Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DISERTAČNÍ PRÁCE

Petr Šimeček

Nezávislostní modely

Katedra pravděpodobnosti a matematické statistiky

Školitel: *RNDr. Milan Studený, DrSc., ÚTIA AV ČR*
Studijní program: *matematika*
Studijní obor: *pravděpodobnost a matematická statistika*

2007

Děkuji svému školiteli, RNDr. Milanu Studenému, DrSc. za starostlivé vedení práce, konzultace a nezměrné množství času a energie, Ing. Františku Matúšovi, CSc. a Mgr. Radimu Lněničkovi za podněty a připomínky, bez kterých by tato práce jen těžko vznikla, a kolegům z ÚTIA AV ČR a KPMS MFF UK za podporu a spolupráci.

Prohlašuji, že jsem svou disertační práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze dne 13. června 2007

Petr Šimeček

Obsah

Název práce: Nezávislostní modely
Autor: Petr Šimeček
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedoucí práce: RNDr. Milan Studený, DrSc.
Email vedoucího práce: studeny@utia.cas.cz

Abstrakt: Důležitou součástí teorie rozhodování za neurčitosti je oblast grafických modelů, kde mnohorozměrné pravděpodobnostní rozdělení je přiřazeno k nezávislostní struktuře. Zajímavou otázkou pocházející od J. Pearla je otázka reprezentovatelnosti takových struktur neboli pro jaké seznamy podmíněných nezávislostí existuje rozdělení splňující tyto a pouze tyto nezávislosti. Tato práce shrnuje známé výsledky a dále je rozšiřuje v několika distribučních rámcích. Zaměříme se především na případ čtyř náhodných veličin, kdy zodpovíme otázku reprezentovatelnosti pro obecné Gaussovo rozdělení (zobecnění výsledku R. Lněničky) a nalezneme částečné řešení v případě diskrétního a binárního rozdělení. Na závěr využijeme dosažené výsledky pro odvození odhadu varianční matice v případě regulárního Gaussova rozdělení.
Klíčová slova: podmíněná nezávislost, grafické a nezávislostní modely

Title: Independence models
Author: Petr Šimeček
Department: Department of probability and mathematical statistics
Supervisor: RNDr. Milan Studený, DrSc.
Supervisor's email: studeny@utia.cas.cz

Abstract: An essential part of probabilistic reasoning is the theory of graphical models endowing a multivariate probabilistic distribution with the independence structure. An interesting question originally coming from J. Pearl is the problem of probabilistic representability, i.e. for which lists of conditional independence constraints there exists a random vector satisfying these and only these independences. In this work the known results are collected and the problem is further studied in several distributional frameworks. The focus is particularly on a special case of four random variables where the solution is found for a general Gaussian distribution (extending the result of R. Lněnička) and partially for discrete and binary distributions. At the end the achieved results are used to derive the variance matrix estimator in a case of regular Gaussian distribution.
Key words: conditional independence, graphical and independence models

Úvod

„A k čemu je to vlastně dobré, to, co děláš?“ byl dotaz, který jsem opakovaně slýchával od svých kolegů. Stručnou odpovědí na tuto otázku zároveň nastíníme, co bude tématem této práce.

V aplikované statistice se často setkáváme s přístupem, že její „uživatelé“ si vyberou svoji „oblíbenou“ míru závislosti mezi dvěmi veličinami a k ní příslušný test (t–test, χ^2–test v kontingenčních tabulkách $m \times n, \ldots$), který následně bez rozmyslu používají. Taktéž často dochází k podcenění analýzy, zda by nebylo vhodné při zkoumání uvedené závislosti zahrnout do studie další důležité faktory. Tento přístup „vezmi-dvě-veličiny-a-změř-závislost“ však vede k absurdním výsledkům jako že nepřítomnost perských koberců v bytě napomáhá častějším astmatickým záchvatům jeho obyvatel[1], že nadměrná konzumace kávy snižuje u mužů středního věku krevní tlak[2] či dokonce, že čápi skutečně nosí děti[3].

Pomoci může kvalitní literatura zaměřená na metodiku výzkumu, viz např. [Dis00], ale také popularizace a lepší pochopení konceptu tzv. „podmíněné nezávislosti náhodných veličin“ jakožto rozšíření „klasické“ nezávislosti známé ze základních kurzů pravděpodobnosti o myšlenku, že dvě náhodné veličiny se stanou nezávislými teprve, když dostaneme informaci o veličině třetí. Na první pohled jednoduchá a intuitivně jasná idea má však řadu úskalí.

Platí, že pokud jsou dvě náhodné veličiny ξ_a a ξ_b podmíněně nezávislé dáno $(\boldsymbol{\xi}_c)_{c\in C}$ (tzn. stanou se nezávislými, pokud již známe hodnotu souboru veličin $(\boldsymbol{\xi}_c)_{c\in C}$), potom jsou nutně ξ_a a ξ_b podmíněně nezávislé, i když podmínku změníme na $(\boldsymbol{\xi}_d)_{d\in D}$, kde $D \supseteq C$ (ev. $D \subseteq C$)? Kdy nezávislostní

[1] Pochopitelně, neb kdo tento koberec v bytě má, těžko bude astmatik.

[2] Bylo prezentováno jako součást [ŠZTJ04]. Kdo má vysoký krevní tlak, zpravidla nepije kávu přesmíru nebo to alespoň nepřizná svému lékaři.

[3] Čápi se vyskytují tam, kde je dobré životní prostředí, a tento faktor je taktéž korelován s vysokou porodností.

tvrzení vyplývá z platnosti jiných nezávislostních tvrzení? Lze tato pravidla nějak jednoduše charakterizovat či axiomatizovat alespoň při omezeném počtu veličin? To bude typ otázek, na něž se budeme snažit v této práci nalézti odpověď.

Často pro jednoduchost předpokládáme, že se dá nezávislostní struktura mezi náhodnými veličinami zobrazit diagramem (grafem), v němž body (vrcholy grafu) odpovídají jednotlivým veličinám a čáry či šipky (orientované či neorientované hrany) přímým závislostem. Takovéto „diagramy" mají neoddiskutovatelnou výhodu a to, že jsou snadno pochopitelné i pro laika. Pro danou třídu modelů se vžil název „grafické modely".

Grafické modely prošly od 70. let minulého století bouřlivým vývojem a dnes mají uplatnění v řadě odlišných oblastí, jako jsou na jedné straně expertní systémy (viz [CLDS99], program Hugin aj.) a systémy pro podporu rozhodování (seznam aplikací na konci [CGH96]) a na straně druhé výpočetní Bayesovská statistika (modul Doodle programu WinBUGS, balíček MCMCpack pro R aj.) a v návaznosti na ni rozpoznávání řeči, obrazu, analýza textu, aplikace v bioinformatice (pro přehled viz [Jor04]) a dobývání znalostí z dat (viz [Cas02]).

Návod k použití

Po tomto spíše neformálním úvodu následují čtyři kapitoly. V první z nich zavedeme základní pojmy jako jsou podmíněná nezávislost, její kritéria v závislosti na uvažovaném distribučním rámci, nezávislostní model (volně přeloženo seznam nezávislostních vztahů), reprezentovatelnost nezávislostních modelů a (ne)existence její konečné charakterizace. Vyjma některých technických záležitostí jako je zobecnění kritéria podmíněné nezávislosti na obecné, ne nutně regulární Gaussovo rozdělení (Lemma 5) a netradiční parametrizace binárního rozdělení a z ní odvozená kritéria pro podmíněnou nezávislost tato část neobsahuje originální výsledky.

V druhé kapitole se soustředíme na Gaussovskou reprezentovatelnost nezávislostních modelů nad čtyřmi veličinami neboli rozšíříme výsledek R. Lněničky pro regulární Gaussovskou reprezentovatelnost [Lně05] na obecný případ (publikováno v [Šim06b]). Na závěr za pomoci analogické myšlenky jako v [Stu92] dokážeme, že ani Gaussovskou reprezentovatelnost nelze konečně charakterizovat ve smyslu seznamu zakázaných minorů (publikováno v [Šim06a]).

Třetí kapitola je zaměřena na diskrétní reprezentovatelnost nezávislostních modelů nad čtyřmi veličinami. Jsou zde shrnuty známé výsledky F. Matúše a M. Studeného (viz [Mat97] a [SB94]) a opraven rozšířený omyl o počtu takto reprezentovatelných modelů (publikováno v [Šim06d]). Dále je diskutována otázka, které z výše zmíněných diskrétně reprezentovatelných modelů jsou i pozitivně reprezentovatelné. Pro čtyři veličiny je určena horní a dolní mez, vyslovena hypotéza, že horní mez je skutečným počtem těchto modelů, a určeno 12 modelů, o jejichž reprezentovatelnosti je třeba rozhodnout, aby hypotéza byla dokázána či vyvrácena.

Ve čtvrté kapitole se snažíme propojit tyto výsledky s teorií grafických modelů. Jsou připomenuty základní pojmy z teorie grafů, odpovídajících nezávislostních modelů a pravděpodobnostních rozdělení. Pro nově příchozího do oboru může být užitečné zkusit si jako cvičení dokončit části důkazů, jež jsou označeny jako snadné. Následně testujeme schopnost určit na základě dat model a identifikovat jeho parametry (část publikována v [Šim06c]). Práci zakončuje simulační studie na případu Gaussova rozdělení nad čtyřmi veličinami: Je ukázáno, že při malém počtu pozorování je identifikace skutečného modelu nepravděpodobná. Studován je též odhad varianční matice, kdy chování odhadů založených pouze na grafických modelech je shledáno neuspokojivým, pokud skutečné rozdělení grafickému modelu neodpovídá. Odhady založené na nezávislostních modelech se ve všech případech pochopitelně chovají lépe.

Veškeré seznamy nezávislostních modelů, jakožto i procedury pro manipulaci s nimi a R–balíček pro výše zmíněné odhady lze nalézt na přiloženém CD.

Protože obtížnost jednotlivých částí textu je poměrně různorodá a lze předpokládat, že vhled čtenářů do dané problematiky se bude dramaticky lišit a jen málokterý z nich bude mít zájem postupovat od začátku do konce, sepsal jsem (zcela nezávazný) seznam doporučení, jak tuto práci číst. Schematicky je tento „návod k použití“ znázorněn na obr. 1.

Začátečníkem je myšlen nově příchozí do této oblasti, který se snaží v rychlosti posbírat potřebné kvantum informací o podmíněné nezávislosti a grafických modelech.

Pokročilý uživatel tohoto textu je již s teorií grafických modelů dobře obeznámen a zajímají jej toliko podrobnosti okolo nezávislostních modelů a jejich reprezentovatelnosti. A pro několik „expertů“, jež hledají pouze nové a originální výsledky, je určena poslední kategorie.

	Začátečník	Pokročilý	Expert
1. kapitola	oddíly 1.1-1.3	oddíly 1.2-1.4	lemma 5 binární par. v odd. 1.2.3 letmo oddíl 1.4
2.kapitola	jen úvod	celá kapitola	lemma 14 + věta 2 lemma 18 + věta 3
3.kapitola	jen úvod	celá kapitola	lemma 20 věta 4 + hypotéza věta 5
4.kapitola	oddíly 4.1-4.2	oddíl 4.3	oddíl 4.3

Obrázek 1: Jak číst tuto práci.

Kapitola 1

Základní pojmy

V této kapitole uvedeme základní definice, tvrzení a pojmy, s kterými se budeme ve zbytku práce setkávat a na které budeme odkazovat. Čtenář mající základní znalosti o pravděpodobnosti a podmíněné nezávislosti ji může přeskočit, případně se omezit na oddíly 1.3 a 1.4.

Předmětem našeho studia bude **náhodný vektor** neboli soubor náhodných veličin $\boldsymbol{\xi} = (\xi_a)_{a \in N}$ indexovaný (konečnou) množinou $N = \{1,2,\ldots,n\}$ a nabývající hodnot z $\mathbb{R}^n$. Sdružené rozdělení tohoto vektoru budeme značit písmenem P. Pokud má toto rozdělení *hustotu* vůči pevně zvolené součinové $\sigma-$konečné míře μ, budeme ji značit $f(\boldsymbol{x})$. Pro podmnožinu indexové množiny $A \subseteq N$ definujeme podvektor $\boldsymbol{\xi}_A$ jako $\boldsymbol{\xi}_A = (\xi_a)_{a \in A}$ za konvence $\boldsymbol{\xi}_\emptyset = 0$ (konstanta). Analogicky budeme chápat marginální rozdělení P_A, jeho hustotu f_A a podvektor hodnot $\boldsymbol{x}_A$.

