovat normální kopulu již s tímto předpokladem, případně jej později doplnit.

10. **Otázka:** *Co přesně znamená aproximační znak na str. 15?*

Odpověď: Aproximace vychází ze článku „Multivariate Dispersion Models Generated From Gaussian Copula“, kde je dané rozdělení společně s aproximací zavedeno (rovnice 12 a 13) a který v práci cituji. V samotném článku není přesně specifikováno, co dané aproximace značí a autoři se dále odkazují na článek „Small Dispersion Asymptotics“. V něm se pracuje primárně s exponenciální rodinou rozdělení, do které patří mj. zde použité gamma rozdělení. Přiznávám, že dále jsem toto již nedohledával, protože daný článek pro mě nebyl srozumitelný.

11. **Otázka:** *Model uvedený na str. 15 padne jen tak z nebe, není jasné, proč je jeho struktura právě taktová a jak souvisí veličiny r_t se sigma algebrou Ψ_t , kterou se podmiňuje v celé práci. Prosím tedy o vysvětlení, co v této fázi práce přesně μ_t vyjadřuje. Podobné je to u modelu na str. 19 - tam opět není jasné, jak souvisí R_t s Ψ_t a proč se model zavádí právě takto. Totéž u vícerozměrné verze.*

Odpověď: Opět se jedná pouze o specifický tvar rovnice střední hodnoty, která se podobá způsobu modelování rozptylu v modelech GARCH (modelování vlastními zpožděnými hodnotami i zpožděnými hodnotami pozorování), případně jeho rozšířením, která uvažují pákový efekt (např. v práci uvedený EGARCH). Pokud bychom modelovali např. cenu aktiva, mohl by výnos představovat klasický logaritmický výnos z jeho držení, tedy by veličina $R_t = \log\left(\frac{Y_t}{Y_{t-1}}\right)$ byla endogenní a patřila by do Ψ_t .

12. **Otázka:** *Na str. 16, 6. řádek zdola je chyba.*

Odpověď Ano, je zde chyba v argumentu, kde je místo \mathbf{q}_t uvedeno $\boldsymbol{\varepsilon}_t$.

13. **Otázka:** *Do části 2.4 je věrohodnost značená s argumentem (někde navíc s jakousi podmínkou), od 2.4 pak je zjevně argument vypuštěn. Co přesně jsou zde parametry, není jasné.*

Odpověď: Znovu jde o stejný problém s podmiňováním jako v předchozích případech. Vektorové odchylky by měly být nezávislé a stejně rozdělené s rozdělením, které je specifikováno normální kopulou a marginálními gamma rozděleními. Logaritmickou věrohodnost musíme podmínit analogickým způsobem jako v bodech 3 a 6. Věrohodnost závisí na parametrech rozdělení odchylek (korelační matice normální kopuly a parametry marginálních gamma rozdělení) a též na parametrech rovnice střední hodnoty.

14. **Otázka** *Text na str. 17 a 18 je zjevně jen přepsaný a přeložený a autor nad ním zřejmě asi nepřemýšlel, ale zeptám se na intuici za zavedením odhadu \mathbf{R} na str. 18. Proč lze tento odhad pak dosadit do věrohodnosti?*

Odpověď: Odhad je odvozen z toho, že pokud by \mathbf{q}_i , $i = 1, \dots, n$ byly nezávislé a stejně rozdělené s K -rozměrným normálním rozdělením $N_K(\mathbf{0}, \mathbf{R})$, pak by jejich

hustota byla tvaru

$$f(\mathbf{q}) = (2\pi)^{-K/2} |\mathbf{R}|^{-1/2} \exp\left(-\frac{1}{2} \mathbf{q}^\top \mathbf{R}^{-1} \mathbf{q}\right), \quad \mathbf{q} \in \mathbb{R}^K,$$

a tedy logaritmická věrohodnost by měla tvar

$$\ell_n(\mathbf{q}_1, \dots, \mathbf{q}_n, \mathbf{R}) = -\frac{Kn}{2} \log(2\pi) + \frac{n}{2} |R^{-1}| - \frac{1}{2} \sum_{t=1}^n \mathbf{q}_t^\top \mathbf{R}^{-1} \mathbf{q}_t.$$

