



**Posudek doktorandské disertační práce**

*Radky Sabolové*

## Statistical inference based on saddlepoint approximations

Předložená práce je věnována aplikaci metod založených na aproximacích pomocí sedlového bodu. Techniky sedlového bodu jsou především rozvíjeny v problematice aproximací hustot a testů pro kvantilovou regresi.

Problematika odhadu a testování hypotéz patří mezi klasická témata, která matematická statistika řeší. Je obvyklé, že je při tom využíváno asymptotického rozdělení odhadů, respektive testových statistik. To se však v řadě praktických situací, kdy je k dispozici málo pozorování, ukazuje být nevhodné. Jedním z možných řešení, jak situaci zlepšit, je využít aproximací na základě metod sedlového bodu. Z tohoto pohledu se práce zabývá velmi zajímavým a přínosným tématem. Disertace se opírá a navazuje na řadu výsledků, které byly publikovány teprve nedávno v kvalitních odborných časopisech. Domnívám se proto, že práce řeší skutečně aktuální téma s mezinárodním dosahem a lze předpokládat i mezinárodní ohlas.

Práce se skládá ze čtyř kapitol. První sumarizuje základní známé výsledky a předpoklady v oblasti aproximací založených na sedlovém bodě. Tato kapitola je doplněna numerickou studií autorky pro speciální třídu odhadů, pro kterou je znám přesný vzorec pro hustoty těchto odhadů. Je proto možné provést srovnání takto získaných hustot s jejich aproximacemi založenými na technikách sedlového bodu. Provedení simulace nebylo rutinní záležitostí, protože bylo potřeba nejen implementovat metodiku sedlového bodu ale i výpočet "přesných" hodnot hustoty, neboť vzorec obsahoval integrály, které bylo nutné řešit numericky.

Za nejcennější část disertace lze pokládat druhou a třetí kapitolu. Ve druhé kapitole je odvozena aproximace hustoty kvantilů a regresních kvantilů. Třetí kapitola je věnována problematice testů. Za nejhodnotnější výsledek považuji konstrukci testů pro regresní kvantily spočívajících na technice sedlového bodu. Čtvrtou kapitolu spojuje s předchozím textem téma aproximací založených na sedlovém bodě, ale jinak se věnuje poměrně odlišnému tématu – Kullbackově-Leiblerově divergenci pro exponenciální třídu rozdělení.

Konstatuji, že autorka úspěšně dokázala odvodit řadu výsledků v oblasti aproximací založených na sedlovém bodě v řadě situací, pro které v literatuře výsledek není známý. Získané výsledky jsou sice očekávané, nicméně důkazy těchto tvrzení nebyly přímočaré, ale naopak často poměrně komplikované. Z práce bezpochyby vyplývá, že R. Sabolová studované problematice dobře porozuměla a že byla schopna existující výsledky využít a tvůrčím způsobem rozšířit na další případy.

V důkazech tvrzení jsem nenalezl žádná závažná a problematická místa a podle mého mínění jsou provedeny korektně. Práce je psána v angličtině, která odpovídá odbornému jazyku v časopisech s tématicky blízkými články. Na druhé straně se domnívám, že autorka mohla vlastnímu zpracování disertace věnovat více pozornosti. V práci jsem našel poměrně mnoho překlepů, řadu nejednoznačností používaného značení, formulačních opomenutí či drobných přehlédnutí v logice vývoje výkladu, které ztěžují čtenářovu orientaci.

Především první kapitola, která by měla shrnovat známé výsledky, metody a motivaci práci působí poněkud neuspořádaným dojmem. Pro čtenáře je například velmi obtížné sledovat, kdy se jedná o jednorozměrnou veličinu či parametr a kdy o vícerozměrný případ. Bylo by vhodné použít bold symboly pro odlišení vícerozměrných případů, jak tomu je v následujících kapitolách. Nalezl jsem zde např. tyto nejasnosti a překlepy:

- str. 6, předpoklady S1-S5: Není zcela jasné použité označení, např. jaký je vztah mezi  $\varphi$  (S4) a  $\varphi_r$  (S1-S3), podle čeho se derivuje při zavedení matice  $A(\theta_0)$  v (S4).
- str. 8, (1.3) a (1.4): Nechybí mínus v (1.4) před  $\frac{f'}{f}$  ?
- str. 9<sub>12</sub>: "...in the Fig. 1.1 ...saddlepoint approximation by the red dashed line and formula (1.4) by the green dotted line", ale na obr. 1.1 je pouze red dotted line a green dashed line.
- str. 9<sub>4</sub>: "...into to formula (1.2) for  $f_n$ ", ale (1.2) obsahuje  $g_n$ .
- ...

Ale i v dalších kapitolách nalézám řadu drobných nejasností:

- str. 26, Theorem 2: V tvrzeních jsou použity symboly  $\beta_\alpha$ ,  $\hat{\beta}_\alpha$  a  $\beta$ , přičemž není zavedeno značení  $\hat{\beta}_\alpha$  a  $\beta_\alpha$ . Z kontextu je patrné, že  $\hat{\beta}_\alpha$  je  $\alpha$ -regresní kvantil, ale co je  $\beta_\alpha$  jasné v daném místě není.
- str. 36, Theorem 5: Ve vzorci aproximované hustoty regresních kvantilů  $\hat{g}_n$  je uveden exponent  $s$ . Co tento parametr označuje, jaké má hodnoty ?
- str. 43, Remark 6: Pokud se odhadne quantile density function v testové statistice, zůstane rozdělení statistiky stejné ?
- str. 44 a 46: V označení empirické distribuční funkce  $F_n$  se jednou používá ostrá, posléze neostrá nerovnost.
- str. 44: Co je  $F_n^{-1}$ , když  $F_n$  je empirické distribuční funkce ?
- str. 50, Theorem 9: Předpokládá se podmínka (A2), která se týká pevné matice, ale v uvažovaném modelu se pracuje s náhodnou maticí.
- str. 71, (4.1.): V integrálu má být místo  $dx$  patrně  $dy$ .
- ...

Za cenné v práci považuji to, že studované metody byly dovedeny ve většině případů až do numerického řešení, které není v dané problematice jednoduché implementovat. Je však trochu škoda, že připravené procedury netvoří nějakou přílohu této práce, aby bylo možné provedené výpočty "zkontrolovat".

Nedomnívám se však, že by výše uvedené kritické poznámky výrazněji ovlivnily celkově dobrou úroveň disertační práce. Mohou být brány v úvahu zejména při další publikaci výsledků.

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, 2. srpna 2014

prof. RNDr. Jan Pícek, CSc.