

## POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

**Název:** Structural Equation Models with Application in Social Sciences

**Autor:** Václav Veselý

### SHRNUTÍ OBSAHU PRÁCE

Diplomová práca študenta Václava Veselého sa venuje tzv. SEM (Structural Equation Models) a EIV (Errors in Variables Models) modelom. SEM modely majú v štatistike široké uplatnenie v prípadoch, keď je potrebné modelovať pomerne zložitú, niekoľkoúrovňovú (lineárnu) závislostnú štruktúru v dátach a špeciálne sú populárne v situáciach, keď sa dá očakávať, že závislá premenná z jedného parciálneho modelu vystupuje ako nezávislá premenná v inom parciálnom modele. Naproti tomu, EIV modely slúžia k modelovaniu (lineárneho) vzťahu medzi závislou a nezávislou premennou, ak sú obe premenné zaťažené náhodnou chybou.

Autor prvé dve kapitoly svojej diplomovej práce venuje práve týmto dvom typom štatistických modelov: popisuje niektoré základné vlastnosti, metódy odhadovania neznámych parametrov a spôsob použitia v praktických aplikáciach. Tretia kapitola prináša inovatívny nápad, ktorý kombinuje ideu SEM a EIV modelov a predstavuje štatistický prístup, pomocou ktorého sú neznáme parametre v SEM modeli odhadované pomocou EIV nástrojov (tzn., že v SEM modeli predpokládame, že závislé aj nezávislé premenné sú merané s nejakou náhodnou chybou). V záverečnej štvrtej kapitole autor využíva navrhnutú metódu kombinácie SEM a EIV modelov na analýzu vlastných sociologických dát o vzťahovom a sexuálnom uspokojení partnerov (sociologický dotazník medzi ženami v ČR a na Slovensku).

Z teoretického hľadiska hodnotím prácu ako náročnú: pochopenie problematiky SEM aj EIV modelov vyžaduje nadštandardné matematické znalosti a schopnosť porozumieť netriviálnym štatistickým výsledkom. V práci autor navyše odvodil niektoré vlastnosti pre navrhované modely, ktoré kombinujú SEM a EIV prístup a prezentuje algoritmické postupy pre odhadovanie neznámych parametrov v takýchto modeloch. Niekoľko ďalších zaujímavých teoretických výziev v tejto súvislosti je načrtnutých ohľadom s možného budúceho výzkumu.

Na druhej strane, bohužiaľ, považujem prácu z formálneho hľadiska za príliš odfláknutú a zbytočne povrchnú: obrovský počet chýb, preklepov, nejasných formulácií, formálnych nepresností, nedostatkov v značení, nekonzistencii písaného textu a nekvalitná angličtina degradujú nadštandardnú teoretickú úroveň práce na pomerne nízky celkový dojem z jej vypracovania.

Na niektorých miestach by si práca taktiež zaslúžila výrazne podrobnejší teoretický základ a zdôvodnenie, než len stručný odkaz na doplnujúcu literatúru.

Prácu doporučujem uznať ako diplomovú prácu na MFF UK.

### PRIPOMIENKY K OBHAJOBE

- Definujte a vysvetlite veličiny  $\mathbb{Y}$ ,  $\mathbb{Y}_j$ ,  $\mathbb{Y}_{(j)}$  a  $\mathbf{Y}_j$  zo str. 9. Čo je  $y_j$  a  $\mathbf{y}_j$ ?
- Vysvetlite vetu "Constraints are simply setting some parameters to zero according to specified model." na str. 10. (Aké obmedzenia? Ktoré parametre? Ktorý model?)
- Vysvetlite vetu "If (2.3) has no solution then the estimator is a fixed matrix." na str. 12. (Aký odhad? Aká pevná matica a ako vyzerá?). Je  $\beta$  vektor, alebo matica? (vid' (2.2) vs. (2.3))

- ❑ Vysvetlite vetu "Once the estimate  $\hat{\beta}$  is found..." na str. 14. Vo výraze (2.6) sa nevyskytuje žiaden parameter  $\hat{\beta}$ .
- ❑ Vysvetlite, ako boli získane štyri odhady na Obr. 2.1 na str. 14.
- ❑ Vysvetlite vetu: "Let rows of the data  $[\mathbb{X}, \mathbb{Y}, \boldsymbol{\epsilon}_x^\top, \boldsymbol{\epsilon}_y^\top]$  be independent identically distributed with generic distribution of random variables  $[\mathbf{y}^\top, \mathbf{x}^\top, \boldsymbol{\epsilon}_x^\top, \boldsymbol{\epsilon}_y^\top]$ ."
- ❑ Vysvetlite, čo znamená  $EIV([\mathbb{X}, \mathbb{Y}])$ .
- ❑ Ako je definovaná matica  $G$  na str. 17? O aké obmedzenia sa jedná?
- ❑ Čo je  $\mathbf{X}$  v rovnici (3.5)? Ako je definované?
- ❑ Čo je  $\mathbf{Y}^0$  a  $\mathbf{Y}^\epsilon$  vo vzťahu (3.7)?
- ❑ Čo je  $\boldsymbol{\epsilon}_{\mathbb{Y}_j}$  vo Vete (Theorem) 8?
- ❑ Vysvetlite rozdiel medzi modelmi označenými ako  $lm1$  a  $lm2$  v Kapitole 4.
- ❑ Na základe čoho boli modely prezentované v Tabulkách 4.7, 4.8 a 4.9 "zredukované"?

Praha, 30.05.2018



RNDr. Matúš Maciak, Ph.D.  
 maciak@karlin.mff.cuni.cz