

Posudek oponenta bakalářské práce

Název: Testy nezávislosti ve čtyřpolní kontingenční tabulce

Autor: Andrej Obuchov

Shrnutí obsahu práce

Bakalářská práce Andreje Obuchova se zabývá testováním nezávislosti ve čtyřpolní kontingenční tabulce. Stručně shrnuje testy založené na porovnání parametrů dvou binomických rozdělení (po řádcích nebo sloupcích tabulky), zmiňuje Pearsonův χ^2 test nezávislosti, poté uvádí Fisherův faktoriálový test a Barnardův test. Na konci práce jsou výsledky simulačních studií.

Celkové hodnocení práce

Téma práce. Téma práce je spíše snazší, spoustu poznatků o tomto problému se studenti dozvědí na přednášce Statistika pro FM. I tak toto téma dává dostatečný prostor k samostatnému rozšíření a prohloubení a je pro bakalářskou práci vhodné.

Vlastní příspěvek. Andrej Obuchov popsal základní testy nezávislosti pro čtyřpolní tabulky; str. 3–11 obsahují poznatky probírané na přednášce, str. 12–18 pojednávají o Fisherově a Barnardově testu, které si autor sám nastudoval z literatury. Autor dále sám provedl simulační studii.

Matematická úroveň. Matematická úroveň práce je slabá. Práce neobsahuje žádná odvození ani důkazy. Pasáže o Fisherově a Barnardově testu obsahují nanejvýš vágně zformulovaná slovní zdůvodnění. Autor si nedělá hlavu s přesným významem symbolů, které používá, takže řada vzorců je uvedena špatně (zejména kvůli obrácenému podmiňování).

Práce se zdroji. Citování zdrojů je v pořádku (s výjimkou jistých typografických nedostatků).

Formální úprava. Formální úprava je celkově přijatelná, ale práce trpí značným množstvím překlepů, které svědčí o tom, že autor po sobě text ani pořádně nepřečetl. Gramatické chyby a nedostatky lze do značné míry pochopit a tolerovat kvůli tomu, že čeština není autorovou rodnou řečí. Horší a méně pochopitelné jsou časté chyby v odborných termínech.

Hlavní připomínky

1. Str. 9: Zde je d_X a r_X definováno obráceně; podmiňovat se má stupněm vzdělání, čili hodnotami veličiny Y , nikoli Z (výsledek zkoušky).
2. Str. 10: Jelikož $p_{ij} = P[Z = i \mid Y = j]$ (viz str. 8), tak vztah $p_{2j} = 1 - p_{1j}$ platí vždy (nejen za nezávislosti) a vztah $p_{11} = p_{21}$ je ekvivalentní s $p_{11} = 1/2$. Kvůli tomu statistiky (1.5) a (1.6) vůbec netestují nezávislost.

3. Str. 12: Chybí odvození vztahu pro $\mathcal{P}(n_{11})$ a zmínka o souvislosti s hypergeometrickým rozdělením.
4. Str. 13: Chybí zdůvodnění výpočtu p-hodnoty. Jakou hladinu má výsledný test?
5. Str. 15, Výpočet P_F pod Tab. 1.8: dvě sumy napravo jsou sčítány přes úplně totáž i .
6. Str. 16: Jestliže $\pi = p_{1|1} = p_{2|1}$, pak nutně $\pi = 1/2$ (viz připomínka 2.). S nezávislostí to nemá nic společného.
7. Str. 17: Vysvětlení T_{ij} v (1.15) je naprosto zmatené. Odkud se vzaly statistiky T_S a T_W uvedené dole? Jak je zdůvodnit? Souvisejí nějak s T_{ij} ? Mimochodem, obě tyto statistiky nabývají hodnoty 0, pokud $\hat{p}_{1|1} = \hat{p}_{2|1} = 1/2$, tj. oba sloupce prvního řádku obsahují stejné četnosti.
8. Str. 18: Jak zdůvodníte tuto proceduru pro odhadování neznámého parametru π ? Má to nějaké rozumné vlastnosti? Nevěřím tomu, že hodnota 0.923 je rozumným odhadem parametru π v Tabulce 1.1.
9. Str. 22: Nevěřím, že Tabulka 2.6 je správně. Proč by se síla testu T_r založeného na logaritmu relativního rizika měla při daných alternativách chovat takto? Proč by χ^2 test měl najednou dosahovat o tolik větší síly než všechny ostatní testy?

Ostatní připomínky

- (i) Str. 3: Co se míní výrokem „ \hat{p} jsou empirické hodnoty spočítané z dat“?
- (ii) Str. 3: Zde se blízko sebe objevují pojmy „asimptotické výsledky“ a „asymptotický test“.
- (iii) Str. 4: Poměr šancí je vysvětlen jako poměr poměrů šancí.
- (iv) Str. 4 a 5: „konzistenční“ odhad.
- (v) Str. 6 a 11: Hodnoty „statisik“.
- (vi) Str. 7: „ve vekrotovém označení“.
- (vii) Str. 10, 12, 13: Není nějaký rozdíl mezi „logaritmicou interakcí“ a „logaritmicou transformací“?
- (viii) Str. 11: Buď vysvětlete, co je Yarnoldovo kritérium nebo o něm raději nemluvte.
- (ix) Str. 19: Dvakrát výskyt slova „alogramus“.
- (x) Str. 23: Co myslíte větou „...testovat, jestli hodnoty nějakého multinomického rozdělení jsou rovny právě nějakým určitým číslům“?

Závěr

Práce Andreje Obuchova obsahově odpovídá zadání, ale její provedení nenaplnuje požadavky na bakalářskou práci zejména proto, že neobsahuje žádné důkazy ani odvození předkládaných vzorců a tvrzení a že se v práci vyskytují závažné faktické chyby. Nemohu tedy doporučit její uznání za práci bakalářskou.

doc. Mgr. Michal Kulich, PhD.
 KPMS MFF UK
 8. srpna 2019