

POSUDEK BAKALÁŘSKÉ PRÁCE

Typ posudku: *oponent*

Autorka práce: Anna Krásná

Název práce: Pokles pití alkoholu a kouření v ČR v letech 2010-2018 ve vztahu k rodinnému zázemí

Vedoucí práce: PhDr. Ing. Petr Soukup, Ph.D.

Oponent: Ivan Petrůšek

Navržené hodnocení: C

Zdůvodnění Vašeho hodnocení práce (zejména výhrad a kritických připomínek) rozepište podle níže uvedených kritérií.

1. Je cíl práce (výzkumná otázka) jasně formulován a odpovídají mu závěry? Je práce vhodně strukturována?

Cíl práce je jasně formulován. Výzkumná otázka si klade za cíl zkoumat kauzalitu („*Jaké rodinné faktory stojí za tím, že mezi lety 2010 a 2018 nastal u mladých lidí v ČR pokles užívání alkoholu a kouření?*“, s. 11). Typ analyzovaných dat (průřezová data z opakovaných dotazníkových šetření) a ani použité metody však kauzalitu analyzovat neumožňují. Věcné závěry práce tudíž nejsou spolehlivé. Jak popisují níže, práce má několik zásadních nedostatků z hlediska argumentace a prováděných analýz, což významným způsobem zpochybňuje závěry práce. Ambiciózní cíl bakalářské práce tedy nebyl naplněn.

Práce je vhodně strukturována.

2. Opírá se autorka o dostatek literatury relevantní tématu práce? Využívá i cizojazyčné texty?

Ano, autorka se opírá o dostatek relevantní literatury. Většina citovaných zdrojů je v anglickém jazyce. Anglické zdroje jsou vhodně doplněny českými texty.

3. Jaká je kvalita použitých dat nebo zdrojů (včetně jejich výběru, sběru a popisu) a metod jejich analýzy?

Bakalářská práce je založena na sekundární analýze kvalitních dat z opakovaných průřezových šetření *Health and Behaviour in School-aged Children* (HBSC). Autorka využívá metody bivariační analýzy, které jsou někdy realizovány uvnitř tří let (2010, 2014, 2018), a jindy na sloučeném datovém souboru. Autorka ne vždy vysvětluje, proč v dané situaci volí jednu z těchto strategií. V práci nejsou popsány rozsahy výběrů v jednotlivých letech šetření. Autorka při testování hypotézy číslo 6 používá lineární regresní model (jako nejkompexnější metodu, která pracuje současně s několika vysvětlujícími proměnnými). Dva regresní modely neobsahují indikátorové proměnné pro rok konání šetření, a tudíž nezohledňují potenciálně různé úrovně pití alkoholu a kouření v jednotlivých letech (což je vlastně tématem bakalářské práce). Jako nekonzistentní považují při testování některých hypotéz používání neparametrických metod, při testování jiných hypotéz naopak používání parametrických metod: všechny analýzy přitom pracují se stejnými proměnnými, které se nakonec potkají v parametrickém regresním modelu. Zároveň oceňují transparentnost autorky v podobě uvedení kompletní syntaxe realizovaných analýz v příloze č. 1.

4. Jaká je kvalita argumentů, o něž autorka opírá závěry, k nimž dospěla?

Kvalita autorčinných argumentů je průměrná. Autorka deklaruje zkoumání změny v čase (pokles užívání alkoholu a kouření; pokles v *family affluence scale*), ale změnu nezkoumá. Změna znamená, že stejní jednotlivci, případně stejné kohorty definované rokem narození, „upraví“ v čase míru užívání alkoholu a kouření. Autorka práce změnu nezkoumá: práce pouze srovnává úroveň užívání alkoholu a kouření ve věkové skupině 13-15 let ve třech časových bodech (2010, 2014, 2018).

Hypotéza H1 předpokládá souvislost mezi komunikací s otcem a pitím alkoholu dětí. Zdůvodnění, proč se autorka zaměřuje pouze na komunikaci s otcem (a nezaměřuje se na komunikaci s matkou a také se nezaměřuje na případnou interakci mezi oběma „komunikačními“ proměnnými) není teoreticky ukotveno. O to překvapivější je pak sledovat, že komunikace s matkou je součástí regresního modelu a příslušný regresní koeficient je u komunikace s matkou statisticky významný (a u komunikace s otcem tomu tak v regresním modelu z tabulky 7 není).

