

## POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

**Název:** Modely rozložených časových zpoždění

**Autor:** Patrik Dian

### SHRNUTÍ OBSAHU PRÁCE

Práce se zabývá regresními modely rozložených časových zpoždění. Část práce je věnována také auto-regresním modelům rozložených časových zpoždění a intervenční analýze. Poslední kapitola obsahuje ilustraci použití představených metod na reálná data.

### CELKOVÉ HODNOCENÍ PRÁCE

Práce vychází zejména z knihy (Cipra, 2008), přičemž text a struktura bakalářské práce se zejména v prvních dvou kapitolách poměrně těsně drží této předlohy. Kapitola 4.4.2. z této knihy je na několika místech doplněna textem čerpaným ze zdroje (Parker, 2012). Zde také text bakalářské práce sleduje předlohu bez dodatečných vlastních rozpisů či odvození. Dokonce i praktická část 4.1. bakalářské práce je v zásadě převzatá z příkladů z knihy (Cipra, 2008). Zároveň v práci chybí základní úvod do problematiky, předpoklady a vůbec jakákoliv zmínka o metodě odhadu představených modelů (viz připomínka 1 níže).

**Téma práce.** Dané téma je pro bakalářskou práci vhodné, nicméně z velké části je v daném rozsahu vyučováno v rámci magisterského studia.

**Vlastní příspěvek.** Vlastních rozšiřujících komentářů nebo rozpisů je málo. Nějaký vlastní příspěvek lze tedy nalézt pouze v částech 4.2 a 4.3, kde byla provedena samostatná analýza dat, která však není bez výhrad (viz konkrétní připomínky níže) a která je spíše přímočarou aplikací za pomoci existujících metod v programu R.

**Matematická úroveň.** Matematické úpravy jsou spíše jednoduché a jsou přepsány z předloh bez větší invence. Práce neobsahuje žádné matematické tvrzení, takže schopnost formulovat tento typ matematického textu není možné hodnotit.

**Práce se zdroji.** Na několika místech jsou citovány monografie bez specifikace konkrétnější části (strana, sekce apod.), což nepovažuji za velmi vhodné (např. str. 11, 12 aj.). Jinak jsou zdroje řádně citovány a uvedeny v seznamu literatury s výjimkou programu R a použitých knihoven, které v seznamu uvedeny nejsou. V seznamu literatury se míchá čeština a angličtina při uvádění vydání knih a údajů o přístupu k online zdrojům.

**Formální úprava.** Po formální stránce je práce kvalitní. Názvy částí v kapitole 4 by bylo podle mého názoru vhodnější volit jinak, stejně tak názvy částí v Příloze A. Obrázky 4.3, 4.4. a 4.7 by mohly obsahovat nějakou legendu, aby bylo jasné, čemu jednotlivé křivky odpovídají.

### PŘIPOMÍNKY A DOTAZY:

1. Zcela zásadní problém vidím v tom, že v práci není jediná zmínka o tom, jakou metodou se uvedené modely odhadují, za jakých předpokladů a jaké jsou vlastnosti výsledných odhadů. Jedná se o hrubý nedostatek zejména s přihlédnutím k tomu, že praktická část poté odhady obsahuje. Za rovnicí (1.13) je navíc jakýsi komentář (přepsaný z (Cipra, 2008)) k potenciální nekonzistenci odhadů (není jasné jakých), zároveň ale v kapitole 4 se tento postup bez dalšího vysvětlení aplikuje. Prosím tedy o vysvětlení, jakým způsobem se odhaduje model (1.1), jak

model (2.1) za zde uvedených předpokladů a jak byly spočteny odhady z příkladu 5 na str. 17 a zda jsou odhady z tabulky 4.2 rozumné vzhledem k poznámce za (1.13).

Práce začíná hned modelem rozložených časových zpoždění, přičemž chybí zmínka o tom, zda vysvětlující proměnnou chápe autor jako náhodnou veličinu nebo jako konstantu. Sám v tom možná nemá úplně jasno, protože předpoklady na náhodné chyby na str. 3 odpovídají spíše situaci, kdy jsou regresory fixní, na straně 7 pak ale hovoří o vysoké korelaci mezi  $x_t$  a zpožděnými hodnotami této řady.

Na str. 10 se píše o tom, že řád modelu lze určovat pomocí statistických testů, ale už nějak chybí informace o tom, o jaký test se jedná, jak se provádí a jaké jsou jeho předpoklady.

V textu se dále často hovoří o „vlivu jednotlivých vysvětlujících proměnných“ (např. str. 3), nicméně v celé práci jsou uvedeny jen modely s jednou vysvětlující proměnnou, která je v modelu zahrnuta se zpožděním.