Lze si všimnout, že tato věrohodnost závisí na korelační matici úplně stejně jako věrohodnost uvedená na straně 17 v práci. Odtud vychází intuice, že \mathbf{R} lze odhadnout jako výběrovou korelační matici vektorů \mathbf{q}_i , $i = 1, \dots, n$, což je právě odhad

$$\hat{\mathbf{R}} = \mathbf{D}_Q^{-\frac{1}{2}} \mathbf{Q} \mathbf{D}_Q^{-\frac{1}{2}},$$

kde $\mathbf{Q} = \frac{\mathbf{q}^\top \mathbf{q}}{n}$, $\mathbf{D}_Q = (Q_{1,1}, \dots, Q_{K,K})$ a $\mathbf{q} = (\mathbf{q}_1^\top, \dots, \mathbf{q}_n^\top)$, prezentovaný v práci. Po jeho dosazení do logaritmické věrohodnosti již tato věrohodnost závisí pouze na parametrech rovnice střední hodnoty a marginálních gamma rozdělení a ty se díky tomu snáze odhadují (tzv. koncentrovaná věrohodnost). Výsledný odhad však kvůli tomu nemůže být interpretován jako maximálně věrohodný.

15. **Otázka:** V části 2.5. chybí explicitní předpoklady na jednotlivé složky. Co přesně se předpokládá o $\{\xi_t\}$ aby platila rovnost na posledním řádku na str. 18? Nebo to platí vždy?

Odpověď: Aby platilo $\mathbb{E}(Y_t | \Psi_{t-1}) = \mathbb{E}(\mu \tau_t \xi_t \epsilon_t | \Psi_{t-1}) = \mu \tau_t \xi_t$, měl by zde být navíc uveden předpoklad, že ξ_t a τ_t patří do Ψ_{t-1} . Další předpoklady jsou potřeba ke konstrukci predikcí (viz bod 17).

16. **Otázka** V čem přesně se liší momentová metoda popsaná v části 2.5.1 a QML metoda popsaná v části 2.5.2? Jaký je rozdíl mezi $\nabla_{\theta} \xi_t$ a $\partial \xi_t / \partial \theta$? Proč se metoda z 2.5.2 nazývá kvazi-maximálně věrohodná, když se vychází z parametrického předpokladu o rozdělení ϵ_t ?

Máte pravdu, že rovnice pro odvození odhadu se v obou případech shodují, protože

$$\sum_{t=1}^n (\epsilon_t - 1) \mathbf{a}_t = \sum_{t=1}^n \left(\frac{y_t^{(\xi)}}{\xi_t} - 1 \right) \frac{1}{\xi_t} \nabla_{\theta} \xi_t = \sum_{t=1}^n \left(\frac{y_t^{(\xi)} - \xi_t}{\xi_t^2} \right) \nabla_{\theta} \xi_t.$$

Musí se proto dojít i ke shodnému odhadu, což jsem si při psaní práce a interpretaci daného článku bohužel neuvědomil. Jediný rozdíl mezi metodami je v tom, jak se k dané soustavě dojde. Zatímco zobecněná momentová metoda využije vztahu $\mathbb{E}(\epsilon_t - 1) = 0$, v případě metody kvazi-maximální věrohodnosti se využije gamma rozdělení odchylek.

Co se týče otázky ohledně značení,

$$\nabla_{\theta} \xi_t = \begin{pmatrix} \frac{\partial \xi_t}{\partial \alpha} \\ \frac{\partial \xi_t}{\partial \beta} \\ \frac{\partial \xi_t}{\partial \gamma} \end{pmatrix}$$

v tomto kontextu znamená to stejné, jako $\frac{\partial \xi_t}{\partial \theta}$. Metoda se zde nazývá kvazi-maximálně věrohodná, protože se předpokládá, že odchylky mají složitější rozdělení než gamma, ale věrohodnost se přesto sestavuje, jako kdyby toto rozdělení měly. To zde asi není úplně nejlépe formulováno.

