

## Posudek na disertační práci Mgr. Markéty Zikmundové „*Interacting spatial particle systems*“, KPMS MFF UK v Praze

Autorka v práci navazuje na svůj víceletý výzkum v oblasti modelů a inference pro prostorové náhodné procesy objektů (částic) a shrnuje své výsledky, včetně nejnovějších připravených k publikaci. Po úvodu do problematiky procesů, tj. modelovaných přes referenční značkové Poissonovy procesy modulované pomocí Gibbsovských hustot závislých na prostorových statistikách (v kapitole 2) je kapitola 3 věnována funkcionalům těchto statistik - charakteristik a jejich momentům. Zpravidla se jedná o U-statistiky. Další kapitola 4 pak přechází k určitému typu těchto procesů vzájemně reagujících objektů. Speciálně se pracuje s procesy kruhů či úseček v  $\mathbb{R}^2$ , případně s rozšířením do  $\mathbb{R}^3$ . Následně se zkoumají a dokazují i limitní vlastnosti momentů příslušných charakteristik, U-statistik. Zatímco kapitoly 3 a 4 jsou víceméně teoretické povahy, pracující s modely a jejich charakteristikami, kapitola 5 se zabývá statistickou inferencí v některých modelech, včetně rozšíření modelu o časový vývoj. Poslední kapitola pak diskutuje výsledky algoritmů z kap. 5.

Nyní tedy mé poznámky a připomínky k práci:

1. Název kap. 5 (Space-time models) je poměrně obecný, ale uvažuje se zde jen jeden způsob zavedení časového vývoje procesu, a to vývojem parametrů „modulující“ hustoty (vzhledem k referenčnímu Poissonovu procesu). Mně připadá, že celý model má 2 úrovně, onen referenční, který je skrytý, ale připadá mi důležitý pro mechanismus vzniku procesu a pak onen pozorovaný, s kterým se v disertaci pracuje. Takže další možností je uvažovat časový vývoj v referenčním (základním) Poissonovu procesu.
2. Ze stejného důvodu bych čekal, že po odhadu parametrů pozorovaného procesu  $(x, a, \sigma)$  se bude pokračovat k identifikaci parametrů  $\rho$  a v našem případě distribuce  $r$ . Jsou nejméně 2 důvody. Jednak to v reálných případech vede k porozumění mechanismu vzniku procesu, a z hlediska statistické inference je to, myslím, třeba k testování fitu modelu. V kapitolce 5.3.3. je naznačen onen způsob, jak porovnat zvolený (a odhadnutý) model s realitou, tj. pomocí generování dat z odhadnutého modelu a porovnání určitých charakteristik. Je to v příkladech kap. 5 i provedeno (tzv. envelopes), tak by mne zajímalo, jak byly vygenerovány bez znalosti referenčního modelu.
3. Metody MC, a spec. MCMC, jsou vyvinuty k generování reprezentací distribucí, tj. užívají se jednak k „simulaci“ dat z modelu, jednak v statistické inferenci, spec. ke generování reprezentace a posteriorního rozdělení v bayesovském kontextu (zde není slovo „simulace“ na místě, i když se tyto výrazy dost zaměňují). V kapitole 5 je to také tak, ale je pravda, že z kontextu vždy vyplyne, který algoritmus je pro generování (tj. „simulaci“) realizace procesu a který slouží pro inferenci.

4. Poznámka ke konvergenci momentů U-statistik: Je to analogie situace, kdy v analýze přežití sledujeme více „stejných“ procesů, čili skutečně intenzita roste do nekonečna. Zde si dovedu představit situaci třeba pozorování seizmologických dat [29], po rostoucí dobu, ale v jiných aplikacích by odpovídající „large sample“ situace vznikla spíš pozorováním ve více oknech. Problém samozřejmě je stacionarita procesu (v čase či lokalitě).
5. Trochu zde postrádám nějakou reálnou aplikaci modelů, hlavně z hlediska statistické inference, včetně testu vhodnosti modelu (viz i pozn. 2). Samozřejmě, v literatuře (i autorčině) se objevuje několik typických aplikací, např. modelu interagujících kruhů na popis rozšíření vegetace. V této souvislosti jsou určité zajímavé i modely zahrnující vliv dalších faktorů (tj. regresní modely). Kde by se objevila ona závislost (v modelech vegetace třeba na nadm. výšce, obsahu C v půdě,...)? Předpokládám, že nejspíš v základním referenčním modelu, spíše než v parametrech  $x$  pro pozorovaný model? Má s takovým rozšířením autorka nějakou zkušenost?
6. Pokud jde o příklady v kap. 5, vlastně se opakuje jediný se stejnou konfigurací a vývojem parametrů  $x$ . Nebylo by od věci mít i jiný, který by vedl k jinému časovému vývoji procesu.
7. Jinak, práce je psána (anglicky) dostatečně pečlivě, nenalezl jsem prakticky žádné překlepy, mám vlastně jen 2 poznámky.
  - 7.1. Na obr. 5.1. je 12 sub-obrázků, a má zobrazovat vývoj procesu v 6-ti různých časech. Co tedy přesně je na onom obrázku?
  - 7.2. Zdá se, že veličina  $l(y)$ , poprvé použitá v 4.2.3., není předtím definovaná (z kontextu jde zřejmě o náhodnou délku úsečky).

Shrnutí: Práce je napsaná přehledně, je logicky uspořádána a obsahuje bohatý materiál s vlastními příspěvky autorky hned v několika směrech. Mé poznámky a připomínky uvedené shora se týkají jen dílčích otázek a nijak nesnižují hodnotu práce. Proto soudím, že práce splňuje požadavky na práci disertační a doporučuji ji k obhajobě.

28.10.2014

Petr Volf  
ÚTIA AV ČR