



POSUDEK MAGISTERSKÉ DIPLOMOVÉ PRÁCE

Posudek vedoucího práce

Autor práce: Bc. Filip Rybín

Název práce: Sampling Fusion: Ekvivalence měření tří módů sběru dat

Autor posudku: Jiří Vinopal

Návrh klasifikace práce: *velmi dobře*

KONCEPČNÍ STRÁNKA PRÁCE

Má práce jasně formulovaný záměr, cíl, výzkumnou otázku a odpovídají jim závěry?

Cíl práce je formulován jasně hned v úvodu, výzkumné otázky a hypotézy jej přehledně upřesňují a dosažené závěry jim odpovídají.

Má práce adekvátně stanovenou a jasně vyloženu metodu a postup řešení?

Základní postup i konkrétní metody řešení jsou stanoveny adekvátně cílům práce i povaze dat a v práci jsou dostatečně podrobně vysvětleny a zdůvodněny.

Jsou jednotlivé části práce (zejm. teoretická a empirická) vyvážené a vhodně propojené, vytváří text konzistentní celek?

Práce má charakter kompaktní výzkumné studie, která při řešení praktické metodologické otázky vychází z koncepcí a teorií výzkumných chyb, aplikuje je na vlastní provedený dotazníkový průzkum a následně vše analyticky završuje.

Jsou argumentace a text práce a vystavěny logicky, jasně, srozumitelně? Má práce přehlednou strukturu?

Práce má přehlednou strukturu a celkově je hlavní idea práce vyložena logicky a srozumitelně. K nejasnostem dochází na úrovni některých dílčích vyjádření, viz dále.

OBSAH PRÁCE

Hodnocení kvality teoretické části práce (relevance k tématu a záměru práce, šíře a hloubka, vhléd autora do problematiky, samostatnost, originalita zpracování,...).

Teoretická část práce se z povahy věci zabývá problematikou výzkumné chyby a jejího měření. Je třeba připomenout, že se jedná o oblast netriviální, pojmově a konceptuálně velmi spletitou (obzvláště při zaměření na měření chyb) a na úrovni magisterského studia poměrně obtížně zvládnutelnou. Proto oceňuji, do jaké míry se v ní autor dokázal zorientovat a že ji pro účely svého výzkumu pojednal v dostatečné šíři, hloubce i přesnosti.

Ocenit je třeba také samostatnost autora při volbě tématu, jeho uchopení a zpracování a současně schopnost propojit vysoce teoretickou úroveň kvality měření s řešením praktického úkolu běžného výzkumného procesu v oblasti výzkumu veřejného mínění a politického chování.

Domnívám se, že větší hloubce a úplnosti by pomohlo zakotvit vše v konceptu Total Survey Error (s větším důrazem na rozlišení systematické a nesystematické povahy většiny druhů chyb) a následně se držet obecnějšího smyslu koncepce Mean Square Error jako nikoli chyby jednoho výzkumu, nýbrž chyby výzkumného designu. O ten totiž i při volbě modu v jakékoli praktické situaci jde – o odhalení a minimalizaci (nebo alespoň evidenci) chyb daného designu při jeho jakémkoli opakování; nikoli (pouze) o změření průměrné chyby v rámci jednoho aktuálního šetření. Koncepce TSE je koneckonců právě takto zamýšlena.

Hodnocení kvality empirické části práce (adekvátnost použitých metod a postupů vzhledem k cíli práce, správnost aplikace metod a postupů, prezentace a interpretace výsledků,...).

Je nasnadě, že pro cílené ověřování ekvivalence modů by bylo vhodné konstruovat precizní výzkumný design tak, aby řada efektů, které (jak autor správně reflektuje) interferují s hlavními sledovanými, byla eliminována, resp. co nejvíce a priori sjednocena. Nicméně luxus takového šetření nebývá k dispozici příliš často. Využít data, která jsou dostupná, je často jediné možné řešení, v jistém smyslu o dost náročnější na možnosti použitých metod a postupů. Osobně považuji použité postupy a metody za adekvátní cílům práce a dostupným datovým zdrojům, a to už na úrovni jednorozměrné analýzy a porovnávání s externími kritérii. Logistické modely pak poskytují vhléd do hlubší povahy rozdílnosti modů a velice oceňuji jejich zařazení jako dobrého věcného obohacení, prohloubení pochopení a nástroje pro argumentaci závěrů. (A, koneckonců, modely také povyšují metodologickou a analytickou úroveň řešení diplomového úkolu.)