Náhodné veličiny a vektory budeme označovat písmeny řecké abecedy, množiny (vyjma N) velkými písmeny ze začátku abecedy, jejich prvky písmeny malými. Vektory a matice budeme tisknout tučným písmem. Determinant čtvercové matice $\boldsymbol{\Sigma}$ budeme značit $|\boldsymbol{\Sigma}|$, její hodnost $\dim(\boldsymbol{\Sigma})$, inverzi a zobecněnou inverzi po řadě $\boldsymbol{\Sigma}^{-1}$ a $\boldsymbol{\Sigma}^{-}$.

V případě, že nebude hrozit nebezpečí omylu, budeme pro zjednodušení značení a větší čitelnost vynechávat znaménko sjednocení, tedy $\boldsymbol{\xi}_{AB} \equiv \boldsymbol{\xi}_{A \cup B}$, a složené závorky kolem jednoprvkových množin, tedy $\boldsymbol{\xi}_{Ab} \equiv \boldsymbol{\xi}_{A \cup \{b\}}$.

1.1 Podmíněná nezávislost

Existuje mnoho způsobů jak zavést pojem podmíněné nezávislosti, od intuitivních, úzce vymezených, viz [Nea03], až po matematicky rigorózní a značně obecné, viz [Stu05].

My se v této práci omezíme na případ rozdělení s hustotou f vůči $\sigma-$konečné součinové míře μ, kde navíc je buďto μ Lebesgueova a f spojitá nebo μ diskrétní. Veškerá níže uvedená tvrzení o hustotě f je třeba chápat ve smyslu $\mu-$skoro všude.

Definice 1. *Řekneme, že pro náhodný vektor $\boldsymbol{\xi} = (\xi_a)_{a\in N}$ a po dvou disjunktní podmnožiny $A, B, C \subseteq N$ je podvektor $\boldsymbol{\xi}_A$* ***podmíněně nezávislý*** *na $\boldsymbol{\xi}_B$ dáno $\boldsymbol{\xi}_C$ (značíme $\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_B | \boldsymbol{\xi}_C$) právě tehdy, když*

$$f_{ABC}(\boldsymbol{x}_{ABC}) \cdot f_C(\boldsymbol{x}_C) = f_{AC}(\boldsymbol{x}_{AC}) \cdot f_{BC}(\boldsymbol{x}_{BC}). \tag{1.1}$$

Pokud $C = \emptyset$, je podle příslušné konvence $f_\emptyset \equiv 1$ a v tomto případě se jedná o (klasickou) nepodmíněnou nezávislost (značíme $\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_B$).

Smysl formální definice $\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_B | \boldsymbol{\xi}_C$ je podobný jako u nepodmíněné nezávislosti: Jestliže již známe hodnotu $\boldsymbol{\xi}_C$, potom je hodnota $\boldsymbol{\xi}_A$ irelevantní (nepodstatná, neužitečná pro predikci, neovlivňující podmíněné rozdělení, apod.) pro hodnotu $\boldsymbol{\xi}_B$ a naopak.

Lemma 1 uvádí několik dalších způsobů zavedení pojmu podmíněné nezávislosti ekvivalentních s (1.1), viz [Lau96], str. 29. *Podmíněnou hustotou* $f_{A|B}(\boldsymbol{x}_A|\boldsymbol{x}_B)$ zde budeme značit funkci $f_{AB}(\boldsymbol{x}_{AB})/f_B(\boldsymbol{x}_B)$, definovanou pokud $f_B(\boldsymbol{x}_B) > 0$, jinak necháváme nedefinováno. Příslušným rovnostem je třeba rozumět tak, že platí na množině, kde jsou podmíněné hustoty definovány.

Lemma 1. *Nechť A, B a C jsou po dvou disjunktní podmnožiny N. Potom následující tvrzení jsou ekvivalentní:*

i) $\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_B | \boldsymbol{\xi}_C$

ii) $f_{AB|C}(\boldsymbol{x}_{AB}|\boldsymbol{x}_C) = f_{A|C}(\boldsymbol{x}_A|\boldsymbol{x}_C) \cdot f_{B|C}(\boldsymbol{x}_B|\boldsymbol{x}_C)$

iii) $f_{A|BC}(\boldsymbol{x}_A|\boldsymbol{x}_{BC}) = f_{A|C}(\boldsymbol{x}_A|\boldsymbol{x}_C)$

iv) $f_{AC|B}(\boldsymbol{x}_{AC}|\boldsymbol{x}_B) = f_{A|C}(\boldsymbol{x}_A|\boldsymbol{x}_C) \cdot f_{C|B}(\boldsymbol{x}_C|\boldsymbol{x}_B)$

v) $f_{ABC}(\boldsymbol{x}_{ABC}) = f_{A|C}(\boldsymbol{x}_A|\boldsymbol{x}_C) \cdot f_{BC}(\boldsymbol{x}_{BC})$

vi) $\exists g, h$ funkce : $f_{ABC}(\boldsymbol{x}_{ABC}) = g(\boldsymbol{x}_{AC}) \cdot h(\boldsymbol{x}_{BC})$

Pokud má rozdělení navíc konečnou **multiinformaci** (MI), tj. Kullback–Leiblerovu divergenci (≡relativní entropii) sdruženého rozdělení vůči součinu jednorozměrných marginál (viz [Stu05], str. 35), lze charakterizovat nezávislostní vztahy za pomoci následujícího lemmatu.

Lemma 2. *Nechť $\boldsymbol{\xi}$ je vektor s konečnou multiinformací a A, B, C jsou po dvou disjunktní podmnožiny N. Potom platí*

$$MI(\boldsymbol{\xi}_{ABC}) + MI(\boldsymbol{\xi}_C) - MI(\boldsymbol{\xi}_{AC}) - MI(\boldsymbol{\xi}_{BC}) \geq 0,$$

a navíc

$$\big(MI(\boldsymbol{\xi}_{ABC}) + MI(\boldsymbol{\xi}_C) - MI(\boldsymbol{\xi}_{AC}) - MI(\boldsymbol{\xi}_{BC}) = 0\big) \iff \big(\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_B | \boldsymbol{\xi}_C\big).$$

Důkaz. Důkaz lze nalézt v [Stu05], str. 27–28. □

Za pomoci podmiňování (viz např. [Lac04]) lze podmíněnou nezávislost mezi $\boldsymbol{\xi}_A$ a $\boldsymbol{\xi}_B$ dáno $\boldsymbol{\xi}_C$ definovat „v širším smyslu“ jako platnost tvrzení, že rozdělení $\boldsymbol{\xi}_{AB}|\boldsymbol{\xi}_C$ je součinem příslušných podmíněných rozdělení, neboli

$$\mathcal{L}(\boldsymbol{\xi}_{AB}|\boldsymbol{\xi}_C = \boldsymbol{x}_C) = \mathcal{L}(\boldsymbol{\xi}_A|\boldsymbol{\xi}_C = \boldsymbol{x}_C) \otimes \mathcal{L}(\boldsymbol{\xi}_B|\boldsymbol{\xi}_C = \boldsymbol{x}_C) \quad P_C\text{–s.v.} \tag{1.2}$$

Snadno se nahlédne, že (1.2) implikuje (1.1). Definici podmíněné nezávislosti v širším smyslu využijeme jen v případě neregulárního Gaussova rozdělení.

Jak ukazuje následující lemma, podmíněnou nezávislost je možné charakterizovat pomocí souboru *elementárních nezávislostních vztahů*, tj. souboru podmíněných nezávislostí $(\boldsymbol{\xi}_{A_i} \perp\!\!\!\perp \boldsymbol{\xi}_{B_i} | \boldsymbol{\xi}_{C_i})_i$ takových, že $|A_i| = |B_i| = 1$.

Lemma 3. *Pro A, B, C po dvou disjunktní podmnožiny N platí*

$$(\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_B | \boldsymbol{\xi}_C) \Leftrightarrow (\forall a \in A, b \in B, D \subseteq ABC \setminus \{a, b\} : \ C \subseteq D \Rightarrow \xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_D)\,.$$

Důkaz. Viz [Mat92], lemma 3. □

Další vlastnosti podmíněné nezávislosti budou odvozeny v oddílu 1.3.

1.2 Gaussovo, diskrétní a binární rozdělení

V tomto oddílu zavedeme tři základní typy rozdělení, se kterými se budeme v této práci setkávat: Gaussovo, diskrétní a binární.

1.2.1 Gaussovo rozdělení

Definice 2. ***Gaussovo rozdělení*** *náhodného vektoru* $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ *je pravděpodobnostní rozdělení dané svou charakteristickou funkcí*

$$\varphi_{\boldsymbol{\xi}}(\mathbf{t}) \equiv \mathrm{E}\exp(i\mathbf{t}'\boldsymbol{\xi}) = \exp\left(i\mathbf{t}'\boldsymbol{\mu} - \frac{\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}{2}\right),$$

které je parametrizované střední hodnotou $\boldsymbol{\mu} \in \mathbb{R}^n$ *a varianční (symetrickou, pozitivně semidefinitní) maticí* $\boldsymbol{\Sigma}$.

Jestliže je matice $\boldsymbol{\Sigma}$ pozitivně definitní ($\equiv$regulární), značíme $\boldsymbol{\Sigma} > 0$, budeme hovořit o **regulárním Gaussově rozdělení**. Regulární Gaussovo rozdělení má spojitou verzi hustoty vůči Lebesgueově míře tvaru

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

a konečnou multiinformaci

$$MI(\boldsymbol{\xi}) = -\frac{1}{2}\left(\log(|\boldsymbol{\Sigma}|) - \sum_{a=1}^{n}\log\sigma_{a\cdot a}\right).$$

Naopak neregulární Gaussovo rozdělení nemá hustotu vůči Lebesgueově míře a obecně ani konečnou multiinformaci. Takovéto rozdělení je degenerované ve smyslu, že existuje nenulový vektor konstant $\boldsymbol{a}$ takový, že

$$\boldsymbol{a}'(\boldsymbol{\xi} - \boldsymbol{\mu}) = 0 \quad \text{(s.j.)}, \tag{1.3}$$

viz [And78], str. 71. Tedy oborem hodnot $\boldsymbol{\xi}$ je vlastní afinní podprostor $\mathbb{R}^n$.

Pro matici $\boldsymbol{\Sigma} = (\sigma_{a\cdot b})_{a,b\in N}$ a A,B neprázdné podmnožiny N definujeme matici $\boldsymbol{\Sigma}_{A\cdot B}$ jako matici, do níž byly ze $\boldsymbol{\Sigma}$ vybrány pouze řádky příslušné prvkům A a sloupce příslušné prvkům B při zachování pořadí řádků a sloupců, neboli

$$\boldsymbol{\Sigma}_{A\cdot B} = (\sigma_{a\cdot b})_{a\in A,\ b\in B}.$$

Jak ukazuje následující lemma, třída Gaussových rozdělení je uzavřená na operace marginalizace a podmiňování.

Lemma 4. *Nechť* $\boldsymbol{\xi}$ *je náhodný vektor s Gaussovým rozdělením,* $\boldsymbol{\Sigma}$ *příslušná varianční matice,* A, B *disjunktní podmnožiny* N*:*

i) *Marginální rozdělení* $\boldsymbol{\xi}_A$ *je Gauss. rozdělení s varianční maticí* $\boldsymbol{\Sigma}_{A \cdot A}$.

ii) *Podmíněné rozdělení* $\boldsymbol{\xi}_A$ *dáno* $\boldsymbol{\xi}_B = \mathbf{x}_B$ *je Gaussovo rozdělení s varianční maticí* $\boldsymbol{\Sigma}_{A \cdot A|B} = \boldsymbol{\Sigma}_{A \cdot A} - \boldsymbol{\Sigma}_{A \cdot B} \boldsymbol{\Sigma}^{-}_{B \cdot B} \boldsymbol{\Sigma}_{B \cdot A}$, *kde* $\boldsymbol{\Sigma}^{-}_{B \cdot B}$ *je libovolná zobecněná inverze* $\boldsymbol{\Sigma}_{B \cdot B}$.

iii) *Navíc, pokud je* $\boldsymbol{\Sigma}$ *regulární, potom jsou regulární i* $\boldsymbol{\Sigma}_{A \cdot A}$ *a* $\boldsymbol{\Sigma}_{A \cdot A|B}$.

Důkaz. Pro i) a ii) viz [Lau96], str. 256. Z předpokladu regularity $\boldsymbol{\Sigma}$ přímo plyne regularita $\boldsymbol{\Sigma}_{A \cdot A}$ a $\boldsymbol{\Sigma}_{B \cdot B}$.

