

Oponentský posudek disertační práce:

MATÚŠ MACIAK: FLEXIBILITY, ROBUSTNESS AND DISCONTINUITY
IN NONPARAMETRIC REGRESSION APPROACHES

Předložená práce se zabývá lokálně polynomickými robustními odhady. Původní výsledky spočívají v odvození asymptotického rozdělení těchto odhadů pro závislá data (část 2.4) a v jejich využití pro modely bodu změny (kapitola 3). Čtvrtá kapitola se zabývá aproximací kritických hodnot metodou bootstrap pro nezávislá i závislá pozorování. V páté kapitole jsou popsány metody volby parametru h_N (bandwidth) a odhady funkce měřítka (scale function). Navržené postupy jsou vyzkoušeny na simulovaných i skutečných datech v šesté kapitole.

Výsledky práce a aktuálnost tématu: Robustní neparametrická regrese a její teoretické vlastnosti jsou námětem značného množství starších i novějších odborných prací: konzistence robustních neparametrických regresních odhadů (lokálně konstantních) pro závislá data je zkoumána například v [Boente, Fraiman: Robust nonparametric regression estimation for dependent observations, *Ann. Statist.* 17, 1989, 1242–1256] nebo ve [Wang, Yu: Robust nonparametric regression with simultaneous scale curve estimation based on dependent observations, *Acta Math. Sci. Ser. A Chin. Ed.*, 30, 2010, 835–847]. Ve druhém zmíněném článku je pro lokálně konstantní robustní odhady odvozena i asymptotická normalita. Robustními lokálně polynomickými odhady se zabývá například disertace [Werthenbach: Untersuchung einer Klasse robuster nichtparametrischer Regressionschätzer und Entwicklung von Varianzschätzverfahren, TU Darmstadt, 1999, Logos Verlag Berlin], kde však byla odvozena konzistence a asymptotická normalita těchto odhadů pouze pro nezávislé náhodné chyby.

Odvození asymptotické normality robustních lokálně polynomických odhadů v předložené práci je tedy aktuální a originální výsledek. Stejně tak jsou původní i výsledky týkající se bodu změny a bootstrapu ve třetí a ve čtvrté kapitole, přestože zde je model pro závislá data opuštěn a autor pracuje již pouze s nezávislými pozorováními (s výjimkou sekce 4.3).

Použité metody a postupy: Použité metody jsou přiměřené zkoumaným problémům. Důkazy jsou většinou založené na podrobném zkoumání řádů konvergence (v pravděpodobnosti) jednotlivých odhadů nebo jejich aproximací a na použití vhodných verzí centrální limitní věty.

Důkazy jsou rozepsány dostatečně podrobně a idea důkazu je vždy jasně popsána. Přesto jsem se však na několika málo místech nedokázal v textu zcela zorientovat a prosím o podrobnější vysvětlení:

Str. 5 Jak se liší „As.Bias“ ve větě 1.1 od „Bias“ ve větě 1.2? Stejně symboly se objevují i ve větách 2.3 a 2.5 a ve větách 2.10 a 2.11.

Str. 15, A7: Může skutečně být $h_N \approx N^{-1/(1+2p)}$?

Str. 29, poznámka pod čarou: Co znamená „mostly really“?

Str. 30, (2.32): Odkud vyplývá řád konvergence $o_P(h_N^{p+1})$? Co přesně se myslí tím, že rozdíly $\hat{\beta}_x^0 - \beta_x^0$ a $H_N^{-1}(\hat{\beta}_X - \beta_x)$ jsou „ekvivalentní“ (viz také třetí odstavec na str. 20)?

Str. 32, (2.35): Z čeho plyne řád konvergence $o_P(h_N^{p+1})$? Jaká je délka vektoru μ ?

Str. 54 Není mi jasné, jak jsou ověřeny podmínky (i) a (ii) věty 2.12. Zřejmě $\sigma_N^2 \rightarrow 0$ a v důkazu by tedy měl být vyšetřen také podíl $\sum E\xi_{Ni}^2/\sigma_N^2$ (nestačí pouze $\sum E\xi_{Ni}^2$).

Str. 58, věta 3.1 Co přesně označuje symbol $m(X_0)$ (je např. funkce $m(\cdot)$ spojitá zleva nebo zprava)? Nebylo by jednodušší a přehlednější psát $1/2$ místo $e_0^T \mu_0$?

Str. 60, věta 3.2 Jak je zde definován „Bias“?

Str. 62, důkaz věty 3.3 V důkazu by mělo být navíc ověřeno, že vychýlení (bias) obou odhadů je stejné.

Str. 74, důkaz lemma 10, poslední věta Není mi úplně jasné, jak se důkaz dokončí. Postup asi nemůže být úplně stejný jako při důkazu lemma 3, kde se dokazovalo, že podobný rozdíl je malý (zatímco lemma 10 říká, že „rozdíl je s pravděpodobností 1 větší než konstanta K^* “).

Str. 76–77 Proč plyne z „malé Mallowsovy vzdálenosti“ tvrzení věty 4.2?

Str. 81, poslední vzorec Nemá zde být \sqrt{n} místo n ?

Str. 90, simulace Jak byly zvoleny parametry k pro Huberovu a Tukeyho funkci? Jak byl zvolen parametr h_N pro jádrovou funkci pro různé hodnoty N ?

Str. 111–112 Proč zde nebyl použit postup z kapitoly 3.3 (bod změny s neznámou polohou)? Použití testu pro předem známý bod změny na bod s maximální hodnotou testové statistiky není příliš vhodné.

Kvalita formálního zpracování: Na velkém množství různých rámečků, čar, textů v záhlaví a zápatí stránky je na první pohled vidět, že autor věnoval formální úpravě textu velkou pozornost. V předložené práci jsem našel pouze několik zbytečných překlepů a formálních opomenutí:

Str. 76 Několikrát Mallow místo Mallows.

Tabulka 5.1 Tabulka je chybně označena jako obrázek. Mrzí mě překlep „Tuckey“ místo „Tukey“ (zvláště, když se tento překlep opakuje ve tabulkách v celé šesté kapitole a autor se zúčastnil konference ICORS2010, kde byla Johnu Tukeymu věnována celá vzpomínková sekce).

Sekce 6.2.1 Autor soustavně uvádí „Squallus cephalus“ místo „Squalius cephalus“. Je délka jelce opravdu maximálně 8cm?

Bibliography Citace „H.Rue“ a „J.Fan“ jsou špatně naformátované a díky tomu se špatně hledají. Místo Fitzenberger (1997) má být Fitzenberger (1998).

I přes výše uvedené drobné připomínky se mi předložená disertační práce líbila a považuji ji za přínosnou.

Závěr: Předložená práce obsahuje aktuální a původní vědecké výsledky a jednoznačně prokazuje předpoklady autora k samostatné vědecké práci. Předložená práce splňuje požadavky kladené na disertační práci a proto ji doporučuji jako disertační práci uznat.

RNDr. Zdeněk Hlávka, Ph.D.