

Posudek dizertační práce

Název dizertační práce: **Tests of statistical hypotheses in measurement error models**

Autor: Mgr. Radim Navrátil

Dizertační práce Mgr. Radima Navrátila se týká testů a odhadů založených na pořadích v regresních modelech. Speciální pozornost je věnována pořadovým procedurám v modelech s chybami měření (v angličtině se používá termín "measurement error models"). Základní otázkou, zda výsledky známé pro modely bez chyb v měřeních zůstanou v platnosti i pro model s chybami v měření, popř. jak je třeba procedury modifikovat. V minulosti byla publikována řada prací pro regresní modely s chybou měření, typicky se jednalo o postupy založené na odhadech metodou nejmenších čtverců. Pokud se týče postupů založených na pořadí pro tuto situaci bylo publikováno jen několik článků zaměřujících se na jednodušší speciální situace. Předložená dizertace je v tomto směru velkým přínosem - řeší úspěšně dosud otevřené problémy a zároveň poskytuje ucelený pohled na pořadové postupy pro modely s chybami měření i bez nich. Teoretické výsledky jsou účelně doplněny numerickými studiemi.

Práce je rozdělena do 6 kapitol. První slouží jako úvod do problematiky, podává přehled některých klasických výsledků o pořadových testech a odhadech a zároveň obsahuje přehled výsledků dosud dosažených v modelech s chybami měření. Autor také upozorňuje na různá úskalí. Teoretické výsledky jsou účelně doplněny numerickými studiemi.

V druhé kapitole jsou navrženy nové pořadové testy pro regresní modely bez i s chybami v měření, je studováno jejich limitní chování pro oba případy. Autor přitom vychází z práce Koula (2002) o odhadech založených na empirických pořadových procesech (pro modely bez chyb v měření). Zkonstruován a studován je test o absolutním členu v lineárním modelu s chybami v měření.

Další kapitola se zaměřuje odhady založené na pořadích v modelech s chybami měření. Je zde odvozeno limitní chování odhadu s podrobným studiem vychýlení tohoto odhadu. Doplněno užitečnou simulační studií se zaměřením na vychýlení. Výrazné prohloubení a doplnění výsledků publikovaných v Saleh et al (2012).

Čtvrtá kapitola se týká pořadových testů v heteroskedastických lineárních modelech, kde heteroskedasticita je uvažována jako speciální model s chybou v měření. Testy se týkají jak hodnot regresního parametru β , tak testu homogenity proti alternativě lokální heteroskedasticity. Opět jsou odvozeny limitní vlastnosti za platnosti nulové hypotézy a lokální (kontiguitní) alternativy a také prezentována simulační studie i aplikace na reálná data.

Pátá a šestá kapitola se týkají modelů polohy s různými typy chyb měření (aditivní chyby v měření, s aditivními symetrickými a asymetrickými chybami v měření). Pátá kapitola se zaměřuje na testy symetrie s aditivními chybami v měření, s aditivními symetrickými a asymetrickými chybami v měření, zatímco šestá se týká odhadů založených na pořadí pro parametr polohy. Podobně jako v předchozích kapitolách autor studuje limitní chování studovaných testů a odhadů pro situace

bez chyb v měření i s nimi. Teoretické výsledky jsou doplněny vhodnými ilustračními simulačními studiemi i praktickou aplikací. V závěru autor shrnuje dosažené výsledky a formuluje otevřené problémy, které je třeba ještě řešit.

Dizertace Mgr. Radima Navrátila na jedné straně přináší mnoho nových zajímavých a užitečných výsledků z oblasti pořadových testů a odhadů založených na pořadích pro regresní modely s chybami v měření, jednak podává ucelené pojednání statistických procedurách založených na pořadích pro regresní modely. Zásadní nové výsledky pro regresní modely jsou obsaženy v kapitolách 2-4, zatímco kapitoly 5 a 6 jsou výsledky pro modely polohy. Každé kapitola obsahuje motivační úvod, popis procedury, odvození limitního chování, simulační studie ilustrující chování studovaných postupů při konečném rozsahu dat a většinou též vhodnou aplikaci na reálná data. Základní nové teoretické výsledky jsou asymptotické (celkový počet pozorování roste nade všechny meze). K jejich odvození autor využívá řadu metod používaných při sledování limitních vlastností pořadových procedur, v řadě případů bylo třeba odvodit netriviální zobecnění. Metody důkazů jsou založené na pokročilých partiích teorie pravděpodobnosti a asymptotické statistiky. Autor do podstaty metod pronikl a dokázal na ně navázat. Navrhl též některé nové testy a odhady. Důležitou součástí práce jsou vhodně zvolené simulace a aplikace na reálná data, které výborně doplňují teoretické výsledky.

K práci mám několik dotazů k uvažovaným předpokladům.

- Základní uvažovaný regresní model má tvar:

$$Y_i = \beta_0 + \mathbf{x}_i^T \boldsymbol{\beta} + e_i$$

$$\mathbf{w}_i = \mathbf{x}_i + \mathbf{v}_i, \quad i = 1, \dots, n,$$

kde na str. 10 je předpoklad, že $\mathbf{v}_1, \dots, \mathbf{v}_n$ jsou nezávislé stejně rozdělené náhodné veličiny nezávislé na e_1, \dots, e_n . Je toto základní předpoklad pro celou práci, pokud je uvažován tento model? V některých částech je tento předpoklad zopakován, někde nikoli. Není to uváděno ani v předpokladech tvrzení. Prosím o vyjasnění. Další otázka je, zda $\mathbf{v}_1, \dots, \mathbf{v}_n$ mohou být i nenáhodné a jak je to s náhodností či nenáhodností $\mathbf{x}_1, \dots, \mathbf{x}_n$. Další otázka je, zda předpoklad, že $\mathbf{v}_1, \dots, \mathbf{v}_n$ jsou nezávislé stejně rozdělení nemůže být nahrazen předpokladem slabším. Str. 34, předpoklad V.3 je zbytečný, pokud \mathbf{x}_i jsou nenáhodné.

- Str. 35, věta 3.1 je skutečně nutné předpokládat, že matice \mathbf{C} je pozitivně definitní? Str. 34, V.2, z předpokladu na str. 33 o $\mathbf{v}_1, \dots, \mathbf{v}_n$ a konečnosti momentů of plyne existence matice \mathbf{C} .

- Str.35, věta 3.1, proč je výchýlení vždy nekladné?

Smyslem těchto dotazů je jen malé upřesnění. V práci jsem našla jen minimální počet překlepů (např. str. 25, (2.5)).

Prezentace výsledků je nadstandarně vysoká. Autor kromě vědecké invence prokázal i pedagogické vlohy. Práce je dobře strukturovaná, dobře se čte, je psána s nadhledem, pěkně vysvětlovány postupy, nechybí interpretace teoretických výsledků, které jsou doplněny řadou motivačních příkladů i numerickými studiemi.

Dizertace Mgr. Radima Navrátila obsahuje mnoho nových zajímavých a užitečných výsledků z oblasti pořadových postupů pro regresní modely s chybami měření. Prokázal, že je schopen samostatné a tvořivé vědecké práce.

Domnívám se, že práce splňuje požadavky kladené na dizertační práci ve studijním oboru Pravděpodobnost a matematická statistika a doporučuji ji za ni uznat.

Prof. RNDr. Marie Hušková, DrSc., oponentka
KPMS MFF UK

2.8.2014