Aplikujeme-li na matici $\boldsymbol{\Sigma}$ Choleského rozklad (předpokládá existenci U^{-1})

$$\begin{pmatrix} R & S \\ T & U \end{pmatrix} = \begin{pmatrix} I & SU^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} R - SU^{-1}T & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} I & 0 \\ U^{-1}T & I \end{pmatrix} \tag{1.4}$$

dosazením $R = \boldsymbol{\Sigma}_{A \cdot A}$, $U = \boldsymbol{\Sigma}_{B \cdot B}$ a $S = T' = \boldsymbol{\Sigma}_{A \cdot B}$, získáme na levé straně (1.4) regulární matici $\boldsymbol{\Sigma}$ a tudíž musí být regulární i všechny matice součinu na pravé straně (1.4). Tedy je regulární i $\boldsymbol{\Sigma}_{A \cdot A|B} = R - SU^{-1}T$. $\square$

Povšimněme si, že varianční matice podmíněného rozdělení $\boldsymbol{\Sigma}_{A \cdot A|B}$ zůstává stejná bez ohledu na to, jakou hodnotou $\mathbf{x}_B$ z oboru hodnot $\boldsymbol{\xi}_B$ podmíníme. To je důležitá speciální vlastnost Gaussova rozdělení, která neplatí v případě dalších rozdělení.

Lemma 5. *Nechť* $\boldsymbol{\xi}$ *je náhodný vektor s Gaussovým rozdělením,* $\boldsymbol{\Sigma}$ *příslušná varianční matice,* A, B *a* C *po dvou disjunktní podmnožiny* N*:*

i) $(\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_C) \Leftrightarrow (\boldsymbol{\Sigma}_{A \cdot C} = \mathbf{0})$.

ii) *Pro* $\boldsymbol{\Sigma}_{B \cdot B} > 0$, $(\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_C | \boldsymbol{\xi}_B) \Leftrightarrow (\forall a \in A, c \in C : \ |\boldsymbol{\Sigma}_{aB \cdot cB}| = 0)$.

iii) *Pro* $D \subset B$ *takovou, že* $\boldsymbol{\Sigma}_{D \cdot D} > 0$ *a* $\dim(\boldsymbol{\Sigma}_{D \cdot D}) = \dim(\boldsymbol{\Sigma}_{B \cdot B})$, *platí* $(\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_C | \boldsymbol{\xi}_B) \Leftrightarrow (\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_C | \boldsymbol{\xi}_D)$.

Důkaz. První část je dobře známý fakt, viz [Lau96], str. 257. Pokud $a \in A, c \in C$ a vyjádříme-li determinant $\Sigma_{aB\cdot cB}$ za dosazení do Choleského rozkladu (1.4) $R = \boldsymbol{\Sigma}_{a\cdot c}$, $U = \boldsymbol{\Sigma}_{B\cdot B}$, $S = \boldsymbol{\Sigma}_{a\cdot B}$ a $S = \boldsymbol{\Sigma}_{B\cdot c}$, zjistíme, že pro $|\boldsymbol{\Sigma}_{B\cdot B}| \neq 0$ platí

$$|\boldsymbol{\Sigma}_{aB\cdot cB}| = 0 \Leftrightarrow \sigma_{a\cdot c} - \boldsymbol{\Sigma}_{a\cdot B}\boldsymbol{\Sigma}_{B\cdot B}^{-1}\boldsymbol{\Sigma}_{B\cdot c} = 0 \Leftrightarrow (\boldsymbol{\Sigma}_{ac\cdot ac|B})_{a\cdot c} = 0 \Leftrightarrow \xi_a \perp\!\!\!\perp \xi_c | \xi_B,$$

přičemž poslední ekvivalence vyplývá z výše zmíněné neměnnosti varianční matice $\boldsymbol{\Sigma}_{ac\cdot ac|B}$ vzhledem k podmínění libovolnou hodnotou $\boldsymbol{\xi}_B$.

Dále již stačí použít část i) tohoto lemmatu, abychom si uvědomili, že

$$\big(\forall a \in A, c \in C : \ \xi_a \perp\!\!\!\perp \xi_c | \xi_B\big) \iff \boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_C | \boldsymbol{\xi}_B.$$

Platnost části iii) snadno nahlédneme za pomoci lemmatu 4, pokud si uvědomíme, že jeden ze způsobů konstrukce zobecněné inverze $\boldsymbol{\Sigma}_{B\cdot B}^{-}$ je nalezenout výše popsanou $\boldsymbol{\Sigma}_{D\cdot D}$ plné hodnosti, invertovat ji a na zbylá místa doplnit nuly, viz [Rao73], oddíl 1b.5. Neboli při vhodném uspořádání prvků N,

$$\boldsymbol{\Sigma}_{B\cdot B}^{-} = \begin{pmatrix} \boldsymbol{\Sigma}_{D\cdot D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Tudíž, $\boldsymbol{\Sigma}_{a\cdot B}\boldsymbol{\Sigma}_{B\cdot B}^{-}\boldsymbol{\Sigma}_{B\cdot c} = \boldsymbol{\Sigma}_{a\cdot D}\boldsymbol{\Sigma}_{D\cdot D}^{-1}\boldsymbol{\Sigma}_{D\cdot c}$ a požadované tvrzení dostáváme jako důsledek charakterizace podmíněného rozdělení v lemmatu 4 ii). □

1.2.2 Diskrétní rozdělení

Definice 3. *Rozdělení náhodného vektoru $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ nazveme* ***diskrétní****, pokud existuje konečná množina taková, že vektor nabývá hodnot mimo tuto množinu s nulovou pravděpodobností (neboli rozdělení mající hustotu vůči nějaké diskrétní míře).*

Analogicky lze zavést diskrétní rozdělení definováním jeho hustoty neboli funkce p ze součinu konečných množin $\boldsymbol{X} = \prod_{a=1}^{n} X_a$ do intervalu $[0,1]$ takové, že

$$\sum_{\boldsymbol{x} \in \boldsymbol{X}} p(\boldsymbol{x}) = 1.$$

Pokud pro všechna $\boldsymbol{x} \in \boldsymbol{X}$ je $p(\boldsymbol{x}) > 0$, řekneme, že dané diskrétní rozdělení je **kladné**.

Definici podmíněné nezávislosti $\boldsymbol{\xi}_A \perp\!\!\!\perp \boldsymbol{\xi}_B | \boldsymbol{\xi}_C$ můžeme v případě diskrétně rozděleného vektoru přepsat jako platnost vztahu

$$P(\boldsymbol{\xi}_{ABC} = \boldsymbol{x}_{ABC})P(\boldsymbol{\xi}_C = \boldsymbol{x}_C) = P(\boldsymbol{\xi}_{AC} = \boldsymbol{x}_{AC})P(\boldsymbol{\xi}_{BC} = \boldsymbol{x}_{BC}) \tag{1.5}$$

pro všechna příslušná $\boldsymbol{x}_{ABC}$. Obdobně lze pro tento případ přeformulovat i všechna ekvivalentní tvrzení z lemmatu 1.

Multiinformace diskrétně rozděleného vektoru $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ je z definice rovna

$$\sum_{\boldsymbol{x}_N \in \boldsymbol{X}} P(\boldsymbol{\xi}_N = \boldsymbol{x}_N) \log \left(\frac{P(\boldsymbol{\xi}_N = \boldsymbol{x}_N)}{\prod_{a=1}^{n} P(\xi_a = x_a)} \right)$$

za konvence $0 \cdot \log 0 = 0$ a je vždy konečná.

1.2.3 Binární rozdělení

Speciálním typem diskrétního rozdělení je rozdělení binární.

Definice 4. *Rozdělení náhodného vektoru $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ nazveme **binární**, pokud pro všechna $a = 1, \ldots, n$ existuje pro ξ_a dvouprvková množina taková, že hodnot mimo tuto množinu nabývá ξ_a s nulovou pravděpodobností.*

Bez újmy na obecnosti budeme dále předpokládat, že ony dvě hodnoty, jichž každá z binárních veličin nabývá, jsou -1 a 1. Tato úmluva nám umožní vhodným způsobem binární rozdělení parametrizovat.

Podmínku (1.5) pro nezávislostní vztah $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_C$ lze v případě binárního rozdělení upravit na tvar

$$\begin{aligned} P(\xi_a = 1, \xi_b = 1, \boldsymbol{\xi}_C = \boldsymbol{x}_C) P(\xi_a = -1, \xi_b = -1, \boldsymbol{\xi}_C = \boldsymbol{x}_C) &= \\ P(\xi_a = 1, \xi_b = -1, \boldsymbol{\xi}_C = \boldsymbol{x}_C) P(\xi_a = -1, \xi_b = 1, \boldsymbol{\xi}_C = \boldsymbol{x}_C), & \end{aligned} \tag{1.6}$$

pro všechna $\boldsymbol{x}_C \in \{-1, 1\}^C$.

Třídu binárních rozdělení budeme parametrizovat pomocí momentů. Pro podmnožinu indexové množiny $A \subseteq N$ zaveďme značení

$$e_A = \mathrm{E} \prod_{a \in A} \xi_a = \sum_{\boldsymbol{x}_N \in \{-1,1\}^N} P(\boldsymbol{\xi}_N = \boldsymbol{x}_N) \prod_{a \in A} x_a, \tag{1.7}$$

přičemž pro $A = \emptyset$ položíme $\prod_{a \in A} x_a = 1$, tudíž nutně $e_\emptyset = 1$.

Lemma 6. *Nechť $(e_A)_{A \subseteq N}$ je daný soubor čísel z intervalu $[-1, 1]$ indexovaný podmnožinami množiny N takový, že $e_\emptyset = 1$. Jestliže*

$$\forall \boldsymbol{x}_N \in \{-1, 1\}^N : \quad p(\boldsymbol{x}_N) := \left(\frac{1}{2}\right)^{|N|} \sum_{A \subseteq N} \left(\prod_{a \in A} x_a \right) e_A \geq 0, \tag{1.8}$$

potom existuje právě jedno binární rozdělení s momenty $(e_A)_{A \subseteq N}$ a platí

$$P(\boldsymbol{\xi}_N = \boldsymbol{x}_N) = p(\boldsymbol{x}_N).$$

Důkaz. Předpokládejme, že nějaké binární rozdělení s danými momenty existuje, a uvažme jedno pevné $\boldsymbol{x}_N^*$. Potom přes všechna $A \subseteq N$ sečtěme $2^{|N|}$ rovnic daných vztahem (1.7) a pronásobených navíc součinem $\prod_{a\in A} x_a^*$. Na levé straně součtu rovnic dostáváme $2^{|N|}p(\boldsymbol{x}_N^*)$, zatímco na pravé straně

$$\sum_{\boldsymbol{x}_N\in\{-1,1\}^N} P(\boldsymbol{\xi}_N = \boldsymbol{x}_N)\left(\sum_{A\subseteq N}\prod_{a\in A} x_a x_a^*\right) = P(\boldsymbol{\xi}_N = \boldsymbol{x}_N^*)2^{|N|},$$

přičemž jsme využili, že pro $\boldsymbol{x} \neq \boldsymbol{x}^*$ je $\sum_{A\subseteq N}\prod_{a\in A} x_a x_a^* = 0$.

Zbývá ukázat, že rozdělení splňující (1.8) bude mít skutečně požadované momenty. Uvažme libovolnou podmnožinu indexové množiny $B \subseteq N$, potom

$$\mathrm{E}\prod_{b\in B}\xi_b = \left(\frac{1}{2}\right)^{|B|}\sum_{\boldsymbol{x}_B\in\{-1,1\}^B}\sum_{A\subseteq B} e_A\left(\prod_{b\in B} x_b\right)\left(\prod_{a\in A} x_a\right) = e_B,$$

kde jsme využili prohození sum a fakt, že pro A vlastní podmnožinu B platí

$$\sum_{\boldsymbol{x}_B\in\{-1,1\}^B}\left(\prod_{b\in B} x_b\right)\left(\prod_{a\in A} x_a\right) = 0.$$

□

Kupříkladu pro dvě náhodné veličiny, neboli $N = \{1,2\}$, dostáváme

$$\begin{array}{lcllcl} P(\xi_1 = 1, \xi_2 = 1) & = & \frac{1+e_1+e_2+e_{12}}{4}, & P(\xi_1 = -1, \xi_2 = -1) & = & \frac{1-e_1-e_2+e_{12}}{4}, \\ P(\xi_1 = 1, \xi_2 = -1) & = & \frac{1+e_1-e_2-e_{12}}{4}, & P(\xi_1 = -1, \xi_2 = 1) & = & \frac{1-e_1+e_2-e_{12}}{4}, \end{array}$$

přičemž, jak lze snadno nahlédnout, je podmínka existence (1.8) ekvivalentní s podmínkou

$$1 - |e_1 - e_2| \geq e_{12} \geq -1 + |e_1 + e_2|.$$

Bohužel pro $|N| > 2$ takovýto elegantní způsob ověření podmínky (1.8) není znám.

Dále si povšimněme, že pokud pro všechny neprázdné $A \subseteq N$ jsou e_A blízko nuly, potom je podmínka (1.8) vždy splněna. Pokud jsou všechny momenty e_A rovny nule, jedná se o rovnoměrné rozdělení.

Ukažme si, jak při parametrizaci pomocí momentů vypadá charakterizace podmíněné nezávislosti.

