

## Posudek oponenta bakalářské práce

Název: Skórové testy v kontingenčních tabulkách

Autor: Martin Jex

### Shrnutí obsahu práce

Bakalářská práce Martina Jexe se zabývá testováním hypotéz v multinomickém rozdělení. Odvozují se zde testové statistiky založené na metodě maximální věrohodnosti, tj. test poměrem věrohodností, Raův skórový test a Waldův test, a to jak pro jednoduchou hypotézu tak pro složenou hypotézu za přítomnosti rušivých parametrů.

### Celkové hodnocení práce

**Téma práce.** Téma práce je pro bakalářskou práci naprosto přiměřené. Téma bylo původně zadáno trochu jinak, než jak je práce postavena, na kontingenční tabulky nedošlo, ale testy pro multinomické rozdělení nepochybně s analýzou kontingenčních tabulek úzce souvisejí. Únik z původního tématu lze podle mého názoru tolerovat, i když předložená práce je poněkud jednoduššího charakteru, než kdyby sledovala původně zadané téma.

**Vlastní příspěvek.** Martin Jex sestavil přehledný soupis výsledků z teorie maximální věrohodnosti pro multinomické rozdělení. Z obecných tvarů testových statistik pro jednoduché hypotézy podrobně odvodil konkrétní vzorce pro multinomické rozdělení. Pro rušivé parametry se pouze odkazuje na monografii J. Shaa. V práci je dále ilustrační příklad s Galtonovou deskou.

**Matematická úroveň.** Matematická úroveň práce je dobrá, je tu snaha o preciznost, některé výsledky jsou podrobně odvozovány a dokazovány. Značení je řádně zavedeno. Odvození hlavních výsledků (Lemma 2 a Lemma 3) je uděláno poněkud otrocky, lze vymyslet daleko jednodušší a elegantnější postup.

**Práce se zdroji.** Zdroje jsou řádně citovány, z citací je zřejmé, které výsledky autor převzal a odkud.

**Formální úprava.** Formální úprava je typograficky slušná, bez zásadních prohřešků. Lze vytknout pár překlepů v textu a neobratnost některých formulací.

### Připomínky

1. Str. 3, komentář k Def. 1: Co si má čtenář představit pod formulací „*n* nezávislých *stejných* experimentů“?
2. Str. 6 uprostřed: ve výrazu  $(X - p_{-K})$  je třeba z  $X$  vyloučit poslední složku. Dále není jasné, odkud se vzal rozptyl limitního normálního rozdělení.
3. Str. 7, řádek 3: Tento vzorec udává  $a_{lk}$ , nikoli  $a_{kl}$ .
4. Str. 8 – 10: Důkaz Lemmatu 2 spočívá v tupém dosazování. Ve skutečnosti lze podat celý důkaz na třech řádcích, kdyby se chytre využilo původního tvaru Waldovy statistiky.
5. Str. 10 – 12: I důkaz Lemmatu 3 by šel zestručnit, ale bylo by asi potřeba o něco více řádků než u Lemmatu 2.
6. Str. 17, řádek 7 odspodu: Skóre  $S_n$  tady asi nemá být derivací celé logaritmické věrohodnosti.

7. Str. 18: Překlepy „sitaci“, „paramaterický“.
8. Str. 19: Příklad nešikovné formulace – „Nyní uvažujme poslední statistiku z maximálně věrohodného přístupu test založený na podílu věrohodností.“
9. Str. 22: Chybějící čárky ve větě „Z uvedeného lze vyvodit, že deska nebyla korektně zkonstruována a naší domněnku, že model neodpovídá realitě jsme si potvrdili.“
10. Str. 24: Nesrozumitelná věta „Jinak řečeno všechny kolíky kuličky neodráží s neznámou stejnou pravděpodobností doleva.“
11. Str. 25: Chybí podmět ve větě „Znovu vyneseme do tabulky 3.5 na konci této kapitoly.“

### Otázky a úkoly k obhajobě

- (A) Předvedte důkaz Lemmatu 2, který se vejde na jeden slide (a přitom jsou v něm všechny rovnosti zdůvodnitelné).
- (B) Je tvar Waldovy statistiky s rušivými parametry uvedený dole na str. 18 v citované literatuře dokázán nebo je uveden bez důkazu? Myslíte si, že Waldova statistika opravdu vypadá takto?

### Závěr

Práce Martina Jexe podle mne stačí na naplnění zadání a svým provedením splňuje stanovené požadavky, doporučuji ji tedy uznat jako bakalářskou práci.

doc. Mgr. Michal Kulich, PhD.  
KPMS MFF UK  
18. června 2019