17. **Otázka:** Proč predikce v modelu probíhá právě podle vzorců na str. 23? Jak se k těmto vztahům dojde?

Protože τ_t je nízkofrekvenční složka, která by se měla v čase měnit pomalu, volí se jako predikce $\tau_{t+h|t}$ (alespoň pro malá h) jednoduše hodnota $\hat{\tau}_t$. Vzorec pro predikci ξ_t je však v práci bohužel uveden chybně. Máme

$$\xi_t = 1 - \left(\beta_1 + \alpha_1 + \frac{\gamma}{2} \right) + \beta_1 \xi_{t-1} + \alpha_1 Y_{t-1}^{(\xi)} + \gamma_1 Y_{t-1}^{(\xi^-)},$$

kde $Y_t^{(\xi)} = \frac{Y_t}{\mu \tau_t}$ a $Y_t^{(\xi^-)} = \frac{Y_t}{\mu \tau_t} \mathbb{1}(R_t < 0)$. Máme tedy:

$$\xi_{t+1|t} = \mathbb{E}(\xi_{t+1} | \Psi_t) = 1 - \left(\beta_1 + \alpha_1 + \frac{\gamma}{2} \right) + \beta_1 \xi_t + \alpha_1 Y_t^{(\xi)} + \gamma_1 Y_t^{(\xi^-)}.$$

A dále pro $h \geq 2$:

$$\begin{aligned} \xi_{t+h|t} &= \mathbb{E}(\xi_{t+h} | \Psi_t) = 1 - \left(\beta_1 + \alpha_1 + \frac{\gamma}{2} \right) \\ &\quad + \beta_1 \mathbb{E}(\xi_{t+h-1} | \Psi_t) + \alpha_1 \mathbb{E}(Y_{t+h-1}^{(\xi)} | \Psi_t) + \gamma_1 \mathbb{E}(Y_{t+h-1}^{(\xi^-)} | \Psi_t). \end{aligned}$$

Dále můžeme odvodit:

$$\mathbb{E}(Y_{t+h}^{(\xi)} | \Psi_t) = \mathbb{E}[\mathbb{E}(Y_{t+h}^{(\xi)} | \Psi_{t+h-1}) | \Psi_t].$$

Máme $Y_t^{(\xi)} = \frac{Y_t}{\mu \tau_t} = \frac{\mu \tau_t \xi_t \varepsilon_t}{\mu \tau_t} = \xi_t \varepsilon_t$, tedy lze psát

$$\mathbb{E}(Y_{t+h}^{(\xi)} | \Psi_t) = \mathbb{E}[\mathbb{E}(\varepsilon_{t+h} \xi_{t+h} | \Psi_{t+h-1}) | \Psi_t] = \mathbb{E}[\xi_{t+h} \mathbb{E}(\varepsilon_{t+h} | \Psi_{t+h-1}) | \Psi_t] = \xi_{t+h|t}$$

a dále analogicky

$$\mathbb{E}(Y_{t+h}^{(\xi^-)} | \Psi_t) = \mathbb{E}[\mathbb{E}(\xi_{t+h} \varepsilon_{t+h} \mathbb{1}(R_{t+h} < 0) | \Psi_{t+h}) | \Psi_t] = \mathbb{E}[\xi_{t+h} \mathbb{1}(R_{t+h} < 0) | \Psi_t],$$

pokud $R_t = \log\left(\frac{Y_t}{Y_{t-1}}\right)$ je endogenní výnos. Dále bychom zřejmě museli předpokládat nezávislost ξ_t na R_t a též, že R_t mají rozdělení symetrické kolem 0. Je pravda, že tyto předpoklady v práci, jakožto i v citovaném článku chybí. Potom by bylo:

$$\mathbb{E}(Y_{t+h}^{(\xi^-)} | \Psi_t) = \frac{1}{2} \xi_{t+h|t}.$$

Celkem bychom proto měli

$$\xi_{t+h|t} = \mathbb{E}(\xi_{t+h} | \Psi_t) = 1 - \left(\beta_1 + \alpha_1 + \frac{\gamma}{2} \right) + \beta_1 \xi_{t+h-1|t} + \alpha_1 \xi_{t+h-1|t} + \frac{1}{2} \gamma_1 \xi_{t+h-1|t}.$$