Lemma 7. *Nechť* $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ *je binární náhodný vektor a* $(e_A)_{A\subseteq N}$ *příslušný soubor jeho momentů. Potom platí, že* $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_C$ *právě tehdy, když pro všechna* $\boldsymbol{x}_C \in \{-1,1\}^C$

$$\left(\sum_{D\subseteq C} e_{abD} \prod_{d\in D} x_d\right)\cdot\left(\sum_{D\subseteq C} e_{D} \prod_{d\in D} x_d\right) = \left(\sum_{D\subseteq C} e_{aD} \prod_{d\in D} x_d\right)\cdot\left(\sum_{D\subseteq C} e_{bD} \prod_{d\in D} x_d\right).$$

Důkaz. Pro marginální rozdělení $\boldsymbol{\xi}_{abC}$ dosadíme za pravděpodobnosti z (1.6) za pomoci vztahu (1.8)

$$\left(\sum_{D\subseteq C} (e_D + e_{aD} + e_{bD} + e_{abD}) \prod_{d\in D} x_d\right)\cdot\left(\sum_{D\subseteq C} (e_D - e_{aD} - e_{bD} + e_{abD}) \prod_{d\in D} x_d\right) =$$
$$\left(\sum_{D\subseteq C} (e_D - e_{aD} + e_{bD} - e_{abD}) \prod_{d\in D} x_d\right)\cdot\left(\sum_{D\subseteq C} (e_D + e_{aD} - e_{bD} - e_{abD}) \prod_{d\in D} x_d\right)$$

Zbývá již jen drobná technická úprava a to každou z řad rozdělit na čtyři řady $R_\emptyset$, R_a, R_b, R_{ab} rozsekáním součtu momentů v závorkách na jednotlivé členy, schematicky dostaneme $(R_\emptyset + R_a + R_b + R_{ab})\cdot(R_\emptyset - R_a - R_b + R_{ab}) = (R_\emptyset - R_a + R_b - R_{ab})\cdot(R_\emptyset + R_a - R_b - R_{ab})$ a odtud odečtením členů vyskytujících se na obou stranách získáme požadovanou rovnost $R_\emptyset \cdot R_{ab} = R_a \cdot R_b$. $\square$

Důkaz lemmatu 7 demonstruje výhodu parametrizace pomocí momentů, kterou je snadný přechod k marginálnímu rozdělení. Bez důkazu zmiňme i možnost (přirozené) konstrukce nestranných, nekorelovaných odhadů jednotlivých parametrů (momentů).

Rozepišme si ještě vztah pro podmíněnou nezávislost $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_C$ z lemmatu 7 pro $|C| = 0,1,2$:

- $\xi_a \perp\!\!\!\perp \xi_b \iff e_{ab} = e_a e_b$
- $\xi_a \perp\!\!\!\perp \xi_b | \xi_c \iff$ $$\begin{aligned} e_{abc} + e_{ab}e_c &= e_a e_{bc} + e_b e_{ac} \\ e_{ab} + e_{abc}e_c &= e_{ac}e_{bc} + e_a e_b \end{aligned}$$
- $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd} \iff$
$$\begin{aligned} e_{abcd} + e_{abc}e_d + e_{abd}e_c + e_{ab}e_{cd} &= e_a e_{bcd} + e_b e_{acd} + e_{ac}e_{bd} + e_{ad}e_{bc} \\ e_{abc} + e_{abd}e_{cd} + e_{ab}e_c + e_d e_{abcd} &= e_{ad}e_{bcd} + e_{bd}e_{acd} + e_a e_{bc} + e_b e_{ac} \\ e_{abd} + e_{abc}e_{cd} + e_{ab}e_d + e_c e_{abcd} &= e_{ac}e_{bcd} + e_{bc}e_{acd} + e_a e_{bd} + e_b e_{ad} \\ e_{ab} + e_c e_{abc} + e_d e_{abd} + e_{cd}e_{abcd} &= e_a e_b + e_{ac}e_{bc} + e_{ad}e_{bd} + e_{acd}e_{bcd} \end{aligned}$$

Pokud místo obyčejných momentů použijeme parametrizaci centrovanými momenty
$$f_A = \mathrm{E} \prod_{a \in A} (\xi_a - e_a),$$
dostaneme po dosazení
$$\begin{aligned} e_{ab} &= f_{ab} + e_a e_b \\ e_{abc} &= f_{abc} + e_a f_{bc} + e_b f_{ac} + e_c f_{ab} + e_a e_b e_c \end{aligned}$$
a následné úpravě

- $\xi_a \perp\!\!\!\perp \xi_b \iff f_{ab} = 0$
- $\xi_a \perp\!\!\!\perp \xi_b | \xi_c \iff \begin{aligned} (1 - e_c^2) f_{ab} &= f_{ac} f_{bc} \\ f_{abc} &= -2 e_c f_{ab} \end{aligned}$

Obdobně lze dosadit
$$e_{abcd} = f_{abcd} + \sum_{i \in \{a,b,c,d\}} (e_i f_{abcd \setminus i}) + \frac{1}{2} \sum_{i,j \in \{a,b,c,d\},\ i \neq j} (e_i e_j f_{abcd \setminus ij}) + e_a e_b e_c e_d$$
do výše uvedených rovnic pro $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, avšak výsledná soustava rovnic je složitější než ta pro necentrované momenty.

Pokud navíc platí, že $e_a = e_b = e_c = e_d = 0$, jsou centrované a necentrované momenty totožné a výše uvedené se zjednoduší vztahy na

- $\xi_a \perp\!\!\!\perp \xi_b \iff e_{ab} = 0$
- $\xi_a \perp\!\!\!\perp \xi_b | \xi_c \iff \begin{aligned} e_{abc} &= 0 \\ e_{ab} &= e_{ac} e_{bc} \end{aligned}$
- $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd} \iff$
$$\begin{aligned} e_{abcd} + e_{ab} e_{cd} &= e_{ac} e_{bd} + e_{ad} e_{bc} \\ e_{abc} + e_{abd} e_{cd} &= e_{ad} e_{bcd} + e_{bd} e_{acd} \\ e_{abd} + e_{abc} e_{cd} &= e_{ac} e_{bcd} + e_{bc} e_{acd} \\ e_{ab} + e_{cd} e_{abcd} &= e_{ac} e_{bc} + e_{ad} e_{bd} + e_{acd} e_{bcd} \end{aligned}$$

Na závěr si povšimněme zajímavého faktu, že v případě $|N| = 4$ při volbě $\forall a, b, c \in N : \; e_a = 0, e_{abc} = 0$ a případně $e_{abcd} = e_{ac} e_{bd} + e_{ad} e_{bc} - e_{ab} e_{cd}$ dostáváme rovnice pro e_{ab} charakterizující podmíněnou nezávislost totožné s rovnicemi pro koeficienty korelační matice $\sigma_{a \cdot b}$ u regulárního Gaussova rozdělení odvozených dle lemmatu 5.

Zajímavá by mohla být též parametrizace za pomoci Möbiovy transformace necentrovaných momentů, to jest $g_D = \sum_{A \subseteq D} (-1)^{|D \setminus A|} e_A$ pro $D \subseteq N$.

1.3 Nezávislostní modely

V tomto oddílu bude zaveden formální „nezávislostní model". Je třeba přiznat, že tento termín není ustálený a v literatuře se ve stejném nebo podobném významu setkáme se „strukturou podmíněné nezávislosti", [Stu97], „objektem podmíněné nezávislosti", [Jir03], případně „seznamem podmíněných nezávislostí", [CGH96]. Značení zde je inspirováno pracemi [RS02] a [Mat97].

Definice 5. *Množinou tripletů* $\mathcal{T}_N$ *nad konečnou množinou* N *budeme rozumět množinu všech dvojic* $\langle ab|C\rangle$ *takových, že* ab *je (neuspořádaná) dvojice dvou různých prvků z* N *a* C *je podmnožinou* $N \setminus ab$.

Jednotlivé triplety odpovídají elementárním nezávislostním vztahům, tak jak byly zavedeny na konci oddílu 1.1.

Definice 6. ***Nezávislostním modelem*** *nad konečnou množinou* N *je myšlena libovolná podmnožina množiny* $\mathcal{T}_N$*. Nezávislostní model* $\mathcal{I}(\boldsymbol{\xi})$ *odpovídající rozdělení náhodného vektoru* $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$ *je nezávislostní model nad* $N = \{1, \ldots, n\}$ *definovaný následovně*

$$\mathcal{I}(\boldsymbol{\xi}) = \{\langle ab|C\rangle;\ \xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_C\}\,.$$

Zdůrazněme, že nezávislostní model $\mathcal{I}(\boldsymbol{\xi})$ určuje podle lemmatu 3 jednoznačně i všechny ostatní podmíněné nezávislosti mezi podvektory $\boldsymbol{\xi}$.

Definice 7. *Pokud pro nezávislostní model* I *existuje náhodný vektor* $\boldsymbol{\xi}$

i) s Gaussovým rozdělením

ii) s diskrétním rozdělením

iii) s binárním rozdělením

takový, že $I = \mathcal{I}(\boldsymbol{\xi})$*, budeme* I *nazývat*

i) ***Gaussovsky reprezentovatelný*** *(g–reprezentovatelný)*

ii) ***diskrétně reprezentovatelný*** *(d–reprezentovatelný)*

iii) ***binárně reprezentovatelný*** *(b–reprezentovatelný)*

a $\boldsymbol{\xi}$*, respektive jeho rozdělení, budeme nazývat* ***reprezentací*** *tohoto modelu.*

Pokud budeme navíc chtít zdůraznit, že nás zajímají pouze regulární, resp. kladné reprezentace, budeme hovořit o **pozitivní** reprezentovatelnosti a nad příslušnou zkratku umístíme znaménko „+“ (g^+–reprezentovatelnost, d^+–reprezentovatelnost, b^+–reprezentovatelnost).

Pokud použijeme pojem „reprezentovatelné“ bez přívlastku, budeme mít na mysli obecnou pravděpodobnostní reprezentovatelnost (pro účely této práce je možné tento pojem chápat omezeně jako reprezentovatelnost jedním z výše uvedených způsobů).

Lemma 8. *Nechť a, b, c jsou různé prvky N a D je podmnožinou $N \setminus abc$. Pro reprezentovatelný model I platí*

$$\big(\{\langle ab|cD\rangle, \langle ac|D\rangle\} \subseteq I\big) \Leftrightarrow \big(\{\langle ac|bD\rangle, \langle ab|D\rangle\} \subseteq I\big). \tag{1.9}$$

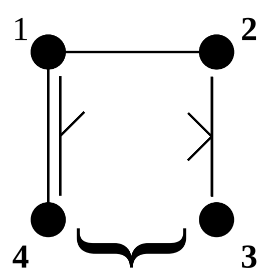
Pokud je model navíc pozitivně reprezentovatelný, potom

$$\big(\{\langle ab|cD\rangle, \langle ac|bD\rangle\} \subseteq I\big) \Rightarrow \big(\{\langle ab|D\rangle, \langle ac|D\rangle\} \subseteq I\big). \tag{1.10}$$

Důkaz. Jedná se o tzv. „semigrafoidovou“, resp. „grafoidovou“ vlastnost, viz [Lau96] nebo podrobněji [Stu05]. □

Popsání nezávislostního modelu výčtem tripletů je špatně uchopitelné a nepříliš názorné. Pro $|N| \leq 4$ můžeme použít grafické znázornění nezávislostního modelu „diagramy“ navrženými R. Lněničkou.

Každému prvku N přiřadíme jeden bod. Jestliže I obsahuje $\langle ab|\emptyset\rangle$, spojíme body odpovídající a a b čarou. Pokud je mezi prvky I triplet $\langle ab|c\rangle$, spojíme a a b čarou a doprostřed přilepíme „zobáček“ směřující k c. Jestliže I obsahuje jak $\langle ab|c\rangle$, tak $\langle ab|d\rangle$, nakreslíme obě čáry přes sebe a připojíme „dvojzobáček“ ukazující k c i d. A konečně pokud $\langle ab|cd\rangle \in I$, spojíme a a b svorkou.



Obrázek 1.1: Příklad grafického znázornění nezávislostního modelu.

