

POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

Název: Lineární regresní model s autokorelovanými rezidui

Autor: Ján Kostka

SHRNUTÍ OBSAHU PRÁCE

Tématem práce je lineární regresní model, ve kterém náhodné chyby nemusí splňovat podmínku nekorelovanosti, ale mohou tvořit obecně nestacionární časovou řadu. V práci je tedy představeno několik testů, pomocí nichž lze testovat hypotézu nekorelovanosti nebo stacionarity. Co se děje se standardními odhady a testovacími procedurami v případě porušení těchto předpokladů, není v práci vůbec zmíněno, a chybí tak motivace k celému výkladu. Podle mého názoru se jedná o dosti pokročilé téma a tomu bohužel odpovídá i jeho zpracování, které je dosti povrchní. V práci jsou pouze přepisovány výsledné testové statistiky a chybí jakékoliv sebemenší odvození nebo vlastní iniciativa. Osobně bych preferovala menší množství prezentovaných testů, které by ale byly popsány a vysvětleny důkladněji.

CELKOVÉ HODNOCENÍ PRÁCE

Téma práce. Téma práce je podle mého názoru na bakalářskou práci poměrně obtížné a příliš obsáhlé.

Vlastní příspěvek. Za vlastní příspěvek autora by se dala považovat simulační studie na konci práce a aplikace dané problematiky na reálná data.

Matematická úroveň. Práce bohužel neobsahuje téměř žádné samostatné matematické odvozování. Když od tohoto odhlédneme, tak je vše formulováno matematicky korektně.

Práce se zdroji. Využité zdroje jsou většinou řádně citovány, ale na několika místech v práci se vyskytuje konstatování „dá se ukázat, že“, „je známo, že“ apod. bez další citace konkrétního zdroje.

Formální úprava. Formální úprava práce je na dobré úrovni. Překlepů je minimum.

PŘIPOMÍNKY:

1. V definici 2 jsou definovány slabý a striktní bílý šum. V dalším (definice 5, 7 atd.) se hovoří již pouze o bílém šumu. Není tedy jasné, který šum má autor na mysli.
2. Co přesně je myšleno znakem \approx na str. 11 a na str. 14 a dále?
3. Je známo, že pokud aplikujeme Ljungův-Boxův test (viz str. 15) na odhadnutá rezidua z ARMA modelu, pak není rozdělení asymptoticky χ_m^2 , ale stupně volnosti musí být sníženy o počet odhadnutých parametrů. Tato skutečnost by (vzhledem k navrženému postupu na str.11) rozhodně měla být zmíněna u popisu testu na straně 15.
4. Značení BŠ pro bílý šum a RW pro náhodnou procházku (random walk) mi přijde trochu schizofrenní.
5. Graf dvojic $\{(X_t, Y_t)\}$ se nenazývá korelogram! (Viz strana 18, obr. 4.1. 4.2, 4.3). Korelogram je něco úplně jiného.

6. V části 4.2: Autor na více místech uvádí, že porovnává histogram s teoretickým rozdělením. To je poněkud nepřesná formulace, protože toto rozdělení je pouze asymptotická aproximace skutečného teoretického rozdělení.
7. Na str. 20 je uvedeno, že „rozdělení Durbinové-Watsonovové statistiky není známo“. Rozdělení této statistiky známé je, je popsáno v literatuře, pouze se nejedná o žádné „běžné“ rozdělení a toto rozdělení závisí na matici X . Je smutné, že výsledky simulací k tomuto testu zcela chybí. Kdyby autor použil jiný software, např. R, pak by s aplikací Durbinova-Watsonova testu neměl problém.
8. Není mi jasné, proč se popisky obrázků 4.6 a 4.7 nazývají „hustoty“, když jsou zde znázorněny histogramy.
9. Při popisu síly v simulacích, např. na str. 22-23, by bylo vhodnější vykreslit sílu v závislosti na parametru θ (např. z intervalu $[0.0, 0.2]$) pro oba testy v jednom obrázku, a nikoliv se pouze omezit na několik voleb θ . Stejný obrázek mohl autor uvést i pro AR(1) model.
10. Matice H na str. 24 není v práci definovaná.

DOPLŇUJÍCÍ OTÁZKY

1. Jaké jsou vlastnosti odhadu metodou nejmenších čtverců v modelu s autokorelovanými chybami? Proč je důležitá detekce autokorelovanosti? A proč je speciálně důležité testovat pozitivní autokorelovanost (např. v Durbinově-Watsonově testu)?
2. Co přesně je myšleno nekonečnou sumou v definici 5? Dále prosím o řádné zdůvodnění úprav na str. 4, kde dochází k záměně nekonečné sumy a střední hodnoty — proč je tato záměna možná?
3. V důsledku na str. 7 zamítáme nulovou hypotézu, pokud daná nerovnost platí pro nějaké k nebo pro všechna $k \dots$? A opravdu se jedná o test na hladině α (jak je uváděno na str. 7)? Proč?
4. Proč simulační studie neobsahuje i to, jak se chovají DW, BG a LB testy pro rezidua s jednotkovým kořenem? Použijeme-li totiž návod ze str. 11, pak máme tyto testy použít hned po odhadu lineárního modelu, tj. i v případě, že chyby jsou nestacionární. Stacionarita je pak až obsahem bodu č. 2.

ZÁVĚR

Předloženou práci doporučuji uznat jako bakalářskou.

RNDr. Šárka Hudecová, PhD.

KPMS MFF UK

26. srpna 2019