

Posudek oponenta diplomové práce

Antonín Koubek: Bodové procesy v čase a prostoru

V předložené diplomové práci jsou studovány modely časoprostorových bodových procesů, které jsou známy jako shot noise Coxovy procesy. Jde o velkou třídu Coxových procesů a zároveň Poissonových shlukových procesů. Nejjednodušší přístup při modelování časoprostorových procesů je uvažovat separabilní strukturu, kde časová a prostorová složka jsou navzájem nezávislé a dají se modelovat každá zvlášť. V práci jsou zvoleny tři různé modely, kromě jednoho separabilního i dva neseperabilní, které jsou založené na tzv. ambitových množinách kuželovitého tvaru. Autor popisuje (přibližnou) simulaci z těchto modelů a zabývá se odhadem funkcionálních popisných charakteristik (funkce intenzity, párové korelační funkce, K -funkce a F -funkce) z replikovaných pozorování. Pro tři vybrané modely je počítačově vygenerováno několik realizací daného časoprostorového procesu a na základě nich jsou provedeny numerické výpočty jádrových odhadů s vhodně zvolenou šířkou pásma. Tyto odhady jsou srovnány se spočtenými teoretickými charakteristikami.

Práce je rozdělena do pěti kapitol. První kapitola shrnuje potřebné základy teorie prostorových bodových procesů a rozšíření na časoprostorový případ. Ve druhé kapitole je nejprve vysvětlen prostorový shot noise Coxův bodový proces a poté zobecněn na situaci procesu v čase a prostoru. Dále jsou definovány tři speciální modely, které se později využijí v numerické části práce. Třetí kapitola popisuje použitý algoritmus pro simulaci a čtvrtá metody odhadu funkcionálních popisných charakteristik z n replikací. Závěrečná kapitola je věnována zčásti teoretickým výpočtům charakteristik pro zvolené modely a zčásti výsledkům simulační studie, která byla prováděna v programu *Mathematica*. Některá obecná teoretická odvození by bylo vhodnější zařadit dříve než až do páté kapitole. Například odvození toho, že funkce F je konstantní v separabilním případě (str. 34), by se lépe hodilo na konec podkapitoly 2.2.

Antonín Koubek ve své diplomové práci zpracovává zajímavé a aktuální téma. Statistice časoprostorových bodových procesů zatím nebylo v literatuře věnováno tolik pozornosti jako jejich modelování. Důležitým přínosem práce je implementace výpočtu odhadů charakteristik a zkoumání jejich chování pro ambitové modely. Výsledky můžou například napovědět, jestli z odhadů F -funkce lze usuzovat, zda daný model je separabilní či ne. Je škoda, že komentáře k obdržným závěrům simulační studie jsou celkem stručné. Provedení všech numerických výpočtů stálo jistě hodně námahy, možná proto nezbylo tolik času k samotnému zpracování práce, které by si dle mého názoru zasloužilo více pozornosti a pečlivosti. Teoretická část obsahuje několik netriviálních výpočtů a odvození, zároveň se v ní ale vyskytuje řada nepřesností a překlepů. Překlepy jsou například ve znění Proposition 6 (chybějící \neq), Proposition 10 (ρ_u a φ_u), Proposition 14 ($Z(u, v)$ má být $Z(u, t)$), vzorci 2.16 (prohozeno k_1 a k_2), vzorci 4.7 (r není z W), vzorci 4.11 (místo r má být t), vzorci 5.8 (přebytečné druhé mocniny v čitateli a dvojka ve jmenovateli) nebo vzorci 5.11 (chybějící mínusy v exponenciálách). Autor by si měl uvědomit rozdíl mezi $B \in \mathcal{B}^d$ a $B \subseteq \mathcal{B}^d$. Na začátku podkapitoly 1.2 není vhodné psát vztah $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}_+$, který neplatí, navíc ve zbytku práce se pro časovou složku používá \mathbb{R} a ne \mathbb{R}_+ . V Note 14 a v Example 2 je Campbellova věta použita chybně. Vztah před odvozením tvaru funkce F na str. 34 rovněž není v pořádku, neboť se předchozí vztah neintegruje přes množinu míry 1. Konkrétní dotazy k obsahu práce mám následující:

1. Z definice ambitové množiny (Definition 28) by plynulo, že $A_0(0) = \{(0, 0)\}$. Jak by měla vypadat správně? Možná tady bylo vhodné zařadit obrázek ilustrující definici.
2. Při odhadu funkce intenzity separabilního modelu se ve jmenovateli objevuje N (str. 27 dole), což je střední počet v pozorovacím okně. V praxi ale tato hodnota není známá. Jak se v tomto případě postupovalo? Použila se skutečná hodnota nebo odhad z dat?
3. U odhadu K -funkce (str. 29) se uvádí, že je přibližně nestranný? V jakém smyslu se má tato přibližnost chápat?
4. Je numerická hodnota c_2 na str. 38 správně? Mně po dosazení vyšlo $c_2 = 4\,777\,600$.

Kromě opravení věcných nedopadnutí by textu prospělo vylepšení struktury. Na několika místech se vzorce zbytečně opakují. Co se týče jazykové stránky, tak práce je psána průměrnou angličtinou, hlavní nedostatky spočívají v používání členů, slovosledu a některých nesrozumitelných formulacích. Dále bych vytkl nepřesně psaná vlastní jména (Neyman, Møller) a z typografických prohřešků psaní pomlček a spojovníků nebo chybějící matematickou sazbu u některých symbolů. Také by se slušelo neodkazovat na neexistující tabulku (str. 39 dole).

I přes uvedené výhrady hodnotím práci pozitivně, student prokázal schopnost nastudovat a samostatně zpracovat zadané téma, provést netriviální výpočty a vše přehledně sepsat v ucelený celek. Posuzovaná diplomová práce splňuje požadavky na ni kladené a **doporučuji ji uznat za diplomovou práci na MFF UK.**

V Praze, 26. srpna 2013

RNDr. Zbyněk Pawlas, Ph.D.