

## POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

**Název:** Statistická inference v modelech s proměnlivými koeficienty

**Autor:** Bc. Martin Splítek

### SHRNUTÍ OBSAHU PRÁCE

Práce pojednává o modelech s proměnlivými koeficienty pomocí splajnov a lokální regrese. Teoretické přístupy jsou aplikovány v simulacích a na reálná data pomocí software R (balíček `mgcv`).

### CELKOVÉ HODNOCENÍ PRÁCE

**Téma práce.** Velmi obtížné téma bylo zpracováno jako triviální kompilace a tak do stejné míry splňuje zadání práce.

**Vlastní příspěvek.** Vlastní příspěvek řešitele práce spočívá v provedené nevhodné simulační studii, v aplikování převzatých statistických přístupů na reálná data a našťudování části balíčku `mgcv`. Řešitel však příliš velkou příležitost věnuje převzaté teorii a malou by věnoval více důrazu (nerovná se příležitost) analýze dat (kterou navíc zredukoval jednoduchým vymazáním podstatné části první neobhájené verze diplomové práce, která byla více než problematická). Řešitel si zvolil bohatý statistický aparát, z čehož potom pramení i nižší uvedené otázky a připomínky.

**Matematická úroveň.** Práce je na podprůměrné matematické úrovni.

**Práce se zdroji.** Bibliografické zdroje v práci jsou správně citovány až na jednu kapitolu knihy, která je považována za článek v neexistujícím časopise.

**Formální úprava.** Grafická úprava je průměrná. Stylistická úprava práce je v pořádku vzhledem k rozsahu práce.

### PŘIPOMÍNKY A OTÁZKY

1. Teoretická část obsahuje prepis dvou článků: Fang and Zhang (1999) a Huang et al. (2004). Je napsána útržkovitě. Působí jako prepisování značení, vzorců a znení vět.
2. Praktická část obsahuje popis už jen jednoho datového souboru (narozdíl od první verze, kde byly datové soubory dva) a velmi skúpe statistické závěry (o analýze nemůže být řeč). Zdrojový kód práce neobsahuje, avšak byl by triviální vzhledem k existenci balíčku `mgcv`.
3. Formulace Věty 1: “Předpokládejme, že podmínky z (Fan a Zhang, 1999) jsou splněny.” je nevhodná. Nie je jasné, o které podmínky sa jedná. Ak o všetkých sedem spomenutých v článku, tak je to nesprávne. Prečo teda nie je predpoklad 7 potrebný pre platnosť záveru Vety 1?
4. Str. 10, vzorec (2.2): Stačí, aby  $\beta_j$  bola “funkce se spojitými derivacemi do řádu 2”? A je to vůbec vhodné?
5. Věta 6 neobsahuje tvrzení, takže sa ani nedá považovať za matematickú vetu. Čo bolo účelom tzv. Vety 6?

6. Simulačná štúdia obsahuje nagenovanie dát z troch modelov, kde prvé dva modely nie sú modelmi s časovo premenlivými koeficientami. Tretí model už je, ale jedná sa o najjednoduchší časovo premenlivý koeficient a to lineárne sa meniaci. Výsledky simulácií pre tento tretí model práca neobsahuje.
7. V simulačnej štúdiu, ak už mala byť nosnou časťou práce, chýbajú netriviálne príklady reflektujúce reálne použitie metodológie v praxi.
8. Strana 41: “V predchádzajúcom textu ale výsledky týchto simulácií nezmiňujeme a to z toho dôvodu, že vyšly téměř shodne s výsledky simulácií na datech s konštantne jednotkovým koeficientom  $\beta(t)$ .” Nemyslím si, že výsledky sú podobné. Vyvráťte moje tvrdenie!
9. Testovací problém (5.2) nie je ekvivalentný problému (5.3). Prečo je tomu tak? Testová štatistika pre test (5.3) je nesprávna.
10. Ako by ste testovali  $H_0 : \{\forall t : \beta(t) = \beta\}$  proti  $H_1 : \{\exists t : \beta(t) \neq \beta\}$ , kde  $\beta$  je neznáma konštanta?

#### ZÁVĚR

Práci považuji za vynikajúcu/velmi dobrou/průměrnou/podprůměrnou/nevyhovujúcu a doporučuji ji uznať ako diplomovú prácu za podmienky, že študent objasní a opraví, čo bolo uvedené v pripomienkach a otázkach.

*Návrh klasifikácie sdělím předsedovi zkušební komise.*

Jméno oponenta: RNDr. Michal Pešta, Ph.D.

Pracoviště: KPMS MFF UK

Datum: 18. 1. 2018