

## Oponentský posudek disertační práce

„*M-Estimations in Nonlinear Regression for Longitudinal Data*“

Mgr. M. Orsákové

Práce je věnována jednomu z modelů pro (nelineární) regresi se vstupy, které se vyvíjejí s časem, k dispozici máme opakování měření v náhodných časech, tyto časy měření se vyskytují s nějakou intenzitou. Z jiného pohledu jde o jedno ze schémat pozorování tzv. longitudinálních dat. Výsledná (závislá, výstupní) veličina je tedy dána dvěma funkcemi, střední hodnotou a rozptylem, do obou vstupuje čas i vyvíjející se kovariáty (predikovatelný proces kovariát), střední hodnota je známá funkce s neznámým parametrem, rozptyl je v první části známá funkce, v další části je pak uvažován její odhad. Tento model od 90. let podrobně studovali T. Scheike a T. Martinussen, práce prvého z r. 1994 byla také hlavní inspirací autorky.

Disertační práce začíná připomenutím pojmu intenzity, čítacího procesu, vztahu s martingaly. Pak se dostaváme k hlavní části práce, v které se modifikují výsledky ze Scheike (1994) pro případ M-odhadu neznámých parametrů regresní funkce. V další části je odhadování parametrů i variance (jako funkce) předvedeno na 2 simulovaných studiích, a jsou porovnány různé typy M-odhadů. Dále (již jako příloha) následuje zhuštěný přehled definic a tvrzení z literatury o martingalech, čítacích a značených bodových procesech relevantních pro teorii a metody uvedené v části 2. Je připomenuta i metoda (Scheike a Zhang) pro neparametrické odhadování regresní funkce ve stejném modelu. K dispozici jsou také S-skripty pro simulační studie vytvořené autorkou.

Protože hlavní částí je kapitola 2, budu se zabývat hlavní jí. Je pravda, že rozložení kapitoly sleduje převážně práci Scheike (1994), někdy až s příliš totožnými slovními obraty, přináší však doplnění právě v rozšíření předpokladů a výsledků pro M-odhady. Jaké předpoklady musely být přidány, je rozebráno v diskusi v části 2.3. Proto lze říci, že práce přináší nové výsledky (s již zmíněnou výhradou někdy až příliš stejných formulací s inspirující prací). Je vidět (jak se i říká v úvodu), že se autorka tematikou zabývala důsledně a do hloubky.

Dále bych zmínil některé své drobné postřehy a dotazy:

1. Ve větě 2.1 se mimo jiné chce dokázat (2.20), ale to se v této formě nedokáže, výsledek (poslední řádek v důkazu) by asi měl být „omezeno v pravděpodobnosti“ (ne „limita = 0“). I tak to stačí k existenci konzistentního odhadu (ne k jednoznačnosti).

2. Obdoba výsledku (2.20) je vztah (2.28) ve větě 2.4, tam je tvrzení již opatrnější, dosazuje se konzistenční odhad – pak je to zřejmě v pořádku. Všimněme si ale, že poslední výraz důkazu věty 2.4 obsahuje skutečně  $O_p(1)$ . Takže by bylo třeba to uvést na pravou míru.
3. V (2.19) by asi mělo být – (tj. minus).
4. Další poznámky jsou spíš terminologické:

Skutečně se říká „quadratic predictable variation“, nebo jen „predictable variation“? Scheike používá tradičnější „variance process“.

Na str. 16 pod (2.22) „... is a martingale with a compensator...“, asi se to myslí jinak, martingal se již nekompenzuje.

Na str. 14 dole  $o_p$  místo  $O_p$  zřejmě.

Několikrát záměna  $S$  a  $s$  (ve (2.3) například).

Pod tím v (2.8) je nedefinovaná podmínka  $H_s$ .

Práce je psaná anglicky na dost slušné jazykové úrovni, ale přeci jen se vyskytují drobné chyby, práce vyžadovala pečlivější redakci (po „is“ chybí člen, např. pod 2.5, po „let“ není infinitiv – začátek předpokladu (B), i další).

Přes uvedené výhrady (které jsou ostatně jen dílčí) je z práce vidět, že autorka zná dobře zkoumanou problematiku, orientuje se v literatuře i v souvislostech a má schopnosti téma i dále teoreticky a metodologicky rozvíjet. Práce přináší nové výsledky do oblasti robustních odhadů parametrů, a to jak teoretické, tak praktické dík simulačním porovnáním různých typů M-odhadů. Soudím, že práce prokazuje dobré předpoklady autorky k samostatné vědecké práci. Doporučuji tedy tuto disertaci k obhajobě jako poklad k získání titulu PhD.

V Praze 19.7.2006