

## Posudek oponenta na diplomovou práci **Gabriely Myšákové: Metody analýzy sezónnosti demografických jevů.**

Předkládaná práce je zaměřena na demografickou aplikaci metod analýzy časových řad. Výchozími daty byly měsíční údaje o narozených, zemřelých a sňatcích za Českou republiku a vybrané evropské státy, které jsou přehledně uvedeny v subkapitole data. Byly použity tři metody analýzy časových řad: dekompozice časové řady umožňující rozložení časové řady na jednotlivé složky, metoda X12-ARIMA, kterou se navíc počítají proměnlivé sezónní indexy a testy a metoda kointegrace časových řad hledající vztah mezi dvěma časovými řadami.

V úvodní části je představena literatura zabývající se sezónností porodnosti, sňatečnosti a úmrtnosti. Jsou diskutovány studie za Českou republiku/České země i ze zahraničí včetně dalších rozmanitých aspektů, jako je věk, rasa, či vůbec celkový kulturní kontext. Vzhledem ke specifčnosti tématu následuje rozsáhlejší třetí kapitola věnovaná metodologii. Autorka zde vytyčuje hlavní charakter, ale i specifčnost časových řad a jejich členění. Začíná s jednoduchou standardizací dat na stejné počty dnů a v případě sňatků i na počet pátků a sobot v měsíci. Následuje pak popis složitějších přístupů: dekompozice časové řady, metoda X-12ARIMA a kointegrace časových řad. Tyto metody jsou srozumitelně popsány včetně popisu jejich praktického výpočtu v softwarovém systému SAS. Na tomto místě mi nepřipadá příliš logické popsat první aplikaci pomocí Enterprise Guide a další dvě přímo pomocí programovacího jazyka SAS. Uvedené procedury jsou součástí modulu SAS/ETS což mělo být uvedeno na počátku popisu praktického výpočtu pomocí SAS. Chybně je uvedeno Base (str.38 obr.4; str.43 obr.6). Pokud jde o grafické výstupy není třeba je dělat v excelu když SAS nabízí nejen modul SAS/GRAPH nebo možnost lepší grafiky pomocí *ods graphics on*. Rovněž v poslední metodologické části popisující shlukovou analýzu v aplikaci SAS by bylo vhodnější uvést tři procedury: *Distance*, *Cluster* a *Tree* a jejich zařazení do modulu SAS/STAT; na obr.7 str.47 je chybně uvedeno Base. Dendrogram není pro tři skupiny („vykreslení pro 3 clustery“ str.47), ale v obr.7 příkaz *nclusters=3* znamená, že se uloží příslušnost ke skupině v datovém souboru *new* jako další/nová proměnná. Toto (*nclusters=3*) nemuselo na obr.7 vůbec být, protože přímo nesouvisí s popisovanou problematikou. V kapitole věnované výsledkům shlukové analýzy měla být lépe popsána vstupní data. Není mi jasná metodika výpočtu sezónního indexu, který se podle obr.17 vztahuje k celému období, tj. do seskupovací analýzy by vstupovala pouze 1 proměnná (podobně u úmrtnosti)? U dalších dvou dendrogramů předpokládám, že proměnnými byly příslušné indexy (složky) za každý kalendářní rok 2001-2009. Práce obsahuje některé nepřesnosti nebo překlepy (str. 16, 20). Strana 12: nedomnívám se, že narození je nevyhnutelná událost, místo potrat mělo být uvedeno interrupce nebo umělé přerušování těhotenství (vyhnutelná událost) a nevím zda je možné se vyhnout každé nemoci. Pokud je o hladinu významnosti (str.39 konec druhého odstavce) je obvyklejší uvádět 0,001 a nikoliv 0,1%. Nevhodné je použití výrazu cluster, clustery v textu na str.46, 47 má být skupina, proto se používá i termín seskupovací analýza vedle termínu shluková.

Těžiště práce lze spatřovat ve dvou dimenzích a to ve zvládnutí v demografii netradičních pokročilých statistických metod a v jejich aplikaci nejen na Českou republiku, ale i na řadu dalších evropských zemí. Hlavní výsledky jsou prezentovány v kapitolách 4 (porodnost), 5 (sňatečnost), 6 (úmrtnost). Tyto analýzy jsou originální a přinášejí nové možnosti studovat vývojové trendy v podrobnějším členění (podle kalendářních měsíců). Nové pohledy mají zejména význam pro Českou republiku od devadesátých let dvacátého století, kdy se rychle měnilo populační klima. Měsíční trendová řada (očištěná od sezónního kolísání) umožňuje lépe specifikovat dopad konkrétních opatření či situací na demografické jevy. Navíc je možné zkoumat eventuální změny v sezónnosti (proměnlivé sezónní indexy) včetně jejich rozpětí (intenzity). Přes uvedené výtky, které zdánlivě zabírají značnou část posudku, lze označit předloženou magisterskou práci Gabriely Myšákové za nadstandardní. Přináší nové výsledky, které jsou nejen užitečné pro hlubší poznání, ale také významné po praxi. Cenná je i datová dokumentace uvedená na příloženém CD. Doporučuji práci k obhajobě.

V Praze dne 25.5.2011

Prof. RNDr. Jitka Rychtaříková, CSc.  
Katedra demografie a geodemografie