Lépe to bude pochopitelné na příkladu: Nezávislostní model zobrazený na obr. 1.1 na str. 19 je model

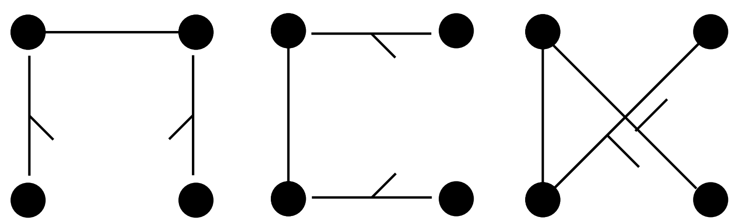
$$\{\langle 12|\emptyset\rangle, \langle 23|1\rangle, \langle 23|4\rangle, \langle 34|12\rangle, \langle 14|\emptyset\rangle, \langle 14|2\rangle\}.$$

Definice 8. *Dva nezávislostní modely I (nad množinou M) a J (nad množinou N) nazveme* ***izomorfní****, jestliže existuje bijektivní zobrazení π množiny M na množinu N takové, že*

$$\langle ab|C\rangle \in I \Leftrightarrow \langle \pi(a)\pi(b)|\pi(C)\rangle \in J,$$

kde pod $\pi(C)$ rozumíme $\{\pi(c);\ c \in C\}$. Třídu rozkladu nezávislostních modelů podle ekvivalence „býti izomorfní“ budeme nazývat ***typ****.*

Izomorfismus je z diagramů snadno rozpoznatelný[1]. Pro ilustraci, všechny modely znázorněné na obr. 1.2 jsou stejného typu, tj. navzájem izomorfní.



Obrázek 1.2: Tři izomorfní nezávislostní modely.

Nečiní problém nahlédnout, že modely daného typu jsou buďto všechny reprezentovatelné nebo naopak není reprezentovatelný žádný z nich. To nám umožňuje hovořit o reprezentovatelnosti typů.

Nezávislostní model I můžeme zúžit na podmnožinu proměnných, případně vybrat pouze ty triplety, jejichž podmínková část zahrnuje zadanou podmnožinu proměnných. Jedná se o analogie operací marginalizace a podmínění pro pravděpodobnostní rozdělení. Výsledné modely budeme nazývat minory[2] I.

[1]Jedná se vlastně o permutaci náhodných veličin.

[2]Raději zdůrazněme, že pojem minor zde používáme tak, jak byl zaveden v [Mat97], nikoli tedy ve významu determinant podmatice.

Definice 9. *Nechť I je nezávislostní model nad N a E, F jsou disjunktní podmnožiny N. Potom definujeme **minor** $I\,[^F_E$ jako nezávislostní model nad $N \setminus EF$, do kterého jsou vybrány právě ty triplety neobsahující žádný z prvků E a obsahující v podmínce F z jejichž podmínkové části je následně F vypuštěno, formálně*

$$I\,[^F_E = \{\langle ab|C\rangle;\ E \cap (abC) = \emptyset, \langle ab|CF\rangle \in I\}\,.$$

Pokud budeme chtít zdůraznit, že se jedná o minor nad množinou o k prvcích (neboli, že $|N| - |EF| = k$), budeme mluvit o k–minorech.

Povšimněme si, že

$$I\,[^\emptyset_\emptyset = I$$

a taktéž, pro $E_1F_1 \cap E_2F_2 = \emptyset$,

$$\left(I\,[^{F_1}_{E_1}\right)[^{F_2}_{E_2} = I\,[^{F_1F_2}_{E_1E_2}\,. \tag{1.11}$$

Třídu nezávislostních modelů nazveme **uzavřenou na minory**, jestliže s každým svým modelem obsahuje i všechny jeho minory. Jak bude ukázáno v dalších dvou kapitolách, třídy Gaussovsky a diskrétně reprezentovatelných modelů jsou uzavřené na minory, zatímco třída binárně reprezentovatelných modelů na minory uzavřená není.

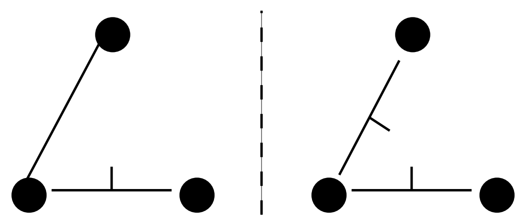
1.4 Charakterizace třídy nezávislostních modelů pomocí seznamu zakázaných minorů

Charakterizovat třídu reprezentovatelných modelů nemusí být vždy jednoduché. Původní hypotéza J. Pearla v [Pea98], že jediné podmínky (pozitivní) reprezentovatelnosti jsou (1.9) a (1.10) uvedené v lemmatu 8, se ukázala býti mylná. Krom řady dalších vlastností reprezentovatelných modelů, publikovaných např. v [Spo94] a [Stu89], existuje důkaz v [Stu92] (pro podrobněji okomentovanou verzi, viz [SV98]), že třída d–reprezentovatelných modelů nemá konečnou charakterizaci pomocí konečného počtu inferenčních pravidel (jistého typu). V této práci se pokusíme k podobnému závěru dojít i pro třídy g–reprezentovatelných a b–reprezentovatelných modelů, avšak místo jazyka inferenčních pravidel budeme používat charakterizaci seznamem zakázaných minorů navrženou v [Mat97].

Definice 10. *O nezávislostním modelu J nad množinou M řekneme, že je **zakázaným minorem** pro třídu nezávislostních modelů $\mathcal{C}$, jestliže žádný z modelů v $\mathcal{C}$ nemá $|M|$–minor izomorfní s modelem J.*

*Třída nezávislostních modelů $\mathcal{C}$ je **charakterizovaná seznamem zakázaných minorů** $\mathcal{D}$, jestliže pro libovolný nezávislostní model I platí, že I je v $\mathcal{C}$ právě tehdy, když žádný z minorů I není izomorfní s žádným z prvků $\mathcal{D}$.*

Kupříkladu dle lemmatu 8 je model $\{\langle 12|3\rangle, \langle 13|\emptyset\rangle\}$, znázorněný vlevo na obr. 1.3, (ale i řada dalších, např. $\{\langle 12|3\rangle, \langle 13|2\rangle, \langle 13|\emptyset\rangle\}$) zakázaným minorem pro třídy nezávislostních modelů, jež jsou d–, b– nebo g–reprezentovatelné. Navíc, model $\{\langle 12|3\rangle, \langle 13|2\rangle\}$, znázorněný vpravo na obr. 1.3, je zakázaným minorem pro třídu kladně reprezentovatelných nezávislostních modelů.



Obrázek 1.3: Dva příklady zakázaných minorů.

Definice 11. *Třída nezávislostních modelů je **konečně charakterizovatelná**, pokud ji lze charakterizovat za pomoci konečného seznamu zakázaných minorů.*

Poznamenejme, že tato definice je mírně odlišná od definice konečné charakterizace uvedené v [Stu92], která je v termínech speciálních inferenčních pravidel.

V našem případě jde o charakterizaci v temínech zakázaných minorů, a tak má pojem rozumný smysl pouze pro třídy nezávislostních modelů, které jsou na minory uzavřené.

Lemma 9. *Existuje-li nekonečný seznam (navzájem neizomorfních) zakázaných minorů $\mathcal{E}$ pro třídu nezávislostních modelů $\mathcal{C}$ takový, že pro každý model $J \in \mathcal{E}$ nad množinou M leží všechny jeho $(|M|-1)$–minory v $\mathcal{C}$, potom třída $\mathcal{C}$ není konečně charakterizovatelná.*

Důkaz. Předpokládejme sporem, že existuje konečný seznam zakázaných minorů $\mathcal{D}$ charakterizující třídu $\mathcal{C}$. Každý model $D \in \mathcal{D}$ je uvažován nad nějakou indexovou množinou M_D. Nechť m je maximální kardinalita M_D přes všechny $D \in \mathcal{D}$. Jelikož $\mathcal{E}$ je nekonečná a složená z neizomorfních modelů, musí v ní existovat model $E \in \mathcal{E}$ nad množinou s kardinalitou $n > m$ a tedy neizomorfní žádnému prvku $\mathcal{D}$. Spor spočívá v tom, že tento model současně patří i nepatří do $\mathcal{C}$.

Závěr $E \notin \mathcal{C}$ plyne z toho, že E je zakázaný minor pro $\mathcal{C}$. Naopak $E \in \mathcal{C}$, neboť předpokládáme, že $(n-1)$–minory E patří do $\mathcal{C}$ a tedy ani ony, a dle identity (1.11) ani jiné vlastní minory E nižší kardinality nemohou mít izomorfní kopii v $\mathcal{D}$. $\square$

Pokud třída nezávislostních modelů není konečně charakterizovatelná, zůstává jedinou nadějí k popsání jejich prvků omezit se na modely nad množinou N o pevně dané velikosti. Jak v následujících dvou kapitolách uvidíme, tato úloha je pro třídy modelů reprezentovatelných v daném distribučním rámci triviální pro $|N| \leq 3$, obtížná pro $|N| = 4$ a prakticky neřešitelná pro $|N| \geq 5$.

Kapitola 2

Gaussovsky reprezentovatelné nezávislostní modely

V této kapitole nalezneme všechny Gaussovsky reprezentovatelné nezávislostní modely pro $|N| \leq 4$. Tato práce navazuje na výsledky z [Mat05] a [Lně05], kde byly charakterizovány všechny g^+–reprezentovatelné modely pro po řadě $|N| = 3$ a $|N| = 4$. Ve srovnání s [Lně05] zde budeme prezentovat alternativní, matematicky méně elegantní zato však více přímočaré řešení charakterizace g^+–reprezentovatelných modelů a toto poté rozšíříme na charakterizaci g–reprezentovatelných modelů (pro $|N| = 4$). Na závěr dokážeme, že obecně třída g–reprezentovatelných nezávislostních modelů nemá konečnou charakterizaci a zmíníme několik zajímavých otevřených problémů.

Pro $|N| \leq 2$ jsou všechny možné modely zjevně reprezentovatelné. Začněme tedy s rozborem pro $|N| = 3$. Modely, jež vyhovují podmínce (1.9) z lemmatu 8 jsou **9** typů, které jsou znázorněny na obr. 2.1 na str. 25.

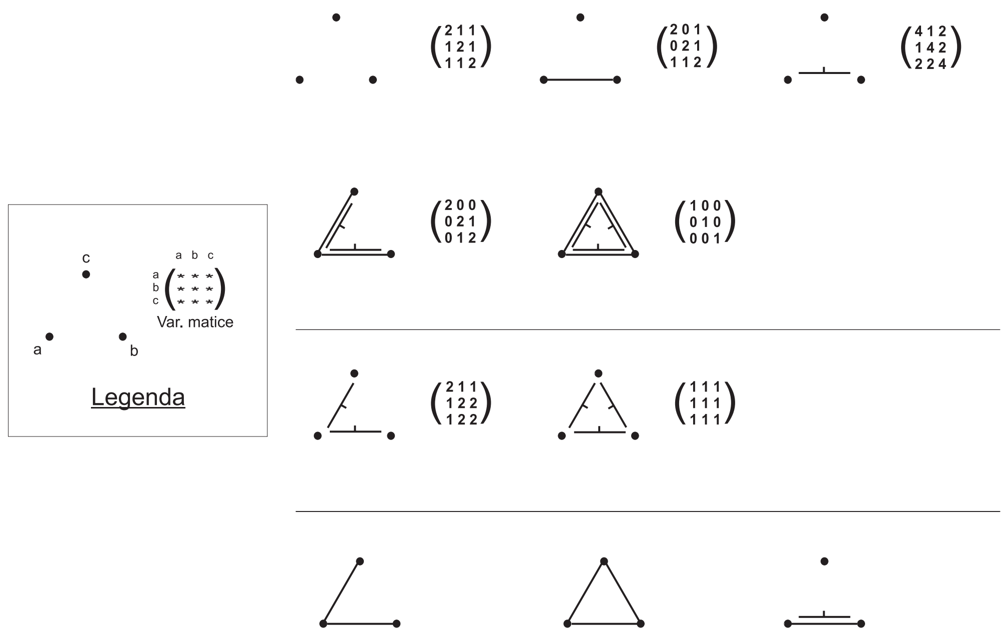
Prvních **5** typů modelů je skutečně g^+–reprezentovatelných. Další **2** pozitivně reprezentovatelné býti nemohou, neboť nesplňují vlastnost (1.10) z lemmatu 8, ale jsou alespoň g–reprezentovatelné. Variační matice příslušných reprezentací na obr. 2.1 odpovídají řazení vrcholů v diagramech proti směru hodinových ručiček s počátkem v levém dolním rohu.

Není obtížné dokázat, že poslední tři typy nejsou g–reprezentovatelné.

Lemma 10. *Nechť a, b a c jsou různé prvky N. Nezávislostní modely*

i) $I_1 = \{\langle ab|\emptyset\rangle, \langle ac|\emptyset\rangle\}$

ii) $I_2 = \{\langle ab|\emptyset\rangle, \langle ac|\emptyset\rangle, \langle bc|\emptyset\rangle\}$



Obrázek 2.1: Po řadě od shora g^+–reprezentovatelné, zbylé g–reprezentovatelné a g–nereprezentovatelné typy pro $|N| = 3$.

iii) $I_3 = \{\langle ab|\emptyset\rangle, \langle ab|c\rangle\}$

nejsou g–reprezentovatelné.

