

## Posudek

vedoucího oponenta

diplomové bakalářské práce

Autor/Autorka: Zuzana Kaderjáková

Název práce: Analýza přežití s programem STATISTICA

Jméno vedoucího/oponenta: Jan Hurt

Matematická úroveň:

vynikající velmi dobrá průměrná podprůměrná nevyhovující

Grafická, jazyková a formální úroveň:

Grafická: velmi dobrá, jazyková a formální: průměrná podprůměrná nevyhovující

Výsledky:

originální původní i převzaté vlastní výpočty citované z literatury opsané

Použité metody:

nestandardní standardní obojí

Aplikovatelnost:

přínos pro teorii přínos pro praxi přínos pro praxi i teorii bez přínosu nedovedu posoudit

Věcné chyby:

téměř žádné vzhledem k rozsahu a pojednávanému tématu přiměřený počet méně podstatné četné závažné

Tiskové chyby:

téměř žádné vzhledem k rozsahu a pojednávanému tématu přiměřený počet četné

Celková úroveň práce:

vynikající velmi dobrá průměrná podprůměrná nevyhovující

Práci

doporučuji nedoporučuji

uznat jako diplomovou/bakalářskou. Návrh klasifikace přikládám na zvláštním papíru.

### Připomínky a vyjádření vedoucího/oponenta:

Str. 4: „krátěná data“. Domnívám se, že i ve slovenštině se používá nějaký ekvivalent termínu „useknuté rozdělení“. Viz např. Janko 1958 Statistické tabulky, Hurt 1984 Teorie spolehlivosti.

Str. 6 a dále: vhodnější termín než „riziková funkce“ pro „hazard function“ by byl vžitý termín „intenzita poruch“.

Str. 10: „... parametr lambda určuje škálovanie funkcií Weibullovoho rozdelenia ...“ asi vyjadřuje jednoduchý fakt, že parametr lambda je parametr měřítka.

Str. 14: Pojednání o mediánu je zcestné. Medián rozhodně není „nejčastěji používaným ukazatelem polohy“. Je v jistém smyslu optimální pro Caychyovo rozdělení, naopak pro normální rozdělení je to střední hodnota, a pro rozdělení někde mezi dvěma uvedenými je to

useknutý průměr. Navíc je nesmysl tvrdit, že se odhaduje z odhadu funkce přežití. V té podobě, jak je popsán v práci je to buď prostřední hodnota v uspořádaném náhodném výběru nebo aritmetický průměr dvou prostředních hodnot v závislosti na tom, zda je počet pozorování lichý nebo sudý.

Str. 25 Jak je definován rozptyl odhadu rizikové funkce? (Volba vah 2.)

Str. 25 a 26: Z obrázků je patrné, že nejjednodušší forma neparametrického odhadu intenzity poruch a hustoty je v tomto případě nevhodná, jako ostatně téměř ve všech případech cenzorovaných modelů. Tvar odhadnuté intenzity poruch mi nepřipadá „složitý“, ale spíše typický. Měla se použít některá z vyhlazovacích metod, např. jádrový odhad. Také nerozumím tomu, že pro stejná data je funkce přežití odhadnuta parametricky (Gompertzovo rozdělení) a pro zbylé dvě pouze neparametricky.

Zadání znělo na použití softwaru STATISTICA. Chtěl bych však na závěr podotknout, že většina řešených úloh by se mohla vyřešit přímým použitím funkcí systému Mathematica (zabudované superfunkce SurvivalDistribution, Censoring, HazardRate, CensoredDistribution) a ty zbylé využitím těchto funkcí. Navíc má tento systém více než 400 zabudovaných pravděpodobnostních rozdělení, řadu metod odhadů parametrických i neparametrických, vše k použití i pro cenzorovaná data.

Místo, datum, podpis vedoucího/oponenta: Praha, 11. 6. 2012, Jan Hurt