

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Název: Poissonův shlukový model

Autor: Bc. Tomáš Růžička

SHRNUTÍ OBSAHU PRÁCE

V práci je zaveden kótovaný Poissonův shlukový bodový proces nejprve na obecném měřitelném prostoru a poté použit ve speciálním jednorozměrném případě pro modelování okamžiků pojistných událostí a následných plateb pojistného plnění. Kóty procesu odpovídají zpoždění plateb v letech a jejich výši. Na základě tohoto modelu autor odvozuje predikce budoucích kumulovaných plateb. Pro porovnání vhodnosti modelu jsou uvedeny i další dva přístupy k získání predikcí pro celkové výše pojistného plnění. Vše je otestováno nejprve na uměle vytvořených a následně i reálných datech.

CELKOVÉ HODNOCENÍ PRÁCE

Diplomová práce má přiměřený rozsah. Jazyková, grafická i matematická prezentace je na velmi dobré úrovni. Autor prokazuje schopnost pracovat s několika zdroji zároveň, vytvořit z nich kompaktní rešerši, doplnit chybějící detaily či zobecnit vybrané výsledky. Jediný prohřešek shledávám v poněkud vágním zacházení s aproximačním znaménkem \approx . Z textu není vůbec jasné, o jak kvalitní aproximace se jedná. Dále bych doporučila trochu citlivější přístup k vynechávání různých indexů a zavádění parametrů funkcí.

Téma práce. Téma bylo v souladu se studijním programem autora, opírá se o základy neživotního pojištění a teorie pravděpodobnosti. Autor si dále musel nastudovat základy bodových procesů. Dle zadání autor přehledně popsal základní vlastnosti Poissonova shlukového procesu a aplikoval jej na data poskytnutá pojišťovnou Allianz. Nad rámec tématu byly dále diskutované i jiné přístupy k problematice predikcí škodních rezerv.

Vlastní příspěvek. Autor přehledně zavádí Poissonův shlukový proces s kótami a diskutuje jeho vlastnosti a možnosti použití v neživotním pojištění. Zde se opírá zejména o článek Jessen a kol. (2011), kde je však Poissonův shlukový proces definován trochu odlišně - rodičovský proces má neznámé rozdělení a velikost jednotlivých shluků se řídí Poissonovým rozdělením s parametrem μ . Na druhé straně v předložené práci je rodičovský proces Poissonův s intenzitou λ a velikost shluků má neznámé rozdělení. Modely nejsou totožné, a proto bylo potřeba myšlenky důkazů přizpůsobit či zobecnit požadované situaci. V tomto ohledu pokládám za největší teoretický přínos Tvzení 2 a 3, kde autor sám odvozuje analytický vzorec pro podmíněnou střední hodnotu celkové výše výplat v nadcházejícím roce. Tento vzorec je ale nevhodný z hlediska výpočetních času, a tak autor volí předpoklad Poissonovských součtů velikostí shluků k aproximaci tohoto vzorce. Tím se dostává zpět do souladu s výše zmíněným článkem. Uvedeno je i odvození víceřadkové predikce, ale už v značně menším detailu. Dalším přínosem je návod na odhady parametrů tohoto modelu, tj. například výše jednotlivých plateb, jejich zpoždění a dále i velikosti shluků za předpokladu, že jejich rozdělení pochází ze speciální (a,b) třídy. Tento odhad je opět inspirován článkem Jessen a kol (2011), kde ale chybí detailnější odvození. To je obsaženo v důkazech Tvzení 9 a 10 předložené práce.

Na datech simulovaných z Poissonova shlukového modelu poté porovnává predikce odvozené v Sekci 2.1. s predikcemi alternativních metod modelování škodních rezerv: chain ladder modelem a zobecněným lineárním modelem. Není příliš překvapivé, že pro tato data vychází predikce odvozené z Poissonova shlukového modelu nejlépe. Zajímavější je ale porovnání přístupů na reálných datech Allianz pojišťovny. Zde nefunguje Poissonův shlukový model až tak dobře, ale

může to být jen nevhodnou kombinací parametrů modelu. V praktické části diplomové práce velmi oceňuji doprovodné komentáře autora ke všem tabulkám a výsledkům včetně vysvětlení, z jakého důvodu je daný přístup vhodný či nevhodný.

Matematická úroveň. Matematické porozumění práce by usnadnilo například použití $:=$ ve vzorcích, kde se definuje nějaký nový objekt. Jak zmiňuji výše, vypuštění indexu i je někdy zavádějící, protože základem jsou trojúhelníková data. Na vícero místech se pak nijak nehlídají hodnoty indexů j a l vzhledem k i a J . Za velmi zbytečné považuji zavádění kótování rodičovského Poissonova procesu výškou platby v době vzniku pojistné události vzhledem k tomu, že tato kóta nehraje žádnou roli. Stejně tak není nutné indexovat něčím, co se v daném objektu nevyskytuje (například $H_{t,k,j+1}$ neobsahuje žádné k , stejně tak působí zvláště člen $k|n$ v definici $p(m, k|n)$ na konci stránky 12.) Nejvíce nejasná je ale aproximace pomocí \approx . Není zde nijak specifikovaná cena použití této aproximace nebo její přínos oproti triviální aproximaci $R_l^C \approx c > 0$. Na druhou stranu vše je včas a korektně značeno.

Práce se zdroji. Práce obsahuje mnoho přejatých myšlenek, značení nebo tvrzení, ale vždy je vše řádně citováno. V několika případech bych ocenila odkaz na konkrétní tvrzení nebo stranu. Nicméně rešeršní část práce rozhodně nebyla provedena doslovným opisem.

Formální úprava. Jazyková úroveň práce je na excelentní úrovni, text je vždy napsán čistě a věcně. Struktura práce je taktéž logická. Všechny tabulky a obrázky jsou řádně očíslovány, přiložen je i jejich seznam a zdrojový kód ke všem simulacím.

PŘIPOMÍNKY A OTÁZKY Níže uvádím další připomínky a otázky k obhajobě.

1. Má smysl uvažovat zobecnění na závisle kótovaný Poissonův proces? Dala by se uvažovat nějaká kovarianční struktura mezi jednotlivými kótami procesu, případně alespoň mezi výškami plateb v rámci jednoho shluku? Který z předpokladů nezávislosti v modelu považujete za nejméně odpovídající reálným datům?
2. Jaká je míra intenzity ve smyslu Definice 3 výsledného Poissonova shlukového procesu popsaného parametry na začátku Sekce 2.1?
3. Závisí nějak Tvrzení 2 na indexu i (řádku dat) nebo bude predikce vycházet pro každé i stejně?
4. V důkazu Tvrzení 2 není úplně přesné, že má vektor (N_0, \dots, N_{j+1}) multinomické rozdělení s parametrem $\sum_{i=1}^m k_i$. To by nastalo pouze za podmínky, že $\sum_{i=0}^{j+1} N_i = \sum_{i=1}^m k_i$. To však nemusí nastat. Předpokládám tedy, že $Q_{j+1}(m, k_1, \dots, k_m)$ značí vzorec (2.3), pak by to bylo dál v pořádku.
5. Jak striktní jsou předpoklady Chain ladder přístupu a zobecněného lineárního modelu v porovnání s Poissonovým shlukovým modelem s použitou aproximací?

ZÁVĚR

Práci považuji za velmi dobrou a rozhodně ji doporučuji uznat jako diplomovou práci.

Daniela Flimmel
KPMS MFF UK
24. května 2024