



Posudek doktorandské disertační práce

Miriam Marušiaková

Tests for multiple changes in linear regression models

Předložená práce je věnována problematice testů pro detekci vícečetných změn v modelech lineární regrese. Tyto testy jsou navrženy jako zobecnění známých výsledků pro situace s nejvíce jednou změnou. Celá disertace se skládá z osmi kapitol a appendixu. Nové teoretické poznatky obsahují především druhá až pátá kapitola. Výsledky simulační studie obsahuje šestá kapitola, po které následuje kapitola s aplikací testů a navržených metod na reálná teplotní data.

Práce se zabývá, jak již bylo řečeno, problematikou detekce změn, což jistě patří mezi důležitá témata, která matematická statistika v současnosti řeší. Zájem o tuto problematiku podtrhuje, že teoretické výsledky jsou dobře aplikovatelné v řadě oblastí a oborů. Disertace se opírá a navazuje na řadu výsledků, které byly publikovány teprve nedávno. Patří sem jednak výsledky publikované školitelkou prof. M. Huškovou, která patří v oblasti change pointu mezi významné autory s velkou mezinárodní reputací. Okruh prací a článků, z kterých M. Marušiaková vychází, je však mnohem širší. Mezi jejich autory patří jména jako J. Bai, P. Perron, C. Kirch, L. Horváth apod., kteří jistě patří ke špičce v oboru. Domnívám se proto, že práce řeší skutečně aktuální téma s mezinárodním dosahem a ohlasem.

Autorce se podařilo úspěšně zobecnit řadu výsledků v oblasti detekce změn v situacích, kdy se uvažuje více změn a nikoliv pouze jedna, jak je obvyklé v publikacích vztahujících se k dané problematice. Důkaz tohoto zobecnění však nebyl přímočarý, ale naopak je často poměrně komplikovaný. Bylo proto nutné úspěšně zvládnout obtížné partie nejen z oblasti "change point", ale i některé kapitoly z problematiky M-odhadů, náhodných procesů a permutačních metod a prostudovat tak velké množství literatury. Z práce bezpochyby vyplývá, že M. Marušiaková dané problematice velmi dobře porozuměla a byla tak schopna existující výsledky využít a tvůrčím způsobem rozšířit na další případy.

Významné výsledky disertace přináší ve všech čtyřech "teoretických" kapitolách. V druhé kapitole autorka uvažuje lineární regresní model s regresory, které jsou buď deterministické nebo stochastické, ale neobsahují trendovou složku. V tomto modelu je odvozeno limitní rozdělení testové statistiky pro rozhodnutí, že nedošlo v modelu k žádné změně proti alternativě, že změn bylo k , kde k je neznámé pevné číslo. Tento výsledek je posléze rozšířen i na situaci, kdy je testována alternativa, že počet změn je neznámý (ale je jich nejvýše M). Jsou zde též odvozeny výsledky, které umožňují použít permutačního principu k aproximaci kritických hodnot testové statistiky. Všechny tyto výsledky jsou následně uvažovány a dokázány v autoregresním modelu (kapitola 3).

Čtvrtá kapitola se zabývá opět lineárním regresním modelem, nyní ale za předpokladu, že regresory mají trend. Autorka řeší podobnou problematiku (testy, permutační princip) jako

v předchozích kapitolách, ale na rozdíl od nich se zaměřuje na techniky, které vychází z M-odhadů a zobecněných M-typů testových statistik. Podařilo se jí dokázat řadu zajímavých výsledků - viz věty 4.4.1 - 4.4.4 a 4.5.1. Pátá kapitola řeší problematiku v modelu polohy za situace, že chyby jsou závislé veličiny. Pochopitelně tato závislost přináší řadu komplikací, přesto se autorce podařilo dokázat několik pěkných výsledků.

Získané teoretické výsledky mají předpokládaný tvar, nicméně jejich důkazy jsou technicky velmi náročné a jsou založeny často na několika dílčích tvrzeních. Řada vět, tvrzení a pojmů, která byla využita v důkazech, je uvedena pro pohodlí čtenáře v appendixu. V důkazech jsem nenalezl žádná problematická místa a domnívám se, že jsou provedeny korektně. Našel jsem pouze několik velmi drobných opomenutí (není např. deklarováno, že použitý symbol je určitá konstanta, je uveden špatný odkaz, chybějí meze integrálu), které však nemají žádný podstatnější význam v logice výkladu. V rámci obhajoby proto není třeba se jimi zabývat. Pokud by autorka tyto připomínky využila např. při publikaci, tak ji jejich seznam včetně nalezených překlepů mohu předat.

M. Marušiaková věnovala zpracování disertace velkou pozornost, práce se velmi dobře čte, je zde minimum překlepů. Pozitivně hodnotím i to, že je v práci nejen dokazováno, ale i vysvětlováno, tj. zdařilým způsobem je komentována motivace a filozofie jednotlivých kroků a pojmů. V tomto ohledu bych měl snad jedinou výtku a to, že by bylo možné více vysvětlit, proč se používá a co znamená pojem α mixing podmínky v kontextu modelu se závislými chybami v kapitole 6. Práce je psána v angličtině, která podle mého mínění zcela odpovídá odborné angličtině v časopisech s tématiky blízkými články. I po formální a grafické stránce působí velmi dobře.

Pečlivě je zpracována i kapitola věnovaná simulacím. Návrh experimentu, jednotlivé kroky i dosažené výsledky jsou vhodně komentovány. Osobně mne překvapil dost rozdílný výsledek pro permutační verzi L_2 testové statistiky týkající se dat generovaných z Laplaceova rozdělení pro rozsahy $n = 120$ a $n = 180$ - viz obrázek 6.3.2 (a) a (c). Rozdíl rozsahů mi nepřipadá zas až tak výrazně rozdílný, aby se asymptotika takto projevila. Rád bych se proto autorky v rámci obhajoby zeptal, zda-li nemá podobnou zkušenost i s jinými rozděleními s "těžšími" chvosty nebo nezkoušela simulace i s jinými rozsahy.

Závěrem konstatuji, že autorka touto prací prokázala hlubokou znalost tématu a schopnost samostatné tvůrčí vědecké práce. Domnívám se, že předložená práce splňuje všechny požadavky kladené na práci disertační, a proto ji doporučuji přijmout k obhajobě.

Liberec, 3. listopadu 2009