Ocenil bych detailnější zdůvodnění pro zařazení posledního modelu pouze s proměnnými politické participace.

V kapitole 4.3 Doprovodné analýzy módů podle mne není šťastně udělána tabulka návratnosti a následně ani komentáře o efektivitě. Kdo jsou v tabulce nekontaktovaní? Počty získaných dotazníků vycházejí ze stejně velkých počtů distribuovaných, ze stejně velkých cílových počtů?

Hodnocení diskuse a závěrů práce (soulad interpretací a závěrů s výsledky empirických analýz, provedení diskuse závěrů, snaha o vysvětlení závěrů a jejich argumentace, zodpovězení výzkumných otázek, naplnění cílů práce).

V první řadě, diskusi bych očekával ještě před závěrem, a nikoli naopak; v závěrech bych pak (mimo jiné) očekával shrnutí a výsledek diskuse.

Samotná diskuse se týká zjištěných výsledků daného šetření a možností jejich zdůvodnění některými koncepty z teorie výzkumných chyb. To je jistě v pořádku; nicméně ocenil bych také myšlenky, závěry, návrhy... na obecnější úrovni, které by se pokusily zjištění (v kontextu koncepcí teorie výzkumných chyb) více zobecnit.

Což by ostatně také lépe vedlo k případným doporučením pro další volby modů. V dané chvíli mi totiž přijde nepodložené např. tvrzení, že „Výsledky tedy poukazují na potřebu sběru dat skrze smíšené módy za účelem minimalizace jednotlivých chyb, kterými jsou dílčí módy více či méně zatíženy. Lze tak využít silných stránek, které módy skýtají a minimalizovat jejich nevýhody.“ Nezpochybňuji, že kombinace modů může mít pozitivní efekty, nicméně výsledky samotné práce mi k tomu podporu neposkytují; ostatně, tímto směrem (ověřování kvality smíšeného modu) se neubíraly ani autorovy výzkumné cíle v této práci.

FORMÁLNÍ ASPEKTY PRÁCE

Hodnocení naplnění kritérií odborného textu

Předložená práce splňuje nároky kladené na odborný text diplomové práce.

Hodnocení vhodnosti a úplnosti využití informačních zdrojů (literární prameny, databáze, zahraniční literatura, datové zdroje,...)

Autor používá vhodné literární i datové zdroje a ve vhodném rozsahu.

Hodnocení jazykové, stylistické a grafické úrovně práce (grafická úprava, členění textu, označení příloh, zpracování tabulek, grafů, schémat atp.)

Jazyková, stylistická ani grafická úroveň textu není nijak vážně narušena. Objevují se občasné překlapy.

Další poznámky

Na několika místech práce se objevuje nejasná, nepřesná nebo možná až chybná vyjádření nebo výrazy. Jsou to sice dílčí a nikoli snad zásadní momenty, ale vzhledem k tomu, že je jich dohromady v práci více, poukazují na tuto věc jako na obecnější nedostatek. Namátkou:

S 9. „Ekvivalence teoretického konceptu je tématem pro srovnání studií napříč zeměmi,...“ Spíše kulturami nebo populacemi, nikoli nutně zeměmi.

s. 11 „...tím pádem nezískáváme naměřené hodnoty náležící ke skutečné pravdivé hodnotě.“ Myslí se tím skutečné pravdivé hodnotě zamýšleného jevu. K nějaké skutečné pravdivé hodnotě se asi dostáváme, ale jiného jevu...

s. 12 „Následující čtyři chyby pokrývají chyby ovlivňující statistické operace na rozdíl od jednotek výběru.“ Nejasné vyjádření...

