

Oponentský posudek diplomové práce
Poissonovská autoregrese
Karolíny Böhmové

V práci se studentka zabývá problematikou časových řad s celočíselnými nezápornými hodnotami. Takové řady se vyskytují například při pozorování denního počtu dopravních nehod v určitém kraji, což je i příklad zkoumaný v práci.

Práce je rozdělena do čtyř částí. Nejprve jsou připomenuty základní pojmy z oblasti procesů a autoregresních procesů, ve druhé části jsou důkladně zkoumány ARCH procesy pro nezáporné celočíselné hodnoty, v krátké třetí kapitole jsou naznačena možná zobecnění a práce končí kapitolou shrnující výsledky simulací a zpracování reálných dat o nehodách.

Autorka v práci podrobně odvozuje hlavní vlastnosti procesů, dále popisuje metody odhadu jejich parametrů a vlastnosti těchto odhadů. Dopustila se některých nepřesností, některé části mohly být rozepsané důkladněji, ale celkově lze říci, že zadání práce bylo splněno a tato práce může posloužit jako základní vodítko pro seznámení s INARCH modely.

Několik poznámek k práci:

- (1) Z několika nejasností a nepřesností v práci vybírám například
 - $b(\zeta) = e^\lambda$ na straně 3
 - V definici striktní stacionarity na straně 3 by mělo být uvedeno, že Y je diskrétní (pro absolutně spojitě procesy je toto splněno triviálně) a nějak omezit indexovou množinu T .
 - V definici podmíněné střední hodnoty na straně 4 je drobná chyba.
 - Na straně 9 se v úvodu píše o řadách s nízkými počty a s vysokými hodnotami. Důležitější než velikost pozorování je spíše množina možných stavů a její mohutnost.
 - Na straně 9 nemůže být pravda, že iid $Y_t \sim \text{Po}(a)$ tvoří bílý šum.
 - Na straně 16 se praví „Nyní si ukážeme ...“, ale ve skutečnosti jde jen o citaci výsledku bez důkazu.
 - Značení \mathbb{Z} na straně 18 bych raději ponechal standardně celým číslem.
 - Diagnostika velmi stručně naznačená v části 2.5.5 není ve čtvrté kapitole vůbec využita, což je škoda a činí tak část 2.5.5 trochu nadbytečnou.
 - U věty 11 na straně 31 by k použitelnosti bodu (iii) pomohlo uvést tvar $I(\theta_0)$. A nemá být v bodě (ii) limitou právě $I(\theta_0)$?
 - Rozšíření uvedené v části 3.2 je opravdu jen stručně naznačené.
 - Formulace posledních dvou odstavců na straně 34 mi přijde poněkud kostrbatá.
 - Na straně 38 dole má být asi uvedeno, že p-hodnota je nižší než 0.05, ne 0.5.
- (2) Jaký je povolený vztah mezi ϵ_t a Y_{t-i} na straně 6?
- (3) Proč na straně 15 platí $E(\epsilon_{t+k} | \mathcal{F}_{t+k-1}) = 0$?
- (4) Proč platí první rovnost na straně 17?
- (5) Proč platí druhá rovnost ve (2.8) na straně 18?
- (6) Na straně 23 nahoře tvrdíte, že Y_t nemusí mít podmíněně Poissonovo rozdělení, ale v důkazu části (e) na následující straně se Poissonovo rozdělení používá. Není to v rozporu?
- (7) Na straně 24: co znamená „není tak eficientní“?
- (8) Na straně 37 vidíme jasně záporné hodnoty autokorelace. Jak to má čtenář chápat v souvislosti s poznámkou na straně 14 dole, že modely INARCH jsou použitelné jen na řady s kladnými autokorelacemi?
- (9) V tabulce 4.4 na straně 38 nejsou vůbec uvedené parametry b_3 a b_4 , ačkoliv v textu jsou v souvislosti s touto tabulkou zmíněné. Jak má čtenář při pohledu na tabulku 4.4 rozumět větě, že „ b_1 , b_3 i b_4 jsou nesignifikantní“?

Práce jistě splnila zadání a přes některé uvedené **drobné** výtky ji považuji za vyhovující standardům diplomové práce. Proto doporučuji její **uznání za diplomovou práci** pro obor Pravděpodobnost, matematická statistika a ekonometrie.

Daniel Hlubinka
v Praze 2.6.2019