

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Název: Joint Models for Longitudinal and Time-to-Event Data
(Sdružené modely pro longitudinální a cenzorovaná data)
Autor: Bc. Jana Vorlíčková

SHRNUTÍ OBSAHU PRÁCE

Předložená práce se věnuje modelům, které umožňují v longitudinálních studiích vyhodnocovat vliv proměnných (např. druh léčby) na stav pacienta (např. jeho mentální schopnosti) a také rizika sledované události (např. propuknutí určité duševní nemoci), která během doby studie může či nemusí nastat (cenzorování). Autorka se zaměřuje zejména na model s latentními třídami a jeho odhad v přístupu Bayesovské statistiky.

Téma práce. Téma odpovídá znalostem a možnostem příslušného oboru magisterského studia a autorka jej nepochybně naplnila.

Vlastní příspěvek. Vlastní příspěvek autorky vidím (kromě přehledného představení modelů) zejména v podrobném popisu modelu latentních tříd a v odvození plně podmíněných rozdělení (full conditional distributions) pro všechny parametry modelu. Dále pak v provedení výpočetně velmi náročné simulační studie a v jejím kritickém zhodnocení.

Matematická úroveň. Z matematického hlediska je práce přehledového charakteru doplněná výše zmíněnými odvozeními plně podmíněných rozdělení. Matematický text je korektně zformulován.

Práce se zdroji. Práce obsahuje vlastní text autorky. Zdroje jsou řádně citovány. Jen občas v referencích chybí nějaký důležitý údaj, viz např. u článku Hastings (1970).

Formální úprava. Formální úroveň práce je velmi dobrá s malým množstvím překlepů. Na kolik dokážu posoudit, tak práce je z hlediska nerodilého mluvčího psána slušnou angličtinou.

CELKOVÉ HODNOCENÍ PRÁCE

Dle mého názoru se jedná o pěknou práci. K dokonalosti jí chybí o trošku větší pečlivost a možná také více času na lepší nastavení simulační studie (viz poslední připomínka níže).

PŘIPOMÍNEK A DOTAZY

- str. 8 the second **Definition** on this page: at t_2, t_2, \dots
- str. 8 the third **Definition** on this page: $\Lambda(t) = \sum_{i:t_i \leq t} \lambda(t_i) ds$
- str. 9 *Remarks*: V textu je odkaz na vzorec (1.5), což asi nebude to, co autorka zamýšlela.
- str. 12: Domnívám, že řada čtenářů by ocenila vysvětlení, proč platí uvedený vzorec pro $p(\theta_j | \mathbf{y}, \theta_{-j})$.
- str. 14: Věta: “*An example of such density is a symmetric random walk*” zní poněkud zvláště.
- str. 29⁵: Ve vzorci hustoty by asi mělo být $\Sigma_{\beta_g}^{-1}$. Dále se na dalším řádku objevuje matice Σ , která je vysvětlena až o dost později (tedy pokud jsem něco nepřehlédl).
- str. 30⁸: Domnívám se, že $\tau^{a\tau}$ by mělo být před $\prod_{g=1}^G$. Dále je drobet matoucí, že autorka již používá N_g , ale o několik řádků níže se vrátí zpět k $\sum_{i:v_i=g} \frac{n_i}{2}$.

8. str. **34**: Poněkud nepřehledný zápis rovností ve (4.5).
9. str. **43**: Ocenil bych nějaký komentář ohledně zvolených modelů pro jednotlivé třídy.
10. str. **45 – 46**: Oceňuji, že se autorka nesnažila zamlčet problematické výsledky simulační části. Na druhou stranu je však poněkud hloupé, že simulace vlastně zejména ukazují to, že použití daného modelu je pro daný rozsah výběru a daná data poněkud problematické. A zatímco některé podivné výsledky s rostoucím rozsahem výběru mizí, jiné naopak přetrvávají. A není vůbec jasné, zda tyto přetrvávající podivnosti zmizí s ještě většími rozsahy výběrů. Měla by autorka nějaký nápad, jak by postupovala, kdyby měla udělat novou simulační studii?

OTÁZKY K OBHAJOBĚ

1. Odpovězte na připomínku č. 10.

ZÁVĚR

I přes výše uvedené výtky pokládám předloženou práci za velmi dobrou. Na kolik jsem schopen posoudit, tak práce splňuje všechny požadavky kladené na diplomovou práci na oboru Pravděpodobnost, matematická statistika a ekonometrie.

doc. Ing. Marek Omelka, Ph.D.
KPMS MFF UK
29. června 2020