Posudek oponenta bakalářské práce

Název: Aproximácia rozdelenia úhrnov škôd
Autor: Viktória Antalicová

Shrnutí obsahu práce

Práce pojednává o různých možnostech aproximace složeného Poissonova rozdělení, které se používá jako model rozdělení úhrnů škod. Pomocí simulační studie autorka porovnává vybrané hodnoty funkce přežití pro různá teoretická rozdělení (gama rozdělení, log-normální rozdělení a posunuté gama rozdělení) a pro rozdělení získané Edgeworthovou aproximací s empirickými hodnotami funkce přežití složeného Poissonova rozdělení. Parametry teoretických rozdělení byly odhadovány momentovou metodou.

Celkové hodnocení práce

Téma práce. Náročnost tématu považuji za přiměřené a odpovídající bakalářské práci. Zadání práce bylo splněno.

Vlastní příspěvek. Vlastní příspěvek autorky není v práci jasně specifikován. Domnívám se, že spočívá v provedení simulační studie a vyhodnocení jejích výsledků. Možná autorka doplnila některé kroky při odvozování Edgeworthovy aproximace.

Matematická úroveň. Práce je zaměřena spíše prakticky, těžiště spočívá v provedení a vyhodnocení simulační studie. Nicméně v teoretické části je nezbytný matematický aparát popsán (až na výjimky zmíněné níže) korektně a rigorózně.

Práce se zdroji. Použité zdroje jsou v práci náležitě citovány, nejsem si vědom doslova zkopírovaných nebo otrocky přeložených pasáží.

Formální úprava. Formální úpravu považuji za velmi zdařilou, práce je psána přehledně a srozumitelně. Práce obsahuje jen minimum překlepů.

Připomínky a otázky

1. Použití momentové metody pro škody podléhající zajištění (str. 13) není dle mého názoru korektní. Výraz pro teoretickou střední hodnotu vyjadřuje skutečnou (nepodmíněnou) střední hodnotu díky faktoru p_1, zatímco výběrový průměr níže odhaduje podmíněnou střední hodnotu (za podmínky, že úrhn škod je kladný) a to kvůli tomu, že výraz n_1 ve jmenovateli zahrnuje jen počet nenulových hodnot.

2. Název kapitoly 2.3. (Testování dobré shody) je poněkud zavádějící. Ve skutečnosti se nejedná o rigorózní statistický test dobré shody v pravém smyslu, nýbrž o pouhý výpočet rozdílů empirických a teoretických hodnot funkce přežití. Navíc, porovnání funkcí přežití pouhým rozdílem nemusí být vhodné pro posouzení chování ve chvostech, tam budou totiž rozdíly vždy spíše malé.

3. Vyloučení metody maximální věrohodnosti z celé studie pouze na základě výsledků jednoho scénáře považuji za unáhlené. Často dává tato metoda lepší odhady než momentová metoda a bylo by zajímavé vidět i výsledky pro ni.

4. Tvrzení na str. 27, že rozdělení získané Edgeworthovou aproximací dobře popisuje skutečné rozdělení simulovaných úhrnů škod, je neopodstatněné v tom smyslu, že autorka neporovnávala Edgeworthovu aproximaci se skutečným rozdělením, nýbrž pouze s empirickým rozdělením. To je ostatně omezení závěrů celé práce, totiž namísto testování shody se skutečným rozdělením se jen porovnávají různé modely se simulovanými daty. Tímto postupem však může docházet k tzv. overfittingu (volbě modelů příliš velkým počtem parametrů).

5. Místy se v práci objevuje nevhodný font pro matematické výrazy (např. výraz "arg max" na str. 11, log místo log pro logaritmus, d místo d pro symbol derivace, symbol "n" na 1. řádku na str. 13, atp.)

6. Značení $r^k \frac{1}{k!} \frac{d^k}{dr^k} log M_Z(r)|_{r=0}$ na str. 23 je poněkud matoucí (krom nevhodných fontů). V bodě $r = 0$ se totiž má vyhodnotit pouze derivace, ne celý výraz.

7. Mohla by se autorka vyjádřit k otázce existence a jednoznačnosti řešení soustav nelineárních rovnic pro momentové odhady na str. 12?

Závěr

Práci považuji za průměrnou a doporučuji ji uznat jako bakalářskou práci.

Pavel Kříž
KPMS, MFF UK
22. srpna 2023