Důkaz. Pokud by $\boldsymbol{\xi} = (\xi_a, \xi_b, \xi_c)$ byl náhodný vektor s Gaussovým rozdělením a varianční maticí $\boldsymbol{\Sigma}$ reprezentující jeden z těchto typů, je zjevné, že žádná z jeho tří složek by nemohla mít nulový rozptyl (neboli být neregulární, čili degenerovaná, konstatní). V opačném případě by totiž tato složka ξ_a byla automaticky nezávislá se zbylými dvěmi složkami ξ_b a ξ_c neboli dle lemmatu 3

$$\xi_a \perp\!\!\!\perp (\xi_b, \xi_c) \Leftrightarrow \big(\xi_a \perp\!\!\!\perp \xi_b \ \&\ \xi_a \perp\!\!\!\perp \xi_c \ \&\ \xi_a \perp\!\!\!\perp \xi_b|\xi_c \ \&\ \xi_a \perp\!\!\!\perp \xi_c|\xi_b\big),$$

což by byl spor.

i,ii) Dále pokud u vektoru bez degenerovaných složek platí $\xi_a \perp\!\!\!\perp \xi_b$ a $\xi_a \perp\!\!\!\perp \xi_c$, což je podle první části lemmatu 5 ekvivalentní s $\sigma_{a\cdot b} = \sigma_{a\cdot c} = 0$, potom nutně platí dle druhé části lemmatu 5 i

$$\sigma_{a\cdot b}\sigma_{c\cdot c} - \sigma_{a\cdot c}\sigma_{b\cdot c} = \sigma_{a\cdot c}\sigma_{b\cdot b} - \sigma_{a\cdot b}\sigma_{b\cdot c} = 0$$

neboli $\xi_a \perp\!\!\!\perp \xi_b | \xi_c$ a $\xi_a \perp\!\!\!\perp \xi_c | \xi_b$. To je ve však sporu s tím, že I_1 a I_2 jsou reprezentovatelné.

iii) Analogicky pokud platí $\xi_a \perp\!\!\!\perp \xi_b$ a $\xi_a \perp\!\!\!\perp \xi_b | \xi_c$, což je podle lemmatu 5 ekvivalentní s $\sigma_{a \cdot b} = 0$ a $\sigma_{a \cdot b} - \sigma_{a \cdot c}\sigma_{b \cdot c} = 0$, potom nutně $\sigma_{a \cdot c} = 0$ nebo $\sigma_{b \cdot c} = 0$ a tudíž $\xi_a \perp\!\!\!\perp \xi_c$ nebo $\xi_b \perp\!\!\!\perp \xi_c$. To je však ve sporu s tím, že I_3 je reprezentovatelný.

□

Jak ukazuje následující lemma, třídy g^+–reprezentovatelných a g–reprezentovatelných nezávislostních modelů jsou uzavřené na minory.

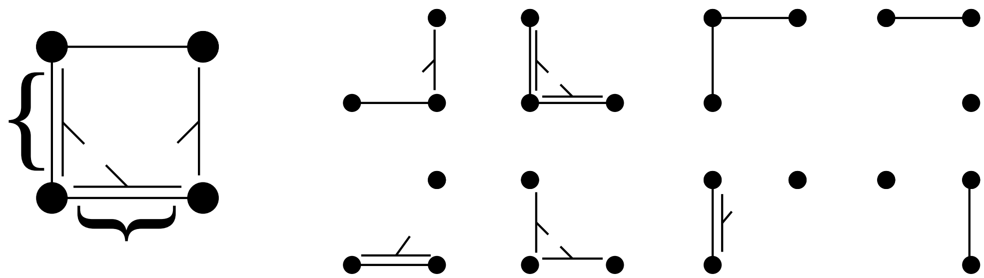
Lemma 11. *Jestliže nezávislostní model I je g–reprezentovatelný (resp. g^+–reprezentovatelný), potom i všechny minory I jsou g–reprezentovatelné (resp. g^+–reprezentovatelné).*

Důkaz. Nechť $\boldsymbol{\xi}$ je g–reprezentace I s varianční maticí $\boldsymbol{\Sigma}$. Potom $I\,[^F_E$ je g–reprezentován Gaussovým rozdělením s varianční maticí $\boldsymbol{\Sigma}_{N \setminus EF \cdot N \setminus EF | F}$, viz lemma 4, první a druhá část. Pokud je navíc $\boldsymbol{\Sigma}$ regulární, potom se podle lemmatu 4 iii) jedná o pozitivní reprezentaci. □

2.1 Regulární Gaussovská reprezentovatelnost

Podle lemmatu 11 musí být minory g^+–reprezentovatelného modelu, taktéž g^+–reprezentovatelné. Připomeňme, že pro $|N| = 3$ existuje 5 g^+–reprezentovatelných typů nezávislostních modelů, viz horní část obr. 2.1. Nutnou podmínkou pro g^+–reprezentovatelnost nezávislostního modelu nad $N = \{1, 2, 3, 4\}$ je tedy, že jeho osm 3–minorů je jednoho z pěti nám již známých typů. Zároveň si uvědomme elementární, ale důležitý fakt, že každý model nad 4-prvkovou množinou je jednoznečně určen svými 3-minory. Pro ilustraci viz obr. 2.2.

Jak lze za pomoci počítače snadno ověřit, existuje **58** typů modelů, jejichž 3–minory jsou g^+–reprezentovatelné (viz M1–M53 na obr. 2.3 a M54–M58 na obr. 2.4). Nastiňme alespoň základní myšlenku, jak byla počítačová implementace realizována: Protože projít všechny nezávislostní modely nad $|N| = 4$ a testovat jejich 3–minory by bývalo příliš časově náročné, postupovali jsme z opačného konce. Snažili jsme se najít osm g^+–reprezentovatelných modelů nad tříprvkovou množinou (tj. vybírali jsme z 11 modelů rozpadajících se do 8 typů, viz obr. 2.2), tak aby tyto tvořily příslušné minory



Obrázek 2.2: Model I spolu s 3–minory $I\,[^{\emptyset}_{e}$ a $I\,[^{f}_{\emptyset}$, kde $e, f = 1, \ldots, 4$.

nezávislostního modelu nad $N = \{1, 2, 3, 4\}$. Stačilo by nám tedy prozkoumat $11^8 = 214358881$ možností výběru minorů. Ve skutečnosti jsme využili některých symetrií a toto číslo řádově snížili.

Dalším krokem bylo nalézt reprezentace co nejvíce z těchto typů, resp. příslučné varianční matice těchto reprezentací. Prověřeny byly všechny symetrické, pozitivně definitní matice mající na všech diagonálních pozicích stejné číslo a to číslo menší nebo rovno 24. Nalezeny byly g^+–reprezentace pro **53** typů, viz M1–M53 na obr. 2.3, str. 39.

Zbylých **5** typů (viz M54–M58 na obr. 2.4) nemůže být g^+–reprezentovatelných, neboť neodpovídají pravidlům odvozeným na základě „nezávislostní implikace“. Tento komplexní nástroj, použitelný pro libovolné rozdělení s konečnou multiinformací, je založený na tzv. „strukturálních imsetech“, celočíselných vektorech reprezentujících nezávislostní modely. Teorie nutná pro jeho pochopení přesahuje rámec této práce, viz [Stu05], str. 114. Avšak bez jakýchkoli formálních znalostí je možné využít „nezávislostní implikace“ díky java appletu

$$\texttt{http}://\texttt{staff.utia.cas.cz/studeny/VerifyView.html}\ .$$

Navíc, v příštím oddílu bude přímo dokázáno, že dotyčných 5 typů není ani g–reprezentovatelných. Shrňme tedy dosažené výsledky:

Věta 1. *Nad množinou $N = \{1, 2, 3, 4\}$ existuje 629 g^+–reprezentovatelných nezávislostních modelů, které se rozpadají do 53 typů.*

Pro inferenční pravidla charakterizující g^+–reprezentovatelné modely viz lemma 16 na konci oddílu 2.3, str. 34.

Grafické znázornění g^+–reprezentovatelných typů spolu s varianční maticí vybrané reprezentace je na obr. 2.3, str. 39. Varianční matice reprezentací odpovídají řazení vrcholů v diagramech po směru hodinových ručiček s počátkem v levém horním rohu. Počet izomorfních modelů příslušných každému z g–reprezentovatelných typů lze nalézt v tabulce 2.1 na straně 30.

2.2 Obecná Gaussovská reprezentovatelnost

V tomto oddílu budou nalezeny všechny g–reprezentovatelné modely nad množinou $N = \{1, 2, 3, 4\}$. Klíčová jsou následující dvě pomocná tvrzení.

Lemma 12. *Jestliže je model I g–reprezentovatelný $I = \mathcal{I}(\boldsymbol{\xi})$, potom existuje jeho g–reprezentace $\boldsymbol{\xi}^* = (\xi_1^*, \ldots, \xi_4^*)$ taková, že*

$$\operatorname{Var} \xi_x^* = 1, \quad x = 1, \ldots, 4.$$

Důkaz. Pokud je nějaká složka degenerovaná (s nulovým rozptylem), potom je z definice nezávislá se zbytkem vektoru. Abychom se vyhnuli komplikacím, zaměníme proto nejprve všechny složky ξ_a s nulovým rozptylem (konstanty) za novou náhodnou náhodnou veličinu s nenulovým rozptylem nezávislou se zbytkem náhodného vektoru. Následně stačí definovat

$$\boldsymbol{\xi}^* = \left((\operatorname{Var} \xi_1)^{-\frac{1}{2}} \cdot \xi_1, (\operatorname{Var} \xi_2)^{-\frac{1}{2}} \cdot \xi_2, (\operatorname{Var} \xi_3)^{-\frac{1}{2}} \cdot \xi_3, (\operatorname{Var} \xi_4)^{-\frac{1}{2}} \cdot \xi_4 \right).$$

Snadno se ověří, že $I(\boldsymbol{\xi}) = I(\boldsymbol{\xi}^*)$. □

Díky lemmatu 12 se tedy lze při hledání maticových reprezentací omezit na matice mající na diagonále jedničky neboli na

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \sigma_{1\cdot 2} & \sigma_{1\cdot 3} & \sigma_{1\cdot 4} \\ \sigma_{2\cdot 1} & 1 & \sigma_{2\cdot 3} & \sigma_{2\cdot 4} \\ \sigma_{3\cdot 1} & \sigma_{3\cdot 2} & 1 & \sigma_{3\cdot 4} \\ \sigma_{4\cdot 1} & \sigma_{4\cdot 2} & \sigma_{4\cdot 3} & 1 \end{pmatrix}, \tag{2.1}$$

kde ze symetrie varianční matice $\sigma_{a\cdot b} = \sigma_{b\cdot a}$ a díky její pozitivní semidefinitnosti $|\sigma_{a\cdot b}| \leq 1$. Pozorný čtenář si zajisté povšimnul, že jsme vlastně přesli od varianční matice k matici korelační.

Pokud jsou dvě náhodné veličiny ξ_a a ξ_b vzájemně **funkčně závislé** (značíme $\xi_a \simeq \xi_b$), což je za výše zmíněného omezení ekvivalentní požadavku $\xi_a = \pm\xi_b$, resp. $\sigma_{a\cdot b} = \pm 1$, potom $\xi_a \perp\!\!\!\perp \xi_c | \xi_b$, $\xi_b \perp\!\!\!\perp \xi_c | \xi_a$, $\xi_a \perp\!\!\!\perp \xi_c | \boldsymbol{\xi}_{bd}$ a $\xi_b \perp\!\!\!\perp \xi_c | \boldsymbol{\xi}_{ad}$, kde a, b, c a d jsou různé prvky N.

Uvedený pojem lze zobecnit i na soubor více veličin: řekneme že $(\xi_a)_{a\in A}$, $A \subseteq N$ jsou vzájemně **funkčně závislé** (značíme $\simeq_{(\boldsymbol{\xi}_A)}$), pokud $|\boldsymbol{\Sigma}_{A\cdot A}| = 0$ neboli dle (1.3) existuje vektor konstant $\boldsymbol{c}_A$ takový, že

$$\sum_{a\in A} c_a \xi_a = konst.$$

Pokud navíc pro žádnou neprázdnou podmnožinu $B \subseteq A$ neplatí $\simeq_{(\boldsymbol{\xi}_B)}$, potom jsou $(c_a)_{a\in A}$ nutně nenulové a tudíž

$$\forall a \in A, \exists \text{ funkce } f_a : \quad \xi_a = f_a(\boldsymbol{\xi}_{A\setminus a}), \tag{2.2}$$

navíc jelikož degenerovaná (konstantní) náhodná veličina je nezávislá s čímkoli, lze ze vztahu (2.2) snadno odvodit

$$\forall a \in A, E \subseteq N \setminus A, F \subseteq N \setminus AE : \quad \xi_a \perp\!\!\!\perp \boldsymbol{\xi}_E | \boldsymbol{\xi}_{AF\setminus a}.$$

Lemma 13. *Nechť I je g–reprezentovatelný nezávislostní model, a, b, c a d jsou různé prvky N.*

i) Jestliže $\{\langle ab|c\rangle, \langle ac|b\rangle\} \subseteq I$, potom $\{\langle ab|\emptyset\rangle, \langle ac|\emptyset\rangle\} \subseteq I$ nebo $\xi_b \simeq \xi_c$.

ii) Jestliže $\{\langle ab|cd\rangle, \langle ac|bd\rangle\} \subseteq I$, potom buďto $\{\langle ab|d\rangle, \langle ac|d\rangle\} \subseteq I$ nebo $\langle ad|bc\rangle \in I$ nebo $\xi_b \simeq \xi_c$.