2. Na straně 5 je uveden jakýsi pokus o zavedení regresní matice. Jak přesně vypadá regresní matice pro model s  $q$  zpožděními?
3. V kapitole 1.3.1 je představeno několik speciálních modelů rozložených časových zpoždění. V praktické aplikaci je z nich uvažován jen polynomiální model bez krajních omezení, ostatní nejsou nijak zmíněny. Není nijak diskutováno, jak rozhodovat mezi volbou jednotlivých modelů. V příkladu 3 se nijak neřeší, jak volit  $m$ , na straně 10 není nijak diskutována volba  $r$ , řádu polynomu.
4. Str. 9, 3. odstavec: Mluvit v daném kontextu o „konvergenci“ mi přijde poněkud úsměvné.
5. Na straně 10 jsou uvedeny vzorce pro informační kritéria AIC a BIC. Tato kritéria mají uvedený tvar (až na konstantu) jen za dodatečných předpokladů. O jaké předpoklady se jedná? Dále by mě zajímalo, kolik je  $K$  (počet odhadovaných parametrů) v modelu s  $q$  zpožděním?
6. Na str. 10 není vysvětleno, co je  $\hat{y}_t$  a jak se spočte.
7. Na straně 13 se píše, že (2.5) odpovídá po přeznačení modelu (1.1). Jsou pro (2.5) splněny předpoklady uvedené pro (1.1)? Jak se tedy vlastně liší geometrický DL model a ARDL(1,0) model?
8. Příklad 5 je převzat z článku (Box a Tiao, 1975), ale v jejich modelu jsou zavedeny ještě další předpoklady, které v práci nejsou uvedené. Odhady uvedené na str. 17 by pak nejspíš neodpovídaly modelu zavedenému v bakalářské práci.

Na str. 17 se píše, že použijeme (3.5) pro  $\alpha_0 = 1$ , nicméně rovnost (3.5) pro  $\alpha_0 = 1$  neplatí.

#### 9. Komentáře k části 4.1:

- Na str. 19 se píše, že „Durbinův-Watsonův test zamítá hypotézu nekorelovanosti reziduí“. Jaká je přesně nulová hypotéza a alternativa tohoto testu, jak se test provádí, jak vychází  $p$ -hodnota a co to znamená pro odhady modelu (4.3) uvedené v Tabulce 4.2?
- Provádění Durbinova-Watsonova testu v modelech se zpožděnou odezvou se v literatuře nedoporučuje, viz např. (Baltagi, 2008, str. 115).
- Jednotlivé části podkapitoly 4.1 nejsou nijak propojeny, jednotlivé modely nejsou porovnány nebo vzájemně diskutovány. Efekt regresoru mohl být porovnán např. graficky jak bývá v podobných aplikacích běžné. Není jasné, proč je pro polynomiální model voleno  $q = 12$  a pro nestruturovaný model  $q = 4$ . Chybí nějaký závěr z analýzy týkající se toho, jaký model je pro daná data nejvíce vhodný.

## 10. Komentáře k částem 4.2 a 4.3:

- Na straně 22 se píše, že „hodnota testové statistiky Durbinova-Watsonova testu naznačuje pozitivní autokorelovanost reziduí“. O pár řádků výše je uvedeno, že výsledný model je zvolen na základě „signifikantnosti jednotlivých parametrů“. Tato signifikantnost je založena na jakých testech vzhledem k detekované autokorelovanosti?
- Co přesně se předpokládá v (4.8) o reziduální složce a jak byl tedy odhadnut model (4.10)?
- Obecně v praktických příkladech v kapitole 4 v zásadě chybí reziduální diagnostika odhadovaných modelů. Uváděn je pouze Durbinův-Watsonův test, který jednak není vhodný pro ADL modely a také testuje pouze proti autokorelovanosti 1.řádu, což je pro čtvrtletní a měsíční data nedostatečné. V části 4.3. není o diagnostice žádná zmínka.
- Je poněkud matoucí, že v části 4.3. značí proměnná REPO něco jiného než v části 4.2. Bylo vhodnější použít jiné značení.
- Jak přesně došlo k identifikaci ARMA(3,3) modelů z grafy výběrové ACF a PACF na straně 25? A tyto výběrové ACF a PACF byly spočteny jen z dat do roku 2020, nebo z celé řady?
- Na straně 26 za (4.13) nejde o pokles o  $\gamma_0$ , jak uvádí autor, ale o nárůst o  $\gamma_0$ , tedy o pokles o  $-\gamma_0$ .
- Ve (4.12) jsou regresor a operátor ve špatném pořadí, což je nesmysl.
- Očekávala bych, že kromě samotných dat bude přílohou k práci i zdrojový R kód provedené analýzy, bohužel tomu tak ale není.

## ZÁVĚR

Rozhodnutí o uznání této práce jako bakalářské doporučuji založit na průběhu obhajoby a na schopnosti uchazeče prokázat pochopení tématu a schopnost reagovat na výše uvedené připomínky a dotazy.

RNDr. Šárka Hudecová, Ph.D.  
KPMS MFF UK  
V Praze dne 3. června 2022

## Reference

- Baltagi, B.H. (2008): *Econometrics*, 4.vydání. Springer.
- Box, G. E. P. a Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, 70(349), 70–79.
- Cipra T. (2008): *Finanční ekonometrie*. Ekopress.
- Parker, A. J. (2012). Distributed-lag models. Online výukový materiál dostupný na webu [https://www.reed.edu/economics/parker/312/tschapters/S13\\_Ch\\_3.pdf](https://www.reed.edu/economics/parker/312/tschapters/S13_Ch_3.pdf).