

POSUDEK VEDOUCÍHO BAKALÁŘSKÉ PRÁCE

Název: Teoretické a empirické kvantily
a jejich využití pro konstrukci predikčních intervalů

Autor: Jakub Šimičák

SHRNUTÍ OBSAHU PRÁCE

Cíľom bakalárskej práce študenta Jakuba Šimičáka bolo oboznámiť sa s pojmami *teoretický kvantil* a *empirický kvantil* používanými v súvislosti s nejakým pravdepodobnostným rozdelením a následne v práci diskutovať využitie oboch typov kvantilov pri konštrukcii predikčných intervalov—t.j. konfidenčných oblastí pre novú (nezávislú) realizáciu náhodnej veličiny z toho istého rozdelenia. Zatiaľ čo *teoretický kvantil*—ako jeden zo základných distribučných charakteristík náhodnej veličiny s vhodným predpokladaným pravdepodobnostným rozdelením—je nutnou súčasťou konštrukcie predikčných intervalov v tzv. frekventistickom zmysle, *empirický kvantil*—získaný z realizovaného náhodného výberu z ľubovoľného (i neznámeho) pravdepodobnostného rozdelenia—slúži ako výberový odhad neznámeho teoretického kvantilu a do značnej miery našiel uplatnenie pri konštrukcii predikčných intervalov s použitím siete minimálnych (až v podstate žiadnych) teoretických predpokladov, ale značnej výpočetnej sily počítačových algoritmov—tzv. konformné predikčné intervaly získavané hlavne pomocou machine learning metód a rôznych neuronových sietí.

Práca je členená do štyroch samostatných kapitol. V prvej kapitole autor uvádza základné (na tretí ročník Bc. štúdia možno až triviálne) definície a vlastnosti teoretického aj empirického kvantilu. V druhej kapitole je stručne popísaný a na jednoduchom príklade ilustrovaný postup konštrukcie frekventistických predikčných intervalov. Tretia kapitola následne popisuje postup konštrukcie konformných predikčných intervalov—opäť aplikovaný na jednoduchý príklad. Záverečná kapitola zhrňuje výsledky simulácii, kde sú pomocou Monte Carlo štúdie porovnané niektoré frekventistické aj konformné predikčné intervaly.

Práca mi na jednej strane pride zaujímavá, keďže hlavne koncept konformnej predikcie je zo štatistického pohľadu pomerne nový, pomerne málo známy a preto z teoretického aj empirického chľadiska značne neprebádaný. Na druhej strane sa ale autor venuje danej problematike len veľmi povrchné a zďaleka nevyužíva ani čiastočný potenciál, ktorý zadaná téma pre účely bakalárskej práce podľa môjho názoru ponúka. Prvé tri kapitoly sú kompilačného charakteru, štvrtá kapitola predstavuje vlastný prínos autora a v určitom zmysle aj zaujímavé empirické zistenia.

Za najväčší nedostatok práce považujem jej formálnu stránku—t.j. formuláciu matematického, ale aj nematematického textu. Formulácie často pôsobia dojmom, že autor nie celkom rozumie význam niektorých základných matematických/štatistických pojmov. To je v podstate aj dôvodom určitého zmätočného, nepresného, či nekonzistentného značenia, ktoré sa v práci objavuje pomerne často. Práca ale splňuje zadanie. Autor pracoval samostatne a použité zdroje sú v práci riadne citované. Prácu doporučujem komisii uznať ako bakalársku prácu na MFF UK.

Z celkového hľadiska mi práca pride síce jednoduchá, teoreticky až triviálna, ale bez zásadných nedostatkov, či problémov, navyše s určitým vlastným prínosom autora. Prácu hodnotím ako priemernú.

Praha, 31.05.2023


RNDr. Matuš Maciak, Ph.D.
maciak@karlin.mff.cuni.cz