

## POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

**Název:** Geometrické rozdělení a jeho mnohorozměrné rozšíření

**Autor:** Diana Pavlovičová

### SHRNUTÍ OBSAHU PRÁCE

Tématem práce je zobecnění geometrického a negativně binomického rozdělení pro situace, kdy uvažujeme více druhů neúspěchů. V práci je nejdříve předložena formální definice dvojrozměrného geometrického rozdělení, pro které jsou následně podrobně odvozeny vektor středních hodnot, varianční matice a pravděpodobnostní vytvořující funkce. Dále jsou odvozena marginální rozdělení, rozdělení celkového počtu neúspěchů před prvním úspěchem a podmíněné rozdělení počtu neúspěchů konkrétního typu před prvním úspěchem při znalosti celkového počtu neúspěchů před prvním úspěchem. Situace je poté rozšířena na tři různé typy neúspěchů a studium sdruženého i marginálních rozdělení kombinovaných neúspěchů, tj. situací, kdy se ptáme na celkový počet neúspěchů  $i$ -tého nebo  $j$ -tého typu před prvním úspěchem. Práce dále zmiňuje možnost rozšíření na vícerozměrné negativně binomické rozdělení. Na závěr práce diskutuje odhad parametrů v dvojrozměrném geometrickém rozdělení metodou maximální věrohodnosti, včetně vyšetření konzistence a konstrukce asymptotických intervalových odhadů. Analytická odvození jsou ilustrována krátkou simulační studií.

### CELKOVÉ HODNOCENÍ PRÁCE

**Téma práce.** Téma práce se věnuje srozumitelnému problému sledování počtu neúspěchů různého typu před prvním úspěchem. I když je práce zaměřena zejména na provedení přímočarých výpočtů, nabízí dostatek prostoru, aby uchazeč ukázal, že rozumí základním konceptům z teorie pravděpodobnosti, je schopen korektně formulovat matematický text a interpretovat dosažené výsledky. Proto ho považuji za vhodné pro bakalářskou práci. Zadání práce bylo uchazečkou splněno, některé výtky ke zpracování jsou uvedeny níže.

**Vlastní příspěvek.** Vlastní příspěvek autorky spočívá v pečlivém odvození základních charakteristik a výpočtu marginálních či podmíněných rozdělení. Detailní odvození tedy doplňuje existující literaturu, kde jsou obdobné výsledky zmíněny bez velkých podrobností.

**Matematická úroveň.** Matematická úroveň práce je dobrá, text formulován srozumitelně. Jsou zde ovšem některé nedostatky, např. při výpočtu pravděpodobnostní vytvořující funkce pro konkrétní rozdělení není uvedeno, pro jaká  $s$  je tato funkce definována. V práci je dále používáno, že pravděpodobnostní vytvořující funkce jednoznačně charakterizuje pravděpodobnostní rozdělení. Tento výsledek není explicitně formulován ani formou odkazu na literaturu. Obdobně je odvozen vzorec pro pravděpodobnostní vytvořující funkci náhodného součtu nezávislých náhodných veličin (toto odvození je v pořádku), který je ale poté bez dalších komentářů používán pro práci s náhodnými vektory. Absence vysvětlujících komentářů, které by propojily jednotlivé pasáže textu (proč jsme volili zrovna tento postup výpočtu, proč pracujeme s pravděpodobnostní vytvořující funkcí), mírně kazí dojem z textu práce. Pozor také na vágní formulace „snadno nahlédneme“, „snadno se přesvědčíme“, atp. Jelikož hlavní přínos práce spočívá v podrobném odvození výsledků dostupných v literatuře, měla by autorka komentovat, jaká pomocná tvrzení k dosažení výsledku použila.

**Práce se zdroji.** V úvodu práce jsou citovány tři zdroje zabývající se danou problematikou. Tyto citace jsou provedeny korektně. V hlavním textu práce autorka provádí odvození nezávisle

na citovaných zdrojích, chybí ovšem podrobnější srovnání dosažených výsledků s uvedenou literaturou.

**Formální úprava.** Formální úprava textu je na vysoké úrovni, která zcela odpovídá úrovni bakalářské práce. Text obsahuje naprosté minimum překlepů, jazykových, stylistických či typografických nedostatků.

#### PŘIPOMÍNKY A OTÁZKY

1. V abstraktu se objevuje formulace „odhady parametrů jednoduchého dvojrozměrného geometrického rozdělení“. Na co zde odkazuje slovo jednoduché?
2. Do budoucna bych zvažila, zda některé formulace nejsou zbytečně zavádějící. Např. v první kapitole připomínáte, jak je zavedeno geometrické rozdělení, jaký je jeho vztah k jiným pravděpodobnostním rozdělením a nabízíte motivaci k zavedení jeho vícerozměrné varianty, včetně příkladů na hody kostkou. Z textu úvodu by ale čtenář mohl očekávat hlubokou diskusi o vícerozměrných rozděleních (fráze „jejich odvození z jiných rozdělení“ a „jejich dalších zobecnění“ jsou matoucí).
3. V první kapitole zmiňujete, že přechod od geometrického rozdělení k vícerozměrnému geometrickému rozdělení se dá srovnat s přechodem od binomického rozdělení k multinomickému. Uměla byste tyto dvě situace srovnat? Co je pro ně podobné? Čím se liší?
4. Zmiňujete, že geometrické rozdělení je rozdělení bez paměti. Uměla byste toto odvodit? Proč je tato vlastnost pro vaši práci důležitá?
5. Ve znění definice 1 by se mělo připomenout, že  $p_1 + p_2 < 1$ .
6. Uměla byste odvodit, jak z pravděpodobnostní vytvořující funkce obdržet rozdělení diskrétní náhodné veličiny?
7. Na straně 8 je proveden výpočet pravděpodobnostní vytvořující funkce pro binomické rozdělení. Pro jaká  $s$  je tato funkce definovaná? Tento tvar ale není dále použit, proč jsme zde nepočítali tvar pro geometrické rozdělení, které se v textu dále používá? Pro jaká  $(s_1, s_2)$  je definovaná funkce  $P_{X_1, N}$ ?
8. Značení  $X_i, i \in \{1, \dots, n\}$  na straně 10 je matoucí. Mluvíte teď o vztahu jednorozměrného geometrického a negativně binomického rozdělení? Vzorec pro pravděpodobnostní vytvořující funkci součtu nezávislých náhodných veličin je v pořádku, jak konkrétně se ale použije na výpočet vytvořující funkce náhodného vektoru  $(X_1, N)^T$  (obdobný dotaz je i k  $P_{X_1, 2, X_2, 3}$  z části 3.2)?
9. Umíte ukázat, že se pro  $P(X_{1,2} = x, X_{2,3} = y)$  opravdu jedná o pravděpodobnostní rozdělení? Jaké jsou zde podmínky na  $p_0, p_1, p_2$  a  $p_3$ ?
10. Jak přesně v důkazu tvrzení 1 použijete větu o úplné pravděpodobnosti?
11. Pro jaká  $\mathbf{x}$  je definovaná  $f_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\theta})$ ?
12. V části 4.2 je formulace „o odhadech získaných metodou maximální věrohodnosti je obecně známo, že bývají (slabě) konzistentní a zpravidla nejsou nestranné“. Umíte formulovat větu, která dává konkrétní podmínky, kdy toto platí? Jsou ve vašem případě tyto podmínky splněny?
13. Na jaké množině je funkce  $k$  ze sekce 4.2 konvexní?

14. Jak se jmenuje věta, kterou používáte pro získání vztahu (4.2)? Jaké jsou její předpoklady? Jakou jinou známou větu používáte ke konstrukci asymptotického odhadu?
15. Napadá vás příklad z praxe (házení kostkou můžeme považovat spíše za školní, ilustrativní příklad), kdy by se dalo vícerozměrné geometrické rozdělení využít? Napadá vás, jakou hypotézu by v daném příkladu mělo smysl testovat?
16. V abstraktu se objevuje formulace, že se na simulovaných datech pokusíte předvést „spolehlivost“ odhadů parametrů ve vícerozměrném geometrickém rozdělení. Je spolehlivost rigorózní pojem? Jak je definovaná? V textu sekce 4.4 se objevuje formulace „odhady jsou velmi přesné“. Jak spolu souvisí „přesné“ a „spolehlivé“? A jak jsou tyto dva pojmy navázány na rozsah vašeho náhodného výběru?
17. Z překlepů pozor na jiné značení pro střední hodnotu na straně 5 a sjednocení značení pro rozptyl (zda  $var(X)$  či  $varX$ ). Pozor také na vzorec pro pravděpodobnosti v binomickém rozdělení na straně 8. Jak se přesně přejde mezi prvním a druhým řádkem u výrazu (3.2)?
18. Z typografického hlediska pozor na číslování u soustavy několika rovnic. Např. u (2.2) není zcela zřejmé, na který řádek se odkazujete. Pro odkaz na rovnici v textu se používá příkaz

`\eqref`

Na rovnici očíslovanou jako (2.3) není dále v textu odkazováno. Také pozor na závorky u intervalového odhadu ze strany 17, zde použít

`\left(`

a

`\right)`

## ZÁVĚR

I přes zmíněné výtky je práce vypracována pečlivě a splňuje všechny požadavky. Doporučuji ji tedy uznat jako bakalářskou práci.

Mgr. Kateřina Pawlasová

Katedra pravděpodobnosti a matematické statistiky MFF UK

13. 6. 2022