Autorovi bych také vytknul ne úplnou důslednost v používání terminologie. Ona je sama o sobě v oblasti kvality výzkumu a výzkumných chyb komplikovaná a nejednoznačná, pro neobeznámeného čtenáře pak je tedy nutné alespoň důsledně dodržovat jednu v rámci práce. Autor přitom poněkud míchá pojmy cílová skupina, základní populace, výběrová populace, oslovená populace atp. viz. např. na s. 12 - 13.

S 10. „Rámec vzorku“ ve schématu, odpovídá anglickému sampling frame (pokud se nepletu). Doporučuji tedy držet se standardního českého termínu opora výběru. A podobně, pro sampling error se v češtině již dávno vžil termín výběrová chyba...

Několikrát je použit termín „očekávané hodnoty“, který ale nebyl v práci definován a čtenář si jen musí domyslet (a věřit, že správně), že jde o totéž, co se na jiných místech označuje jako "skutečná hodnota".

Sám asi dovedu intuitivně porozumět sdělení na s. 13.: „Pokud se ale nejedná o systematické odchylky, je chyba způsobena odlišnými charakteristikami jednotlivých respondentů z výběrové populace, která může být výzkumem oslovena a která nemůže být výzkumem oslovena, vyřešena skrze agregaci statistik na celý výzkumný vzorek.“ Nicméně pro neobeznámeného čtenáře se fakticky jedná o konstatování zavádějící, nebo přinejmenším málo srozumitelné – o jakou agregaci na jaký celý výzkumný vzorek se zde jedná, když mluvíme o části populace, která není ani pokryta oporou výzkumu?

U příkladů systematických a nesystematických chyb (s. 14-15) chybí dokončení, tj. vysvětlení mechanismu, jak se na agregované úrovni chyby vyruší, resp. proč dojde k posunu výsledku. Uvedený příklad vždy zůstává na úrovni individuální odpovědi jednoho respondenta, což není pro ilustraci systematickosti/náhodnosti chyb na úrovni celého výzkumu adekvátní.

Co autor myslí „makro“ hlediskem ve větě na s. 18: „Oba tyto přístupy jsou tedy využívány a jejich přesnost lze z makro hlediska považovat za *srovnatelnou*.“

s. 18 „Klíčovým tématem, které rezonuje akademickou a výzkumnou sférou a je úzce spojené s módy sběru a typy výběru, je postupné globální snižování návratnosti a s tím spojené riziko nárůstu systematické chyby spojené s výpadkem návratnosti, zejména u pravděpodobnostních výběrů.“ U nepravděpodobnostních (CAPI, PAPI) nikoli? Proč?

s. 29 Nerozumím podtržené části věty: „První přístup staví ekvivalenci měření na úroveň analýz prvního a druhého stupně, zaměřuje se na základní otázku, zdali jsou data sebraná různými způsoby (ať už v případě sběru mezi zeměmi, v čase či skrze jiný mód) srovnatelná...“

Nerozumím vyjádření na s. 49: „Je ale jasné, že pokud zohledníme situaci a voličskou strukturu v období měření (září 2020) a nikoli reálné volební výsledky z roku 2017, změní se přesnost jednotlivých módů sběru dat.“

Celkové hodnocení práce

Celkově hodnotím diplomovou práci Filipa Rybína jako zdařilou empirickou studii na téma ekvivalence módů sběru dat. Oceňuji dobré zakotvení v teorii kvality měření a výzkumných chyb, originální a samostatný přístup a dobrou praktickou využitelnost závěrů.

Otázky a náměty k obhajobě

Řada otázek vyplývá z poznámek uvedených výše, zde bych doplnil ještě jednu:

Autor formuluje závěr, že „Výsledky tedy poukazují na potřebu sběru dat skrze smíšené módy za účelem minimalizace jednotlivých chyb, kterými jsou dílčí módy více či méně zatíženy. Lze tak využít silných stránek, které módy skýtají a minimalizovat jejich nevýhody.“ Ovšem nedoje takto spíše (také) ke kumulaci nevýhod?

Datum: 22. 8. 2021

Podpis: