

doc. RNDr. Karel Hron, Ph.D.

Katedra matematické analýzy a aplikací matematiky
Přírodovědecká fakulta UP
17. listopadu 12, 771 46 Olomouc

Posudek vedoucího diplomové práce
Bc. Jiří Válek: Alternativní přístupy ke statistické analýze
kompozičních dat

Diplomová práce studenta druhého ročníku oboru Aplikace matematiky v ekonomii má 73 stran včetně pěti příloh, uvádí 26 literárních pramenů a je rozčleněna do tří kapitol. První z nich (str. 10-29) se věnuje různým metodickým přístupům ke statistické analýze kompozičních dat, zejména srovnání v současnosti nejužívanější logpodílové metodiky a tzv. sférického přístupu. Dvě konkrétní metody využívající alternativních přístupů jsou potom rozebrány v následujících kapitolách. Druhá kapitola (str. 30-42) představuje regresní model založený na odmocninové transformaci. Kapitola třetí (str. 43-63) se věnuje robustní metodě hlavních komponent v tangentském prostoru včetně aplikace na reálná data.

Přestože v současné době představuje logpodílová metodika zdaleka nejrozšířenější přístup ke statistické analýze kompozičních dat, pozorování nesusobících relativní informací, historicky se tato rozvíjela vedle dalších metodických směrů, které se místo klíčového principu obsaženého v invarianci na změnu měřítka fixovaly na konkrétní, proporcionální reprezentaci kompozičních dat. Ačkoli jsou neoddělitelnou součástí všech přístupů k analýze kompozičních dat jejich určité transformace z původního výběrového prostoru do reálného prostoru, jejich chápání je odlišné. Na rozdíl od logpodílové metodiky, která je vnímá spíše jako vyjádření ve vhodných souřadnicích vzhledem k přirozené geometrii kompozic (tzv. Aitchisonově geometrii) s možností následné aplikace standardních statistických nástrojů, je totiž v tomto druhém případě cílem spíše využít konkrétního tvaru transformací (např. do tangentského prostoru) pro vývoj specifických statistických metod, nezřídka vyžadujících zapojení pokročilých matematicko-statistických nástrojů. To činí samozřejmě použití a interpretaci těchto metod obecně náročnější, než je tomu v případě logpodílové metodiky. Proto bylo cílem diplomové práce soustředit se pouze na vybrané, snadněji uchopitelné metody, na kterých by bylo možno některá výše zmíněná specifika dobře ilustrovat, a i numerická realizace na vlastním reálném problému byla z téhož důvodu provedena

pouze pro případ robustní metody hlavních komponent v tangentovém prostoru, kde bylo již možno alespoň částečně využít předchozí implementace metody v softwaru R.

Lze konstatovat, že uvedeného cíle bylo dosaženo. Diplomant prokázal schopnost zpracovat v ucelené formě netriviální metody, které jak v případě uvedeného regresního modelu s kompoziční vysvětlovanou proměnnou, tak i u robustní metody hlavních komponent výrazněji přesahovaly jeho dosavadní znalosti. Oceňuji též první kapitolu, věnovanou srovnání různých metodik pro kompoziční data, které vyžadovalo hlubší vhled do jemných rozdílů mezi různými způsoby chápání kompozičních dat, které ovšem nakonec v důsledku vedly ke zcela odlišnému pojetí jejich statistického zpracování.

K práci nemám žádných zásadních připomínek, přesto, že si diplomant zvláště ve druhé kapitole při zjednodušené prezentaci regresního modelu zjevně sáhl až na hranice svých schopností. Přesto, či právě proto navrhuji do obhajoby hodnocení *A*.

V Olomouci, dne 20. dubna 2017