

Oponentský posudek diplomové práce
Center-outward ranks and signs and their application in statistical tests
Bc. Veroniky Roubínové

Diplomantka zpracovala práci na téma mnohorozměrných pořadí a znamének a to konkrétně pořadí a znamének založených na konceptu optimálního přenosu míry. Jedná se o téma z posledního desetiletí, které se velmi rychle vyvíjí a každoročně je v něm publikováno mnoho nových výsledků a odvozených metod.

Nejprve se uchazečka musela seznámit s klasickými výsledky, které se objevují již v knize Hájka a Šidáka a v navazujících pracích. Tyto základy jsou shrnuté v první kapitole předložené práce. Poměrně rozsáhlá druhá kapitola již obsahuje i některé vlastní nápady autorky, zejména v oblasti konstrukce pravidelných sítí bodů v jednotkové kouli. Čtenář je nejprve seznámen s pojmem vnitřně-vnějšího (nebo též centrálního) uspořádání v jednorozměrném a vícerozměrném případě. Po motivaci v jednorozměrném případě je představeno přímé zobecnění konstrukce centrálního uspořádání a znamének pomocí optimálního přenosu míry, inspirovaného distribuční funkcí jednorozměrné náhodné veličiny. Pomocí přenosu rozdělení d -rozměrného náhodného vektoru na rovnoměrné rozdělení na d -rozměrné jednotkové kouli je definována centrální distribuční funkce, která však na rozdíl od obvyklé definice nemá hodnoty v intervalu $[0, 1]$, ale její hodnoty jsou ve zmíněné jednotkové kouli. Autorka poté shrnuje základní vlastnosti populační i empirické centrální distribuční funkce. Velkou pozornost věnuje konstrukci empirické verze, která vyžaduje síť bodů na jednotkové kouli co nejlépe imitující rovnoměrné rozdělení.

Ve třetí kapitole se autorka věnuje lineárním pořadovým testům založeným na centrálním pořadí a znaménkách. Nejprve věnuje pozornost asymptotické normalitě testových statistik, kterou odvodila vedoucí práce, a poté popisuje dvouvýběrové a jednovýběrové testy. U dvouvýběrových testů doplnila zejména permutační variantu odvození kritických hodnot v případě, že pro malý počet pozorování není vhodné spoléhat se na asymptotické rozdělení testové statistiky. Část o jednovýběrových testech je do značné míry vlastní prací autorky, byť inspirovanou prací vedoucí diplomové práce a článkem Hallin, Liu and Verdebout zveřejněným zatím na stránkách arxiv.org.

Čtvrtá kapitola je věnovaná přehledné a detailní simulační studii, ve které jsou představené testy důkladně prozkoumány a porovnány s jinými známými dvouvýběrovými a jednovýběrovými testy pro náhodné vektory.

Práce je napsaná velmi čtivě, angličtinou na velmi dobré úrovni a je z ní poznat, že diplomantka pracovala s chutí a porozuměním. Při psaní si nejenom musela přečíst a rozmyslet mnoho nových poznatků, ale také vymyslet a odvodit vlastnosti několika testových statistik, navrhnout vhodné postupy výpočtů, konstrukcí pravidelných sítí v jednotkové kouli v různých rozměrech. Výsledkem je pěkná práce kombinující jak teoretické, tak praktické výsledky, která může sloužit jako vhodný úvod do problematiky pro každého zájemce.

K práci mám jen několik dotazů a poznámek, nejde však o závažné chyby, spíš jenom dotazy vyžadující drobné upřesnění.

- (1) Na straně 7 se píše o lokálně nejsilnějším testu, nenašel jsem ale formální definici tohoto pojmu.
- (2) V části 2.1 hned za definicí 2.1 se nediskutuje možnost $S_{\pm} = 0$, přitom s definice tato hodnota nastat může.
- (3) Popis řazení na konci strany 12 a začátku strany 13 je trochu mlhavý. Tím, že dosud není přesně definováno $R_{\pm, i}^{(n)}$, je těžké si udělat představu, jak přesně jsou pozorování řazena pro záporná znaménka.
- (4) Na straně 15 se mluví o „propíchnuté“ kouli a potřebě dodefinovat $Q_{\pm}(0)$. Znamená to, že $Q_{\pm}(0)$ není dobře definovaná, nebo jde o množinu—vzor 0 v zobrazení F_{\pm} ?
- (5) Nevím, jestli tvrzení o implikaci $n_0/n \rightarrow 0$ za vzorcem (2.2) plyne opravdu z předpokladů na n_R a n_S . Neměl by toto být spíš předpoklad na n_0 ?
- (6) V tvrzení 2.1 jsou nějaké překlepy. Zdá se mi, že místo \bar{S}_{d-1} má být spíš \bar{S}_{d-1} , v bodě 7 tohoto tvrzení je zvláštní zápis $F(x) \geq (1+u) = 2$, asi má být $(1+u)/2$.
- (7) Bod 3 tvrzení 2.3 se zdá nedořečený.
- (8) Na straně 25 u konstrukce S_i a C_i chybí nějaký vztah s k a d .

- (9) Na straně 36 dole se píše o p -hodnotě 0, ale v definici o něco výše je v čitateli vždy hodnota alespoň 1, jak může tedy být dosaženo nuly?
- (10) U testu popsaného v části 3.3.1 hraje roli náhoda přidělení znaménka. Mohlo by se proto uvažovat, jakou roli může tato náhoda hrát ve výsledku testu, zda třeba počet přidělení znamének vedoucí k chybnému závěru je či není významný vzhledem k celkovému počtu možností.

Uvedené otázky a poznámky se týkají spíše drobných nejasností a překlepů. Práce je psaná s porozuměním a přehledně a dobře se čte. Autorka prokázala schopnost zpracovat netriviální téma a přispět k jeho zkoumání. Doporučuji tuto práci **uznat za diplomovou práci** pro obor Pravděpodobnost, matematická statistika a ekonometrie.

Daniel Hlubinka
V Lovani 5.6.2023