

## Nezávislostní modely

*Mgr. Petra Šimečka*

Předložená disertace je věnována studiu podmíněné nezávislosti mezi složkami náhodného vektoru a statistickým modelům, které jsou zadané takovými nezávislostmi. Tato tematika je předmětem výzkumu zhruba už dvacet let na nejméně dvou desítkách pracovišť po celém světě. Hlavní pozornost v práci P. Šimečka je soustředěna na problém, které nezávislosti mohou nastat simultánně. Autor uvažuje třídy náhodných vektorů, jejichž distribuce je gausovská nebo nesena konečnou množinou, nebo jejichž složky jsou binární. Tento typ problémů se ukázal být velice složitý. Známé výsledky jsou jenom dílčí; nedávno byly shrnuty knižně (M. Studený: Probabilistic conditional independence structures, Springer 2005). Inspirací P. Šimečka byly také některé moje publikace k této tématice.

Členění disertační práce na čtyři kapitoly je funkční. První z nich uvádí základné pojmy, terminologii a značení. Parametrizace binárních rozdělání momenty odpovídá Fourier-Stieltjesově transformaci na grupě  $\mathbb{Z}_2^n$ , což není zmíněno. Překvapením je Definice 10, která se liší od terminologie v mém článku [Mat97, str. 104], a tudíž od běžné terminologie v teorii grafů.

Kapitola 2 je věnována neregulárním gausovským modelům. Prezentuje výsledky počítačových experimentů s maticemi  $4 \times 4$ , které potvrzují Větu 1 R. Lněničky [Lně05], a rozšiřují ji na neregulární případ, viz Věta 2. Originální pozorování jsou dokázána v Lemmatu 14 (vi)–(vii) a 15 (xi). Hlavním výsledkem je netriviální Věta 3 o neexistenci konečné množiny zakázaných (minimálních) minorů. Stojí za zmínku, že podobný výsledek o neexistenci konečné axiomatizace byl dokázán letos odlišnými algebraickými metodami Setem Sullivantem ve Stanfordu.

Na úvod Kapitoly 3 byly zařazeny výsledky počítačových experimentů s rozděleními čtyř binárních veličin. Lemma 20 formuluje nové, dosud neznámé vlastnosti pozitivně binárních modelů. Cenné jsou opravy chybných tvrzení v literatuře a hypotéza o neexistenci dalších vlastností podmíněné nezávislosti v pozitivních distribucích. Věta 4 dává nové odhady na počet  $d^+$ -reprezentovatelných modelů s čtyřmi veličinami. Netriviální je první část Lemmatu 26, které se použilo k důkazu Věty 5 pro případ  $d^+$ -reprezentovatelných modelů. Tvrzení Věty 5 pro  $b^-$  a  $b^+$ -reprezentovatelné modely je však triviální, jelikož žádná třída neuzavřená na minory není konečně charakterizovatelná dle Definice 10.

V Kapitole 4 se studují statistické modely zadané podmíněnými nezávislostmi. Poté, co jsou připomenuty standartní grafické modely nad neorientovanými a orientovanými grafy a informační kritéria pro výběr modelů (AIC, BIC), diskuse se zužuje na modely z Věty 1. Jsou vypracovány parametrizace těchto modelů. Model se pak nalezne maximalizací informačních kritérií. Tato strategie je použita na reálná data. Z výsledků simulace vyplývá, že v některých případech jsou obecné modely zadané podmíněnými nezávislostmi v lepším souladu z daty, než jakékoliv známé grafické modely. Nezbývá než

litovat, že práce nebyla napsána anglicky, neboť tato originální statistická aplikace může být zajímavá pro širší okruh zahraničních specialistů.

K práci je přiložen také CD s její pdf-verzí, seznamy modelů, dalšími pracemi autora a datovým souborem.

Dále bych zmínil některé své drobné postřehy a dotazy:

- (1) Úvod je přehnaně beletristický.
- (2) Poslední věta před Definicí 7 platí jenom pro nezávislosti s disjunktními množinami.
- (3) Na str. 27 se pracuje s maticemi s celočíselnými prvky, ačkoli to není zdůrazněno.
- (4) Na str. 29 není definováno  $\xi_b \simeq \xi_c$ .
- (5) Důkaz Lemmatu 18 je možno ještě zjednodušit.
- (6) Jelikož binární modely nejsou uzavřené na minory, první odstavec na str. 43 nedává smysl, ledaže by byly uvažovány jenom ty 3-minory, které jsou restrikcemi.

S potěšením mohu konstatovat, že autor pronikl do zkoumané problematiky, zorientoval se v literatuře i jemných souvislostech a prokázal schopnost samostatné tvořivé vědecké práce. Přes uvedené výhrady, které jsou ostatně jen dílčí, tato disertace přináší nové výsledky, a to jak teoretické, tak praktické. Proto ji považuji za dostačující k udělení hodnosti PhD.

V Praze, 19.7.2007.



Ing. František Matůš, CSc.