

Posudek na diplomovou práci

Bc. Tomáš Hanzlíček: Odhady parametrů smíšených distribučních funkcí náhodného vektoru

Student 2. ročníku oboru AME

Diplomová práce vychází z výzkumu Laboratoře experimentální medicíny při Dětské klinice LF UP a FN v Olomouci. Cílem prováděné studie je výzkum vlivu proteinů mnohočetné lékové rezistence (P-gp, MRP, LRP) u léčby astma bronchiale a nespecifických střevních zánětů. Výsledky, které byly získány v pilotních experimentech ukazují, že množství P-glykoproteinu (P-gp) se na lymfocytech periferní krve liší a stratifikuje je do několika skupin s nízkým, středním a vysokým množstvím P-gp. Přitom každý pacient je odlišný, jak množstvím P-gp, tak množstvím skupin. Dosud běžně používaný software však neumožňuje mnohorozměrnou analýzu křivek za použití matematického modelování. Proto je potřeba vytvořit matematický nástroj, který bude umět proložit těmito histogramy křivky a odlišit jednotlivé skupiny. Tato práce si klade za cíl pouze dvourozměrnou analýzu, tj. zjistit počet skupin a rozdělení P-gp a MRP do těchto skupin pro každého pacienta. Při řešení daného problému se ukazuje, že naměřené veličiny získané z krve pacientů mají strukturu tzv. smíšených distribucí a právě tato struktura pak nejvíce ovlivňuje léčbu.

Diplomová práce má rozsah 69 stran, je rozdělena do 4 kapitol a v seznamu literatury je uvedeno 8 položek. Autor zde přímo navazuje na svoji bakalářskou práci „Odhad parametrů smíšených distribucí“, kde se věnoval vysvětlení základních vztahů z oblasti smíšených distribucí a metodám odhadů jejich parametrů (metoda maximální věrohodnosti a její iterační řešení pomocí EM algoritmu). V diplomové práci prezentuje poznatky získané z jednorozměrné analýzy, které následně zobecňuje pro případ náhodných vektorů a řeší nové problémy, které vznikají u těchto vícerozměrných smíšených distribucí (identifikace složek jednotlivých distribucí, charakteristiky přesnosti odhadů). Teoretické výsledky jsou poté aplikovány na zkoumanou problematiku při hledání odpovědi na to, jak specifikovat léčbu pro každého pacienta. Praktická část práce je omezena pouze na jednorozměrnou a dvourozměrnou analýzu. Výpočty jsou realizovány s využitím programu Matlab. Všechny potřebné algoritmy pro výpočty a grafické výstupy jsou podrobně uvedeny v příloze. Obdržené výsledky jsou velmi vhodně a výstižně interpretovány.

Autor prokázal, že je schopen uceleně zpracovat teoretické poznatky a získané znalosti aplikovat při řešení praktických úloh. Je pouze škoda, že se v praktické části u vícerozměrné analýzy omezil jen na dvousložkové distribuce. Toto omezení je značně limitující a neumožnilo jinak podrobnou analýzu souboru dat precizně dořešit.

Práce je psána jasně a srozumitelně, matematicky korektně, typograficky velmi pěkně upravena (až na občas chybějící interpunkci). K předložené práci mám několik připomínek, které jsou uvedeny níže.

Podrobné připomínky:

- str. 18, Definice 2.2: co znamená Wishartova matice s $n-1$ stupni volnosti
- str. 18, Definice 2.3: co znamená hustota obecného rozdělení
- str. 22₁: parciální derivace věrohodnostní funkce $\ln L(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ podle $\boldsymbol{\Sigma}$ je uvedena špatně. Jelikož se nejedná o standardně známé postupy, je třeba tuto derivaci podrobně spočítat a uvést vztahy, které se při výpočtu užívají.
- str. 43-46, výpočet empirické Fisherovy informační matice: tato část by si zasluhovala důslednější zpracování; proč je zde počítána parciální derivace $f(\mathbf{y})$ podle $\boldsymbol{\Sigma}^{-1}$, potřebný výpočet opět podrobně rozepsat, str. 46²⁻³ vysvětlit, proč „Každý řádek v tomto vektoru vydělíme hodnotou $f(\mathbf{y})$ “, zavedené označení $\mathbf{Y}$ je trochu matoucí, výsledný vzorec (53) je zapsán nepřesně a s použitím nedefinovaného symbolu.
- V praktické části se objevuje zásadní vliv identifikace jednotlivých složek smíšené distribuce na výsledné odhady. Byly zkoušeny i jiné postupy než uvedená metoda dělicí křivky, která se na prezentovaných příkladech nejeví příliš dobře? Výsledky však samozřejmě mohly být dosti ovlivněny omezením na dvousložkové distribuce.

Vzhledem k rozsahu a kvalitě diplomovou práci **doporučuji** k obhajobě. Jestliže diplomant u obhajoby uspokojivě zodpoví jednotlivé připomínky, navrhuji hodnocení **v ý b o r n ě**.

V Olomouci, 12. května 2010

Doc. RNDr. Eva Fišerová, Ph.D.
oponent práce