Odůvodnění pro formulování hypotézy 2 je také velmi chabé. Evidence prezentovaná v rámci výsledků zobrazuje koláčové grafy pouze za rok 2018 a čtenář tak neví, jak u hypotézy „*U mladistvých se od roku 2010 zvyšuje pocit, že si se svými rodiči mohou promluvit o tom, co je trápí.*“ vypadal pocit, že si se svými rodiči mohou promluvit v dvou předchozích šetřeních.

U hypotézy 3 jsou hodnoty korelací tak nízké, že se prakticky jedná o nulovou asociaci mezi proměnnými. Korelační koeficienty jsou statisticky významné pouze z důvodu, že autorka pracuje s v každém roce s více než 1000 respondentů.

Hypotéza 4 opět nesprávně pracuje s konceptem změny. Na analyzovaných datech lze pouze provést srovnání, jestli se liší socioekonomické zázemí dětí ve věkové skupině 13-15 ve třech analyzovaných letech, a nelze zkoumat „*zda se množství vlastněného majetku zvýšil*“ (s. 14).

Hypotéza 6 není v textu vůbec odůvodněna (ani teoreticky, ani výsledky předchozích empirických výzkumů). Autorka tudíž neformuluje vůbec žádný argument, proč by „*Konzumace alkoholu a cigaret je nejsilněji ovlivněna FAS skórem...*“ (s. 14).

Neodůvodněný je následující závěr z výsledků práce: „*V tabulce je zajímavé, že proměnná pro úplnost rodiny, přítomnost otce a matky v místě bydliště, a kvalita komunikace s otcem vyšly jako statisticky nesignifikantní, tudíž na pití alkoholu nemají vliv.*“ (s. 28) Proměnná *fullfam* představuje interakci mezi přítomností otce a matky v rodině. Mezi všemi třemi proměnnými v regresním modelu existuje multikolinearita, která zvyšuje směrodatné chyby odhadů tří příslušných regresních koeficientů. Jak naznačují prezentované modely, po vyřazení proměnných (*Mother in main home* a *Father in main home*) z modelů předpokládám, že děti z kompletních rodin budou významně méně konzumovat alkohol a kouřit. Citované autorčino konstatování tedy neobstojí.

5. Jsou v práci autorčina tvrzení a zjištění jasně odlišena od tvrzení a zjištění převzatých?

Ano, ve většině textu bakalářské práce jsou autorčina tvrzení a zjištění jasně odlišena od tvrzení převzatých. Pouze v teoretické části místy není jasné, jestli se jedná o tvrzení autorky, nebo jiných autorů (např. první odstavec sekce 1.2 nebo tvrzení typu: „*Školní prospěch se odvíjí od zděděné schopnosti se učit a udržovat pozornost, tudíž vliv má genetická informace rodičů.*“; s. 4).

6. Jaká je úroveň odkazového aparátu, jazyka a dalších formálních náležitostí?

Odkazový aparát je zcela v pořádku. Jelikož nejsem rodilý mluvčí, nedokážu posoudit gramatiku textu. Styl psaní místy (spíše výjimečně) neodpovídá akademickému textu. Celková úroveň jazyka je podle mého soudu na velmi dobré úrovni.

7. Jiné přednosti a/nebo nedostatky, které neodpovídají výše uvedeným kritériím (jsou-li jaké).

Na tomto místě zmíním několik menších nedostatků bakalářské práce:

- Bakalářská práce na mnoha místech deklaruje, že úroveň pití alkoholu a kouření se u mladistvých mezi třemi analyzovanými lety liší. Toto tvrzení je přebíráno z dalších zdrojů a na samotných datech HBSC není téměř analyzováno. Až v samotném závěru práce (s. 30)

autorka prezentuje průměry frekvence užívání alkoholu a kouření. Graf 6 přitom vůbec nenaznačuje, že se četnost kouření ve věkové skupině 13-15 let napříč roky šetření liší.