Důkaz. Snadno nahlédneme, že má-li náhodný vektor $\boldsymbol{\xi}_{abcd}$ některou ze složek degenerovanou (konstantní), je pravá strana implikací i) a ii) splněna automaticky. Bez újmy na obecnosti se tedy opět omezíme na důkaz pro rozdělení s varianční maticí typu (2.1).

Požadované nezávislosti $\xi_a \perp\!\!\!\perp \xi_b | \xi_c$ a $\xi_a \perp\!\!\!\perp \xi_c | \xi_b$ neboli $\sigma_{a\cdot b} = \sigma_{a\cdot c}\sigma_{b\cdot c}$ a $\sigma_{a\cdot c} = \sigma_{a\cdot b}\sigma_{b\cdot c}$ dávají po dosazení $\sigma_{a\cdot b} = \sigma_{a\cdot b}\sigma_{b\cdot c}^2$, což implikuje $\sigma_{a\cdot b} = \sigma_{a\cdot c} = 0$ ($\xi_1 \perp\!\!\!\perp \xi_2$ a $\xi_1 \perp\!\!\!\perp \xi_3$) nebo $|\sigma_{b\cdot c}| = 1$ ($\xi_b \simeq \xi_c$).

Důkaz druhé části je myšlenkově analogický. Předpokládejme nejprve, že $\simeq_{(\boldsymbol{\xi}_{bcd})}$ neboli existují konstanty r, s, t a u takové, že

$$r = s\xi_b + t\xi_c + u\xi_d.$$

Pokud $u = 0$, potom $\xi_b \simeq \xi_c$. Naopak je-li $u \neq 0$, lze vyjádřit ξ_d jako funkci zbylých dvou složek $\xi_d = f_d(\xi_b, \xi_c)$, a proto $\xi_d \perp\!\!\!\perp \xi_a | \boldsymbol{\xi}_{bc}$.

Pokud naopak $|\boldsymbol{\Sigma}_{bcd\cdot bcd}| \neq 0$ a předpokládáme-li $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$ a $\xi_a \perp\!\!\!\perp \xi_c | \boldsymbol{\xi}_{bd}$, dostáváme po podmínění $\xi_d = x_d$ z část i) že $\xi_a \perp\!\!\!\perp \xi_b | \xi_d$ a $\xi_a \perp\!\!\!\perp \xi_c | \xi_d$ (možnost vzájemné funkční závislosti ξ_b a ξ_c po podmínění ξ_d je ve sporu s nenulovostí determinantu $\boldsymbol{\Sigma}_{bcd\cdot bcd}$). $\square$

M1: 6	**M2:** 12	**M3:** 6	**M4:** 6	**M5:** 24	**M6:** 6
M7: 12	**M8:** 12	**M9:** 3	**M10:** 12	**M11:** 24	**M12:** 3
M13: 6	**M14:** 24	**M15:** 24	**M16:** 12	**M17:** 24	**M18:** 24
M19: 12	**M20:** 6	**M21:** 12	**M22:** 6	**M23:** 12	**M24:** 24
M25: 24	**M26:** 12	**M27:** 24	**M28:** 24	**M29:** 12	**M30:** 6
M31: 24	**M32:** 6	**M33:** 12	**M34:** 12	**M35:** 3	**M36:** 3
M37: 12	**M38:** 12	**M39:** 12	**M40:** 12	**M41:** 12	**M42:** 4
M43: 4	**M44:** 24	**M45:** 12	**M46:** 12	**M47:** 4	**M48:** 12
M49: 12	**M50:** 3	**M51:** 6	**M52:** 1	**M53:** 1	
M59: 4	**M60:** 12	**M61:** 12	**M62:** 12	**M63:** 12	**M64:** 12
M65: 12	**M66:** 24	**M67:** 1	**M68:** 12	**M69:** 6	**M70:** 6
M71: 3	**M72:** 24	**M73:** 6	**M74:** 6	**M75:** 12	**M76:** 6
M77: 12	**M78:** 6	**M79:** 12	**M80:** 3	**M81:** 4	**M82:** 12
M83: 1	**M84:** 4	**M85:** 12			

Tabulka 2.1: Počet izomorfních modelů nad $N = \{1, 2, 3, 4\}$ pro každý z g–reprezentovatelných typů na obr. 2.3 a 2.4.

Existuje **178** typů nezávislostních modelů, jejichž 3–minory jsou jednoho ze sedmi typů g–reprezentovatelných pro $|N| = 3$ (viz obr. 2.1). Lemma 13 nám umožňuje ukázat, že **90** z nich nemůže mít g–reprezentaci. Zbýva tedy rozhodnout o g–reprezentovatelnosti **88** typů (M1–M88, obr. 2.3 a obr. 2.4).

Počítačovým prohledáváním prostoru celočíselných symetrických pozitivně semidefinitních matric bylo nalezeno **79** g–reprezentací a **1** další byla nalezena později bez použití počítače (M1–M53, obr. 2.3, a M59–M85, obr. 2.4). Varianční matice reprezentací na obr. 2.3 a 2.4 odpovídají řazení vrcholů v diagramech po směru hodinových ručiček s počátkem v levém horním rohu. Počet izomorfních modelů pro každý typ lze nalézt v tabulce 2.1.

Zbývajících 8 typů (M54–M58 a M86–M88, obr. 2.4) není g–reprezentovatelných, jak dokážeme v následujícím lemmatu.

Lemma 14. *Nechť a, b, c a d jsou různé prvky N. Nezávislostní modely*

i) $I_1 = \{\langle cd|a\rangle, \langle ad|b\rangle, \langle bd|c\rangle\}$

ii) $I_2 = \{\langle ab|c\rangle, \langle cd|b\rangle, \langle bd|a\rangle, \langle ac|d\rangle\}$

iii) $I_3 = \{\langle cd|\emptyset\rangle, \langle ab|c\rangle, \langle bd|a\rangle, \langle ac|bd\rangle\}$

iv) $I_4 = \{\langle cd|\emptyset\rangle, \langle ab|c\rangle, \langle ab|d\rangle, \langle cd|ab\rangle\}$

v) $I_5 = \{\langle ab|\emptyset\rangle, \langle cd|\emptyset\rangle, \langle ac|bd\rangle, \langle bd|ac\rangle\}$

vi) $I_6 = \{\langle ab|d\rangle, \langle bc|d\rangle, \langle ac|d\rangle, \langle ab|cd\rangle, \langle bc|ad\rangle, \langle cd|ab\rangle, \langle ac|bd\rangle\}$

vii) $I_7 = \{\langle ab|d\rangle, \langle bc|d\rangle, \langle ac|d\rangle, \langle ab|cd\rangle, \langle bc|ad\rangle, \langle cd|ab\rangle, \langle ad|bc\rangle, \langle ac|bd\rangle, \langle bd|ac\rangle\}$

viii) $I_8 = \{\langle ab|d\rangle, \langle bc|d\rangle, \langle ac|d\rangle, \langle ab|c\rangle, \langle bd|c\rangle, \langle ad|c\rangle, \langle ab|cd\rangle, \langle bc|ad\rangle, \langle cd|ab\rangle, \langle ad|bc\rangle, \langle ac|bd\rangle, \langle bd|ac\rangle\}$

nejsou g–reprezentovatelné.

Důkaz. Předpokládejme, že $\boldsymbol{\xi}$ je g–reprezentací některého z výše uvedených modelů, $\boldsymbol{\Sigma}$ je příslušná varianční matrice typu (2.1). Přičemž pokud zkoumaný model I neobsahuje triplet $\langle xy|\emptyset\rangle$, můžeme automaticky předpokládát $\sigma_{x\cdot y} \neq 0$, a pokud pro dané x, y neobsahuje všechny triplety typu $\langle xu|y\rangle$, $\langle yu|x\rangle$, $\langle xu|yt\rangle$ a $\langle yu|xt\rangle$, lze předpokládat $\xi_x \not\simeq \xi_y$.

V částech i)-viii) dospějeme ke sporu pro modely po řadě $I_1, \ldots, I_8$ a to způsobem, že si vypíšeme vztahy odpovídající tripletům v modelu a z nich vyvodíme nezávislostní vztah v modelu neobsažený, případně funkční závislost, kterou model nedovoluje.

i) Požadované nezávislosti dle lemmatu 5 implikují: $\sigma_{c\cdot d} = \sigma_{a\cdot c}\sigma_{a\cdot d}$, $\sigma_{a\cdot d} = \sigma_{a\cdot b}\sigma_{b\cdot d}$ a $\sigma_{b\cdot d} = \sigma_{b\cdot c}\sigma_{c\cdot d}$. Dosadíme-li do sebe tyto rovnice, dostaneme

$$\sigma_{c\cdot d} = \sigma_{a\cdot c}\sigma_{a\cdot b}\sigma_{b\cdot c}\sigma_{c\cdot d},$$

a tudíž $\sigma_{c\cdot d} = \sigma_{a\cdot d} = \sigma_{b\cdot d} = 0$ ($\xi_d \perp\!\!\!\perp \boldsymbol{\xi}_{abc}$) nebo $\xi_a \simeq \xi_b \simeq \xi_c$, což je spor.

ii) Analogicky lze z $\sigma_{a\cdot b} = \sigma_{a\cdot c}\sigma_{b\cdot c}$, $\sigma_{a\cdot c} = \sigma_{a\cdot d}\sigma_{c\cdot d}$, $\sigma_{c\cdot d} = \sigma_{b\cdot c}\sigma_{b\cdot d}$ a $\sigma_{b\cdot d} = \sigma_{a\cdot b}\sigma_{a\cdot d}$ odvodit

$$\sigma_{a\cdot b} = (\sigma_{b\cdot c}\sigma_{a\cdot d})^2\,\sigma_{a\cdot b}$$

a odtud $\sigma_{a\cdot b} = 0$ ($\xi_a \perp\!\!\!\perp \xi_b$) nebo $\xi_b \simeq \xi_c$ a $\xi_a \simeq \xi_d$.

iii) Pokud vyloučíme $\xi_b \simeq \xi_d$ a dosadíme-li $\sigma_{c\cdot d} = 0$, $\sigma_{a\cdot b} = \sigma_{a\cdot c}\sigma_{b\cdot c}$ a $\sigma_{b\cdot d} = \sigma_{a\cdot b}\sigma_{a\cdot d}$ do

$$\sigma_{a\cdot c} + \sigma_{b\cdot d}\sigma_{c\cdot d}\sigma_{a\cdot b} + \sigma_{b\cdot d}\sigma_{b\cdot c}\sigma_{a\cdot d} - \sigma_{a\cdot d}\sigma_{c\cdot d} - \sigma_{a\cdot b}\sigma_{b\cdot c} - \sigma_{a\cdot c}\sigma_{b\cdot d}^2 = 0$$

dostaneme po úpravě

$$\sigma_{a\cdot c}((1 - \sigma_{b\cdot c}^2) + \sigma_{a\cdot d}^2\sigma_{b\cdot c}^2(1 - \sigma_{a\cdot c}^2)) = 0.$$

Odtud plyne $\sigma_{a\cdot c} = 0$ ($\xi_a \perp\!\!\!\perp \xi_c$) nebo $\sigma_{b\cdot c}^2 = 1$ ($\xi_b \simeq \xi_c$).

iv) Analogicky, dosazením $\sigma_{c\cdot d} = 0$, $\sigma_{a\cdot b} = \sigma_{a\cdot c}\sigma_{b\cdot c}$ a $\sigma_{b\cdot d} = \sigma_{a\cdot b}/\sigma_{a\cdot d} = \sigma_{a\cdot c}\sigma_{b\cdot c}/\sigma_{a\cdot d}$ do rovnice příslušné tripletu $\langle cd|ab\rangle$

$$\sigma_{a\cdot b}\sigma_{b\cdot c}\sigma_{a\cdot d} + \sigma_{a\cdot b}\sigma_{b\cdot d}\sigma_{a\cdot c} - \sigma_{a\cdot c}\sigma_{a\cdot d} - \sigma_{b\cdot c}\sigma_{b\cdot d} = 0$$

dostaneme

$$\frac{\sigma_{a\cdot c}}{\sigma_{a\cdot d}}(\sigma^2_{a\cdot d}(1-\sigma^2_{b\cdot c}) + \sigma^2_{b\cdot c}(1-\sigma^2_{a\cdot c})) = 0$$

a odtud plyne $\sigma_{a\cdot c} = 0$ ($\xi_a \perp\!\!\!\perp \xi_c$) nebo $\sigma_{a\cdot d} = 0$ ($\xi_a \perp\!\!\!\perp \xi_d$) nebo $|\sigma_{b\cdot c}| = 1$ ($\xi_b \simeq \xi_c$).

