Posudek oponenta diplomové práce

Zuzana Skřivanová:

Stochastické modelování úmrtnosti pro více populací

Předložená diplomová práce se zabývá chováním doporučovaných stochastických modelů úmrtnosti (Lee-Carter, Renshaw-Haberman a Cairns-Blake-Dowd) v případě současného modelování úmrtnosti více navzájem korelovaných populací. Tomu odpovídá i struktura práce: po popisu zmíněných modelů pro „jednorozměrný“ případ následuje popis jejich rozšíření na případ korelovaných populací a vše je v závěrečné kapitole numericky aplikováno na příkladě společného modelování české a slovenské populace. Hned v úvodu si dovoluji podotknout, že jsem byl již oponentem bakalářské práce téže autorky s podobným tématickým zaměřením.

Řešená problematika je velice aktuální a numerický příklad je vhodně zvolený (stupeň korelovanosti úmrtnostního chování populace v ČR a SR je nepochybně zajímavý problém pro řadu institucí).

V práci jsem nenašel žádné systematické chyby. Výklad je srozumitelný a logicky uspořádaný. Grafická úroveň práce je dobrá. Ekonometrickou analýzu provedenou v praktické části práce považuji za velmi kvalitní a zajímavou. K práci mám nicméně řadu připomínek a dotazů:

$5_{2\text{-}1}$: Vztah (1.2) měl být rozdělen do dvou, přičemž pro platnost prvního z nich je nutné předpokládat existenci hustoty pravděpodobnosti.

7^3: Věk sledovaných může je případně větší než ω, neboť poslední věková kategorie v úmrtnostní tabulce znamená „věk ω a více“.

8^8: Formulace „za předpokladu konstantní intenzity úmrtnosti“ není jasná.

8_1: Označení matice uvedené na str. 9 jako „matice pravděpodobností úmrtí ve věku x“ není správné a je matoucí.

9^{12}: Proč cílem musí být *hladká* křivka?

12_3: Proč věky $x_1, \ldots, x_m$ nemusí být po sobě jdoucí (jako roky $t_1, \ldots, t_n$, u nichž navíc dostupnost dat o úmrtnosti ze souvislého časového bloku může představovat v praxi problém)? Kromě toho použité sumační symboly (např. v (2.2) na str. 12) vypadají, jako by se sčítalo přes $x_1, x_1 + 1, \ldots, x_m - 1, x_m$.

$15_{2\text{-}1}$: Proč má uvedená optimalizační funkce vyhlazující účinek a jak se volí koeficient p v praxi?

2.1-2.3: V popisu odhadových procedur pro všechny tři uvedené modely není ani zmínka o statistických vlastnostech získaných odhadů: to je záměrně? Navíc složitost jejich

analytického vyjádření (viz komentář na str. 12) tedy znamená, že i intervalové předpovědi úmrtnosti, které jsou v praxi (na rozdíl od bodových) velmi žádané, se konstruují pouze simulačně?

3.5.1-2: Proč se v některých paragrafech, jako je např. 3.5.1 a 3.5.2, předpokládá (bez odůvodnění) normalita (je hlavním důvodem možnost aplikace metody maximální věrohodnosti)?

31_1: … pro *logaritmus* věkově specifické míry úmrtnosti …

$33^{4\text{-}6}$: Co jsou standardizovaná Pearsonova rezidua, devianční rezidua a (nestandardizovaná?) Pearsonova rezidua a proč byla právě tato rezidua zvolena pro jednotlivé modely?

54: Dle mého názoru ze shrnutí provedeného v závěru diplomové práce vyplývá značná závislost výsledků (především demografických projekcí) na volbě modelu pro více populací; zastává diplomantka stejný názor?

− Domnívá se autorka, že získané závěry týkající se korelované úmrtnosti jsou obecné nebo jsou specifické pouze pro česko-slovenskou populaci (viz také následující dotaz)?

− Následující dotazy jsou „nepovinné“, neboť nesouvisí s matematikou, ale úzce se dotýkají praktické problematiky řešené v diplomové práci. Má nějaké racionální vysvětlení, že „od roku 1990 je specifická míra úmrtnosti ve sledovaných věcích v české populaci nižší“ (viz str. 35: přitom se jedná o věky 60 a 70 let na obr. 4.6, které jsou významné např. z hlediska starobních penzí)? Jakým způsobem je registrovaná úmrtnost Slováků žijících v ČR (jedná se patrně o velmi významné počty, které mohou mít vliv na populační demografii, přičemž dle mého názoru zde proti sobě působí počet let aktivně strávených v ČR a datum získání občanství ČR)?

Závěr: Předloženou práci doporučuji uznat jako práci diplomovou.

15. 8. 2016 Prof. RNDr. Tomáš Cipra, DrSc.