prof. RNDr. Karel Hron, Ph.D.

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

Posudek oponenta bakalářské práce
Jáchym Jelínek: Klinické studie a statistika

Bakalářská práce studenta třetího ročníku programu Aplikovaná matematika, specializace Matematika v ekonomické praxi, má 56 stran, uvádí 11 literárních pramenů a je rozčleněna do čtyř kapitol (bez atypicky číslovaného úvodu a závěru). První z nich (str. 8-11) seznamuje s klinickými studiemi jako takovými, včetně průběhu a cíle studie, výběru testovací skupiny a uspořádání studie. Kapitola druhá (str. 12-17) přestavuje obecně cíl práce, tedy určení velikosti testovací skupiny při daném statistickém testu. Ve třetí kapitole (str. 18-37) je tato velikost odvozena pro konkrétní testy (jednovýběrový z-test, dvouvýběrový t-test, jednofaktorová ANOVA, chí-kvadrát test nezávislosti, McNemarův test efektu ošetření) a v kapitole čtvrté (str. 38-52) jsou teoretické úvahy a výpočty dle nich konfrontovány s příslušnou implementací ve statistických softwarech.

Hlavním cílem práce bylo odvodit velikost testovací skupiny subjektů v rámci klinické studie při volbě konkrétních statistických testů a porovnat dosažené výsledky s příslušnou softwarovou implementací v analytických nástrojích SAS Studio a RStudio. Určení této velikosti je totiž klíčové pro zajištění průkaznosti výsledků celé klinické studie.

Lze konstatovat, že uvedeného cíle bylo dosaženo. Diplomant podrobně odvodil velikosti testovací skupiny subjektů pro zvolené testy, práce je navíc psána velmi přehledným stylem a s příkladnou prací s citacemi. Do obhajoby mě tak napadají pouze následující otázky:

- Na koncích jednotlivých podkapitol s odvozením velikostí testovací skupiny (kapitoly 4.1-4.5) je vždy formulace „Postup a výsledný vzorec ověřen podle knihy Sample Size Calculations in Clinical Research. . . “. Mám tomu tedy rozumět tak, že byla tato odvozování pojata jako určitá (pokročilá) cvičení s, předpokládám, následnou kontrolou dle uvedené citace [1], nebo byla tato odvozování též nějakým způsobem optimalizována? Pokud ano, v jakých případech a jak?

- Na konci kapitoly 5.1 (resp. i potom na konci strany 42) je uvedeno, že kvantily t-rozdělení aproximují kvantily normálního rozdělení. Není

tomu spíše naopak, tedy že použití t-rozdělení je při neznámém rozptylu tou realističtější variantou? Mimochodem, asi by pak bylo vzhledem k předpokladu známého σ^2 přesnější i u dvouvýběrového testu v kapitole 4.2 použití označení z-test, podobně jako v kapitole 4.1.

- kapitola 5.5: Proč se výsledek funkce `sampleSizeMcNemar` liší od alternativních výpočtů? Lze to na základě kódu nějak vysvětlit?

Práci se mi celkově líbí a při předpokladu vysvětlení uvedených drobných nejasností navrhuji do obhajoby hodnocení A.

V Olomouci, dne 6. května 2024