Označíme-li $\beta^* = \beta_1 + \alpha_1 + \frac{\gamma}{2}$, pak

$$\xi_{t+h|t} = \mathbb{E}(\xi_{t+h} | \Psi_t) = (1 - \beta_1^*) + \beta_1^* \xi_{t+h-1|t}, \quad h \geq 2.$$

Co se týče celkové predikce, ještě by se zde zřejmě hodil předpoklad nezávislosti ξ_t na τ_t tak, aby

$$\mathbb{E}(Y_{t+h} | \Psi_t) = \mathbb{E}(\mu \tau_{t+h} \xi_{t+h} \varepsilon_{t+h} | \Psi_t) = \mu \mathbb{E}(\tau_{t+h} | \Psi_t) \mathbb{E}(\xi_{t+h} | \Psi_t) = \mu \tau_{t+h|t} \xi_{t+h|t}$$

18. **Otázka:** *Praktická část je sepsaná spíše jednoduchým jazykem s naivními formulacemi o tom, co, jak a kde se spustí v programu R. Autor se neodkazuje na obrázky a tabulky číselně, ale vkládá je přímo na potřebné místo do textu. Rozlišení některých obrázků je nedostatečné. Celkově se jedná spíše o jednodušší aplikaci, jediný zajímavý moment pak shledávám na str. 40-41.*

Odpověď: Každý krok v praktické části jsem velmi podrobně konzultoval s panem docentem Peštou, který její podobu schválil. Nemám proto důvod se domnívat, že je její kvalita nedostatečná.

V souladu se zadáním a cílem jsem práce některé představené postupy aplikoval na reálná data v podobě časové řady nahlášených a vyplacených škod jedné z velkých českých pojišťoven. Získat tato data nebyl úplně jednoduchý proces a již jejich explorativní analýza je z mého pohledu zajímavá. V části ohledně modelů MEM rozšířených v nule jsem udělal chybu, protože jsem si neuvědomil, že po převedení časové řady na čtvrtletní frekvenci zřejmě nemám dostatek pozorování pro odhad modelu se zobecněným F-rozdělením, které samo o sobě obsahuje mnoho parametrů. Parametry modelu jsem však nakonec úspěšně odhadl za předpokladu exponenciálního rozdělení odchylek na kladné části nosiče. Věrohodnosti zde poměrně podrobně komentuji a rozepisuji, z čehož vyplývá, že jejich odvození rozumím.

V části o semiparametrických modelech MEM jsem opět vše podrobně popsal a okomentoval a výsledky tohoto přístupu jsem porovnal se základními modely MEM. I přes určité nedostatky si proto myslím, že praktická část představuje netriviální softwarovou aplikaci představených postupů, která se rozhodně neomezuje jen na použití již připravených balíčků a ukazuje, že jsem tématu porozuměl a jsem s modely schopen pracovat.

Doplňující otázky k praktické části:

- **Otázka** *Na str. 31 tvrdíte, že nemusí být $\alpha + \beta < 1$. Jaké omezení na parametry tedy bylo voleno v numerické optimalizaci a proč?*

Odpověď: Maximalizace proběhla za předpokladu nezápornosti parametrů, aby byla zaručena nezápornost μ_t .

- **Otázka:** *Na str. 32 (a posléze i jinde) je na řadu reziduí použit jednovýběrový t-test. Proč jej zde lze použít?*

Odpověď: Je pravda, že vypočtená rezidua nejsou obecně nezávislá, ale vzhledem k dostatečnému počtu pozorování by to nemělo vadit. Normalita též není v asymptotické verzi t-testu nutná.