- Hypotéza 5 postuluje souvislost mezi přítomností biologických rodičů v domácnosti a pitím alkoholu/kouřením. Příslušné otázky z dotazníku (které jsou autorkou značeny jako q1 a q2, s. 16) zároveň uvažují také nevlastní rodiče. Není jasné, jestli analyzované proměnné skutečně pracují výhradně s biologickými rodiči.
- Autorka v sekci výsledků u hypotézy 4 uvádí, že „Pro zjištění vztahu konzumace alkoholu, cigaret a úplnosti rodiny byly proměnné převedeny do dummy proměnných a vyhodnoceny Mann-Whitney *t*-testem.“ Tabulka 5 (a také syntax v příloze) ale prezentují výsledky parametrického *t*-testu pro dva nezávislé výběry.
- FAS nepředstavuje škálu, která měří sociální nerovnosti (jak je v práci nesprávně uvedeno, s. 16). Jedná se o škálu, která měří socioekonomický status rodiny dítěte. Určitě se také nejedná o spojitou proměnnou (jak uvádí autorka).
- Grafická úprava výstupů prováděných analýz není na požadované úrovni. Autorka přejímá výstupy z SPSS bez překladu anglických termínů do českého jazyka. Tabulky těchto výstupů často přesahují okraje textu.
- Tabulka 3 představuje korelační matici mezi dvěma proměnnými, a tudíž je vlastně zbytečná. Stačilo by do textu práce napsat hodnotu příslušného Spearmanova korelačního koeficientu (a ne dvakrát uvádět tuto korelaci v korelační matici, společně s korelacemi dvou proměnných se sebou samými).
- Rozsah výběrového souboru značíme malým písmenem *n* (velké *N* se používá k značení velikosti populace).
- Věk respondentů není v datech HBCS měřen jako spojitá proměnná, ale jako proměnná diskretní.
- Proměnné zobrazené v grafech 1 a 2 jsou ordinální, tudíž pro zobrazení jejich četností by byl vhodnější sloupcový graf (vhodnější než prezentovaný koláčový graf).
- Po podrobném přečtení práce si nejsem jistý, jestli autorka plně rozumí konceptu statistické inference z výběrových dat na cílovou populaci. Na několika místech totiž formuluje úsudek z „výběru na výběr“: „**Nulovou hypotézu tedy nelze zamítnout, a v datovém souboru není pravda, že by rozdíl mezi množstvím vypitého alkoholu a vykouřených cigaret byl významně ovlivněn úplností rodiny.**“ (s. 26).

8. Náměty na diskuzi při obhajobě práce.

Jaké další aspekty „rodinného zázemí“, které autorka v bakalářské práci neměří, mohou korelovat s pitím alkoholu a/nebo kouřením cigaret u mladistvých? Jak budou tyto korelace pravděpodobně vypadat a proč?

Považuje autorka práce regresní koeficienty u klíčových vysvětlujících proměnných za nevychýlené? Jak by se změnilы hodnoty regresních koeficientů klíčových vysvětlujících proměnných v regresním modelu, pokud by jako vysvětlující proměnné model obsahoval proměnné měřící „tlak spolužáků a kamarádů“, skutečnost, jestli „je v domácnosti nějaký dospělý, který kouří“ (s. 2) nebo užívání internetu mladistvými?

Na kolika respondentech jsou při *listwise deletion* vypočteny dva regresní modely? Jaký podíl z celkového počtu respondentů vstupuje do regresních modelů? Jakým způsobem může vyřazování respondentů s chybějícími hodnotami z regresního modelu ovlivnit věcné závěry modelu?

9. Prohlašuji, že jsem se seznámil s výsledkem kontroly originality textu závěrečné práce v systému: [X] Theses [] Turnitin [] Ouriginal (Urkund)

Komentář k výsledku kontroly:

Celková podobnost s dalšími dokumenty byla systémem *Theses* indikována na úrovni pouze 2 %. Po přečtení textu usuzuji, že ho autorka vypracovala samostatně.

Celkové hodnocení práce:

Práci celkově hodnotím jako průměrnou a doporučuji ji hodnotit známkou C. Formální stánky práce (struktura textu, práce se zdroji, data management) odpovídá požadavkům kladeným na tento typ prací a je na velmi dobré úrovni. Jako zásadní slabiny práce považuji kladené výzkumné otázky ve vztahu na typ analyzovaných dat, způsob argumentace ve vztahu na realizované analýzy a deklarované závěry práce. Základní výzkumná otázka (s. 11) mi přijde pro studenta bakalářského studia jako příliš ambiciózní (dle mého názoru by stačovalo zkoumat souvislosti mezi pitím alkoholu/kouřením a rodinným zázemím v jednom časovém bodě). V bakalářské práci vidím nenaplněný potenciál - škoda, že u tohoto typu textů neexistuje možnost *revise and resubmit*.

Datum: 22. 5. 2023

Podpis: