

Oponentský posudek na diplomovou práci:

Marika Stroukalová : Modelování systémů bonus-malus

Diplomová práce slečny Stroukalové se zabývá systémy bonus-malus, matematickým modelem je popisujícím, odhadem parametrů tohoto modelu z pojistných dat a otázkou určení optimálních sazeb pojistného. Představené teoretické postupy jsou aplikovány na data jedné konkrétní české pojišťovny a získané teoretické optimální sazby pojistného porovnány s aktuálními sazbami této pojišťovny.

V práci představený model vychází z předpokladu Poissonovského počtu škod pro jednotlivé pojištěnce. Intenzita tohoto Poissonova rozdělení je ovšem součinem očekávané škodní frekvence pro daného pojištěnce, jež je regresní funkcí známých charakteristik pojištěnce zařazujících ho do apriorní tarifní třídy, a (neznámé) individuální rizikové úrovně tohoto pojištěnce, u níž se předpokládá Gamma rozdělení. Tento model a odvození maximálně věrohodných odhadů parametrů tohoto modelu ze škodních dat je popsáno v kapitole 1. V kapitole 2 je stručně zopakováno, jak použití bonus-malus pravidel a uvažovaného modelu na počet škod vede na Markovův řetězec popisující průchod pojištěnce mezi bonusovými třídami v po sobě jdoucích letech.

Třetí kapitola pak řeší ústřední otázku práce – jak optimálně nastavit bonusy pro jednotlivé třídy bonus-malus systému tak, aby co nejlépe aproximovaly neznámou individuální rizikovou úroveň jednotlivých pojištěnců. Je popsán přístup s kvadratickou ztrátovou funkcí, která je minimalizována podmíněnou střední hodnotou neznámé individuální rizikové úrovně (podmiňuje se bonusovou třídou pojištěnce). Aby bylo možné výpočet provést, předpokládá se ergodicita celého systému, a skutečné rozdělení pojištěnců mezi bonusovými třídami se nahrazuje rozdělením stacionárním.

Vlastní přínos diplomantky je obsažen ve čtvrté kapitole, kde je představený teoretický postup aplikovaný na data konkrétní české pojišťovny. Nicméně skutečná pravidla bonus-malus systému v České republice je potřeba poněkud modifikovat, aby bylo možno je modelovat pomocí Markovského řetězce. Diplomantka uvádí dvě z těchto možných modifikací a odvozuje optimální bonusové sazby pro tyto dva modely a alternativní model s výraznější penalizací. V závěru komentuje příčiny značné odlišnosti teoretických optimálních bonusových sazeb a reality.

Práce je zpracována velmi pečlivě, s dobrou úpravou a logickým uspořádáním, překlepů je minimum.

Problémem nicméně je styl psaní typu "vyprávěnka". Popis pravděpodobnostního modelu, se kterým se pracuje, je volně roztroušen po celých prvních dvou kapitolách, předpoklady jsou uvedeny tu v mimochodné poznámce, tu explicitně, někdy není zřejmé, které všechny předpoklady tedy v dané situaci máme, tvrzení (někdy triviální někdy netriviální) jsou volně konstatována v textu, aniž je zřejmo, do které kategorie tvrzení spadá a v druhém případě, proč by mělo platit, nebo kde lze dohledat důkaz daného tvrzení. Častokrát není z textu zřejmé, co je tvrzení, co je předpoklad, co je důsledek, a jak to všechno spolu souvisí. Takovýto přístup je u diplomové práce na matematice na MFF UK značně nevhodný! Matematická úroveň práce je tím velmi snížena a to navzdory faktu, že odvození do práce zahrnutá jsou korektně provedená a bez chyb. Nicméně některá další odvození (která absolvent magisterského studia musí umět odvodit sám) v práci chybí a u aplikací složitějších vět z jiných zdrojů lze pochybovat o tom, že je autorka opravdu provedla, když odkazuje na konkrétní použitou větu odkazem na knihu o několika stech stran. Spíše to působí jako přepsání faktů z hlavního použitého zdroje bez jejich podrobnějšího pochopení. (Konkrétní dotazy a připomínky viz níže – minimálně na dotazy označené * by bylo dobré reagovat během obhajoby.)

Rovněž ve čtvrté kapitole by zpracování mohlo být v některých částech trochu stručnější. Popis pojmenovávání jednotlivých proměnných v R a vypisování příkazů, jimiž voláme odhadovací funkci zobecněného lineárního modelu, patří nanejvýš do přílohy diplomové práce.

Na druhou stranu analýza ve čtvrté kapitole je provedena pečlivě a je podrobně popsána, výsledky jsou dobře komentovány, v návržení modifikace bonus-malus pravidel tak, aby vedly na Markovský řetězec, musela diplomantka uplatnit vlastní invenci.

Mohu tedy konstatovat, že práce splňuje požadavky kladené na diplomovou práci na MFF UK a proto ji doporučuji jako takovou uznat.

Dotazy a připomínky:

* str. 5, ř.3 – Opravdu tvrdíte, že fakt, že výskyty pojistných událostí tvoří Poissonův proces, vyplývá z předpokladů v daném odstavci?

značení v kapitole 1 a 2 – n značí pokaždé něco jiného – jednou počet let a jednou počet pojištěnců, to by bylo dobré odlišit.

str. 11, pozn. 2 – Opravdu? Vždy se dostanete z jednoho stavu do druhého? Pro každé ω z pravděpodobnostního prostoru?

* str. 12, ř. -9 – Proč je ten řetězec definovaný pomocí (2.2) homogenní? A on může být nehomogenní, pokud ho budete "uvažovat" nehomogenní?

str. 14, ř. 7 – Díky kterým předpokladům? Jak zní věta, na základě které tvrdíte (2.4)?

* str. 14, ř. -7 až -5 – Tvrzení, že matice pravděpodobností přechodu nerozložitelného řetězce splňuje definici 2 je zřejmé? Definice 2 vyžaduje nenulové pravděpodobnosti přechodu pro nějaké konkrétní n_0 a všechny kombinace stavů najednou.

důkaz věty 1 – Odkaz na důkaz konkrétního tvrzení nelze řešit odkazem na celou knihu! Navíc tento důkaz je i v základním použitém zdroji [5].

kapitola 3.1 – Náhodná veličina Θ představená v kapitole 1 zcela jistě není relativní výše pojistného. Kdyby byla relativní výší pojistného, proč by měla být škodní frekvence $\Lambda\Theta$? To by se pojištěný, podle toho, kolik mu určí pojistné, rozhodoval bourat více nebo méně? Rizikovitost a relativní výše pojistného jsou samozřejmě dvě různé věci (viz též [5], str. 185).

* tvrzení kolem vzorce (3.1) – Proč? Proč by Λ nemohla nabývat jiných hodnot než těch několika $\hat{\lambda}_1, \dots, \hat{\lambda}_K$? To je opět váš pravděpodobnostní model pro zkoumanou situaci. A teď, neboť používáte Bayesovský přístup, předpokládáte (tedy vy jako autor modelu jste si zvolila), že rozdělení Λ bude odpovídat empirické distribuční funkci Λ pro náhodně zvoleného pojištěnce z pojistného kmene, kterou jste odhadla z dat. Dotaz – ten náhodný pojištěnec by se v tomto případě vybíral náhodně jak? Pokud by měl mít ono odpovídající empirické rozdělení Λ ?

str. 18, odstavec druhý – Proč je argument minima kvadratické ztrátové funkce roven podmíněné střední hodnotě? To je tvrzení? Zřejmé nebo ne? Nemá pro tento případ kniha [5] nějakou větu?

* str. 21, (3.9) – To byste neuměla odvodit?

str. 31, ř.1 – Jaké je rozdělení testové statistiky, když je nestandardní?

RNDr. Michaela Prokešová, Ph.D.