- **Otázka:** *Volba parametru zpoždění pro Ljungův Boxův test ad hoc na základě naměřených dat je krajně nevhodná. Navíc mi není jasné, jak přesně byl tento test proveden pro rezidua z MEM modelu. Jaký byl uvažován kritický obor a co přesně je nulová hypotéza?*

Odpověď: Ljungův Boxův test testuje významnost prvních K korelací v časové řadě. Jeho testová statistika je

$$Q = n(n+2) \sum_{t=1}^K \frac{\hat{r}^2(k)}{n-k},$$

kde n je počet pozorování a \hat{r}_k značí odhadnutou autokorelační funkci v bodě zpoždění k . Nulová hypotéza je, že data nejsou korelovaná, resp. že prvních K hodnot autokorelační funkce je nulových. Testová statistika má za nulové hypotézy rozdělení chí-kvadrát s K stupni volnosti. Co se týče volby zpoždění, uznávám, že v literatuře je standardní volit zpoždění spíše na základě počtu pozorování (často se volí např. $K = \log n$). Toto byla moje neznalost, avšak domnívám se, že výsledky v práci to významně neovlivnilo.

- **Otázka** *Z kolika pozorování byl odhadován model v části 3.3? Je to dostatečný počet pro tolik parametrů? Proč nebyla uvažována i jiná struktura modelu?*

Odpověď: Model byl odhadován z 52 pozorování. Uznávám, že to zřejmě není dost pro takový počet parametrů, což jsem si bohužel neuvědomil. Odhady v tomto modelu jsem mnohokrát diskutoval s panem docentem Peštou, který mi radil, ať se spíše držím kvazi-maximálně věrohodných přístupů, které neobsahují příliš velké množství zbytečných parametrů. Jinou strukturu jsem zvažoval, ale nakonec mi nezbyl dostatek času pro její implementaci.

- **Otázka:** *Proč by v části 3.3. mělo být ω omezené? Obecně daný přístup ke konstrukci omezení pro parametry mi přijde hodně nestandardní.*

Odpověď: Parametr ω je nutné omezit, protože jinak při numerické optimalizaci střední hodnota μ_t , která v sobě obsahuje exponenciálu, rychle konverguje do nekonečna a odhady není možné získat. I toto téma jsem opakovaně diskutoval s panem vedoucím, který mi potvrdil, že ve článku, ze kterého jsem čerpal teoretický základ pro tuto část práce, zřejmě chybí nějaká omezení na parametry. Teoretické odvození takových omezení nebylo bohužel v mých silách, a proto jsem (opět po konzultaci s vedoucím) použil omezení bez exaktního teoretického základu.

- **Otázka:** *Jaká jádrová funkce a proč byla použita pro odhad τ_t ? A jak přesně byla spočtena šířka pásma?*

Odpověď: Je použito standardní Gaussovské jádro. Nepletu-li se, není při jádrových odhadech až tolik důležité, jaké konkrétní jádro se zvolí. Výběr šířky pásma zprostředkovala funkce, která používá minimalizaci tzv. „asymptotic mean integrated square error“ (**AMISE**). Blíže jsem se výpočtu šířky pásma nevěnoval, protože jádrové odhady nejsou zaměřením práce.

- **Otázka:** *Vložení použitého kódu přímo do pdf, bez dodatečných komentářů, mi přijde spíše nestandardní.*

Odpověď: Vložení důležitých částí kódu přímo do práce bylo doporučení vedoucího a finální podobu práce mi schválil. Stručné komentáře jsou uvedeny přímo v kódu. Celý kód je potom vložen v SISu jako příloha práce.

ZÁVĚR

Předloženou práci považuji za podprůměrnou. Moje hodnocení bude záviset na průběhu obhajoby a odpovědích na připomínky a otázky položené výše.