

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Název: Zobecněné lineární a aditivní modely v pojišťovnictví

Autor: Bc. Peter Rusnák

SHRNUTÍ OBSAHU PRÁCE

Práce pojednává o GLM a GAM s aplikáciou v neživotnom poistení. Teoretické prístupy sú aplikované na reálne dáta pomocou software R.

CELKOVÉ HODNOCENÍ PRÁCE

Téma práce. Obtiažne téma bolo spracované štandardne a tak spĺňa zadanie práce.

Vlastní příspěvek. Vlastný príspevok riešiteľa práce spočíva v aplikovaní štandardných štatistických prístupov na reálne dáta z poisťovníctva. Riešiteľ však príliš veľa priestoru venuje ustálenej teórii a mal by venovať viac dôrazu (nerovná sa priestoru) analýze dát. V tomto konkrétnom prípade je analýza náročná a riešiteľ si zvolil bohatý štatistický aparát, z čoho potom pramenia i nižšie uvedené otázky a pripomienky.

Matematická úroveň. Práca je na priemernej matematickej úrovni.

Práce se zdroji. Bibliografické zdroje v práci sú všetky až na zdroj [4] nesprávne citované, vid'. [6] (napr. chýba autor). U zvyšných kníh chýba miesto vydania. Časopisy neobsahujú číslo zväzku, prípadne číslo výtlačku.

Formální úprava. Grafická úprava je podpriemerná. Štylistická úprava práce je v poriadku vzhľadom na rozsah práce, ale rušivé sú časté preklepy v angličtine. Kontrola pravopisu chýba.

PŘIPOMÍNKY A OTÁZKY

1. Prvých 25 strán práce je otrocké a zbytočné opisovanie teórie GLM (vid'. [7]), ktorou autor umelo rozširuje prácu. Referencie a stručný úvod so značením by boli postačujúce. Opisovanie dôkazov je v diplomovej práci absolútne nevhodné. Navyše chýbajú pri zkopírovaných dôkazoch citácie, odkiaľ bol dôkaz prevzatý.
2. Kapitola 2 obsahuje prevzaté vety, súvetia a tabuľky z knihy Fox, J. (2002): An R and S-Plus Companion to Applied Regression, resp. Fox, J. (2008): Applied Regression Analysis and Generalized Linear Models. Ani jedna z týchto kníh nie je v zozname bibliografických zdrojov.
3. Dôkaz vety na str. 16 je nesprávny. Čo by sa stalo, ak by nosič $S(Y)$ závisel na parametri θ ?
4. Má štatistika na str. 42, r. 2 vždy asymptoticky χ^2 -rozdelenie za platnosti H_0 zo str. 41? Čo ak pre i -tu zložku vektora β_j platí $\beta_j^{(i)} \geq 0$ (tj. testovaná hodnota je na okraji parametrického priestoru)?
5. V Kapitole 5 nie je uvedené, aká je veľkosť dátového súboru. Ak by sa jednalo o desiatky tisíc pozorovaní, je rozumné používať formálne štatistické testy na testovanie signifikantnosti parametrov modelu? Aká bude sila týchto testov? A čo to prakticky znamená?
6. Podkapitola 5.2 uvažuje pre vek 7-uzlový spline. Ako sú volené uzly? Nebolo by vhodnejšie použiť rovno 7 vekových kategórií? V čom je teda prínos GAM v tomto prípade od GLM?

7. Prečo Tabuľky 5.3 a 5.5 obsahujú neceločíselné stupne voľnosti? Ako boli spočítané?
8. Ako vznikol Obr. 1, z ktorého modelu? Aká je jeho spojitosť s modelom (16)? Čo je znázornené bodkovanou čiarou a ako bola skonštruovaná/napočítaná? Oranžová a zelená čiara pre vyššie x -ové hodnoty značne líšia. Prečo? Je teda odpovedajúci model vhodný? Analogické otázky sa týkajú Obr. 2 a modelu (17).
9. Ako bol skonštruovaný Obr. 4 (vľavo)? Čomu odpovedajú hranice *No resolution* a *No skill*? Existuje nejaké prepojenie medzi niektorými krivkami z obrázku a niektorým vzorcom z odprezentovanej teórie? Je odpovedajúci model vhodný na základe tohoto obrázku, keď červená čiara vôbec neodpovedá čiernej a navyše vybieha zo sivého regiónu?
10. Je vôbec niektorý z odprezentovaných modelov vhodný vhlľadom na ľavé časti Obr. 5, 6 a 7?
11. Str. 46 obsahuje chybný odkaz na Table 4.1 namiesto na Table 5.1.
12. Namiesto Picture 1 sa píše Figure 1, atď.
13. Maticové násobenie je spočiatku neštandardne značené bodkou, neskôr už nie. Značenie by malo byť konzistentné počas celej práce.
14. Chýba záver práce.

ZÁVĚR

Práci považuji za ~~vynikajúcej/~~velmi dobrú/průměrnou/podprůměrnou/nevyhovujúcej a doporučuji ji uznať jako diplomovou práci.

Návrh klasifikace sdělím předsedovi zkušební komise.

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

Pracoviště: KPMS MFF UK

Datum: 19. 8. 2013