v) Pokud nenastanou funkcionální závislosti $\xi_a \simeq \xi_c$ nebo $\xi_b \simeq \xi_d$, dostaneme dosazením $\sigma_{a\cdot b} = \sigma_{c\cdot d} = 0$ do rovnicí přísluČným tripletům $\langle ac|bd\rangle$ a $\langle bd|ac\rangle$

$$\begin{aligned} \sigma_{a\cdot c} - \sigma_{a\cdot c}\sigma^2_{b\cdot d} + \sigma_{a\cdot d}\sigma_{b\cdot c}\sigma_{b\cdot d} &= 0, \\ \sigma_{b\cdot d} - \sigma_{b\cdot d}\sigma^2_{a\cdot c} + \sigma_{a\cdot d}\sigma_{b\cdot c}\sigma_{a\cdot c} &= 0. \end{aligned}$$

Odečtením první rovnice vynásobené $\sigma_{b\cdot d}$ od druhé rovnice vynásobené $\sigma_{a\cdot c}$ dostaneme $\sigma_{a\cdot c}\sigma_{b\cdot d}(\sigma_{a\cdot c} - \sigma_{b\cdot d})(\sigma_{a\cdot c} + \sigma_{b\cdot d}) = 0$. Odtud buďto $\sigma_{a\cdot c} = 0$ nebo $\sigma_{b\cdot d} = 0$ nebo $\sigma_{a\cdot c} = \pm\sigma_{b\cdot d}$. Pro $\sigma_{a\cdot c} = \pm\sigma_{b\cdot d}$ lze dále z pozitivní semidefinitnosti matic $\mathbf{\Sigma}_{abc\cdot abc}$ a $\mathbf{\Sigma}_{acd\cdot acd}$ ukázat, že $1 - \sigma^2_{a\cdot c} \geq \max(\sigma^2_{a\cdot d}, \sigma^2_{b\cdot c}) \geq |\sigma_{a\cdot d}\sigma_{b\cdot c}|$, přičemž v druhém případě nastává rovnost jen pro $|\sigma_{a\cdot d}| = |\sigma_{b\cdot c}|$. Toto dosadíme zpět do rovnic příslušných $\langle ac|bd\rangle$ a $\langle bd|ac\rangle$ a dostáváme $\sigma_{a\cdot d} = \mp\sigma_{b\cdot c}$, což vede za předpokladu $\xi_c \not\simeq \xi_d$ na $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$.

vi)+vii) Dosadíme-li dle předpokladů $\sigma_{a\cdot b} = \sigma_{b\cdot d}\sigma_{a\cdot d}$, $\sigma_{b\cdot c} = \sigma_{b\cdot d}\sigma_{c\cdot d}$ a $\sigma_{a\cdot c} = \sigma_{a\cdot d}\sigma_{c\cdot d}$ do rovnice příslušné tripletu $\langle cd|ab\rangle$ (můžeme předpokládat $\xi_a \not\simeq \xi_b$), dostaneme

$$\sigma_{c\cdot d}(1-\sigma^2_{b\cdot d})(1-\sigma^2_{a\cdot d}) = 0,$$

odtud $\xi_c \perp\!\!\!\perp \xi_d$ nebo $\xi_a \simeq \xi_d$ nebo $\xi_b \simeq \xi_d$, což je spor.

viii) Jestliže $\xi_c \simeq \xi_d$, potom lze z nezávislosti $\xi_c \perp\!\!\!\perp \xi_d | \boldsymbol{\xi}_{ab}$ odvodit $\simeq_{(\boldsymbol{\xi}_{abc})}$. Uvážíme-li dále $\xi_a \perp\!\!\!\perp \xi_b | \xi_c$, dostáváme $\xi_a \simeq \xi_c$ nebo $\xi_b \simeq \xi_c$, což je spor. V opačném případě ($\xi_c \not\simeq \xi_d$) dostaneme spor způsobem popsaným v části „vi)+vii)“.

$\square$

Shrňme tedy dosažené poznatky do závěrečné věty:

Věta 2. *Nad množinou $N = \{1,2,3,4\}$ existuje 877 g–reprezentovatelných nezávislostních modelů 80 typů. Pozitivně reprezentovatelných je 629 z nich (53 typů) a jsou to právě ty splňující podmínku (1.10).*

2.3 Inferenční pravidla

Někteří autoři dávají přednost charakterizaci reprezentovatelných modelů za pomoci inferenčních pravidel. Přestože je zjevné, že tato práce se neubírá tímto směrem, uveďme pro srovnání seznam inferenčních pravidel vyplynuvších z dosažené charakterizace Gaussovsky reprezentovatelných modelů nad čtyřprvkovou množinou.

Lemma 15. *Nechť N je konečná množina a I je g–reprezentovatelný nezávislostní model nad N. Potom pro a,b,c,d různé prvky N a množiny $E \subseteq N \setminus abc$ a $F \subseteq N \setminus abcd$ platí:*

i) $\{\langle ab|E\rangle, \langle ac|bE\rangle\} \subseteq I \iff \{\langle ac|E\rangle, \langle ab|cE\rangle\} \subseteq I$

ii) $\{\langle ab|E\rangle, \langle ac|E\rangle\} \subseteq I \implies \langle ab|cE\rangle \in I$

iii) $\{\langle ab|E\rangle, \langle ab|cE\rangle\} \subseteq I \implies \big(\langle ac|E\rangle \in I\big) \vee \big(\langle bc|E\rangle \in I\big)$

iv) $\{\langle ab|cF\rangle, \langle ac|bF\rangle\} \subseteq I \implies \big(\langle ab|F\rangle \in I\big) \vee \big(\{\langle bd|cF\rangle, \langle cd|bF\rangle\} \subseteq I\big)$

v) $\{\langle ab|cdF\rangle, \langle ac|bdF\rangle\} \subseteq I \implies \big(\langle ab|dF\rangle \in I\big) \vee \big(\langle ad|bcF\rangle \in I\big) \vee \big(\{\langle bd|cF\rangle, \langle cd|bF\rangle\} \subseteq I\big)$

vi) $\{\langle ab|F\rangle, \langle cd|F\rangle, \langle ac|bdF\rangle, \langle bd|acF\rangle\} \subseteq I \implies \big(\langle ac|F\rangle \in I\big) \vee \big(\{\langle ab|cdF\rangle, \langle ad|bcF\rangle\} \subseteq I\big)$

vii) $\{\langle ab|F\rangle, \langle cd|aF\rangle, \langle cd|bF\rangle, \langle ab|cdF\rangle\} \subseteq I \implies \langle cd|F\rangle \in I$

viii) $\{\langle ab|F\rangle, \langle bd|cF\rangle, \langle cd|aF\rangle, \langle ac|bdF\rangle\} \subseteq I \implies \langle bd|F\rangle \in I$

ix) $\{\langle ab|cF\rangle, \langle ac|dF\rangle, \langle ad|bF\rangle\} \subseteq I \implies \langle ab|dF\rangle \in I$

x) $\{\langle ab|cF\rangle, \langle bc|dF\rangle, \langle cd|aF\rangle, \langle ad|bF\rangle\} \subseteq I \implies \langle ab|dF\rangle \in I$

xi) $\{\langle ab|dF\rangle, \langle ac|dF\rangle, \langle bc|dF\rangle, \langle cd|abF\rangle\} \subseteq I \implies \big(\langle cd|bF\rangle \in I\big) \vee \big(\langle cd|aF\rangle \in I\big)$

Důkaz. První část je semigrafoidová vlastnostnost z lemmatu 8, druhá a třetí jsou důsledky důkazu lemmatu 10, další dvě důsledkem lemmatu 13.

Dalších šest dostaneme rozborem g–reprezentovatelných nadmodelů modelů z lemmatu 14 (M54–58 a M86–88) za pomoci obr. 2.3 a obr. 2.4. Ke každému g–nereprezentovatelnému modelu I nalezneme všechny g–reprezentovatelné nadmodely $J_i \supset I$ a jako nové inferenční pravidlo přidáme „Pokud je I podmnožinou g–reprezentovatelného modelu, potom daný model obsahuje i (libovolně zvolený) triplet z průniku $\cap_i J_i$“.

Dodejme, že takto dostaneme jedno pravidlo pro každý z modelů M54–58, zatímco modely M86–88 jsou zaráz vyloučeny jediným pravidlem xi).

□

Pro $|N| \leq 4$ dávají inferenční pravidla z lemmatu 15 charakterizaci úplnou, tzn. nezávislostní model je g–reprezentovatelný právě tehdy, když splňuje pravidla i)-xi).

Regulárně Gaussovsky reprezentovatelné modely nad čtyřprvkovou množinou již byly charakterizovány pomocí inferenčních pravidel v [Lně05]. Pro úplnost uvedeme i tento výsledek (bez dalšího zdůvodnění).

Lemma 16. *Nechť N je konečná množina a I je g^+–reprezentovatelný nezávislostní model nad N. Potom pro a,b,c,d různé prvky N a množiny $E \subseteq N \setminus abc$ a $F \subseteq N \setminus abcd$ platí:*

i) $\{\langle ab|E\rangle, \langle ac|bE\rangle\} \subseteq I \iff \{\langle ac|E\rangle, \langle ab|cE\rangle\} \subseteq I$

ii) $\{\langle ab|cE\rangle, \langle ac|bE\rangle\} \subseteq I \iff \{\langle ab|E\rangle, \langle ac|E\rangle\} \subseteq I$

iii) $\{\langle ab|E\rangle, \langle ab|cE\rangle\} \subseteq I \implies \big(\langle ac|E\rangle \in I\big) \vee \big(\langle bc|E\rangle \in I\big)$

iv) $\{\langle ab|F\rangle, \langle cd|F\rangle, \langle ac|bdF\rangle, \langle bd|acF\rangle\} \subseteq I \implies \langle ac|F\rangle \in I$

v) $\{\langle ab|F\rangle, \langle cd|aF\rangle, \langle cd|bF\rangle, \langle ab|cdF\rangle\} \subseteq I \implies \langle cd|F\rangle \in I$

vi) $\{\langle ab|F\rangle, \langle bd|cF\rangle, \langle cd|aF\rangle, \langle ac|bdF\rangle\} \subseteq I \implies \langle ac|F\rangle \in I$

vii) $\{\langle ab|cF\rangle, \langle ac|dF\rangle, \langle ad|bF\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

viii) $\{\langle ab|cF\rangle, \langle bc|dF\rangle, \langle cd|aF\rangle, \langle ad|bF\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

Snad by bylo vhodné poznamenat, že vlastnosti $v) - viii)$ v lemmatu 16 jsou díky vlastnostem $i) - iii)$ z lemmatu 16 vlastně ekvivalentní vlastnostem $vii) - x)$ z lemmatu 15 a že pro $|N| \leq 4$ dává lemma 16 úplnou charakterizaci g^+–reprezentovatelných modelů.

2.4 Neexistence konečné charakterizace

V tomto oddílu dokážeme, že Gaussovsky reprezentovatelné nezávislostní modely nelze charakterizovat konečným seznamem zakázaných minorů. Postupovat budeme podle návodu z lemmatu 9 v oddílu 1.4.

Lemma 17. *Pro libovolné $n \geq 4$ je nezávislostní model*

$$J_n = \{\langle 12|3\rangle, \langle 13|4\rangle, \ldots, \langle 1(n-1)|n\rangle, \langle 1n|2\rangle\} \tag{2.3}$$

zakázaným minorem pro třídu Gaussovsky reprezentovatelných modelů.

Důkaz. Předpokládejme, že by existoval g–reprezentovatelný nezávislostní model I nad množinou $\{1, \ldots, n\} \cup EF$ takový, že $I\,[_E^F = J_n$. Potom by dle lemmatu 11 byl g–reprezentovatelný i model J_n. Označme tedy $\mathbf{\Sigma}$ varianční matici typu $\forall a: \ \sigma_{a \cdot a} = 1$ (viz lemma 12) příslušné g–reprezentace $\boldsymbol{\xi}$.

Požadované nezávislosti

$$\xi_1 \perp\!\!\!\perp \xi_2|\xi_3, \ \xi_1 \perp\!\!\!\perp \xi_3|\xi_4, \ \ldots, \ \xi_1 \perp\!\!\!\perp \xi_{n-1}|\xi_n, \ \xi_1 \perp\!\!\!\perp \xi_n|\xi_2$$

odpovídají dle lemmatu 5 vztahům

$$\sigma_{1\cdot 2} = \sigma_{1\cdot 3}\sigma_{2\cdot 3}, \ \sigma_{1\cdot 3} = \sigma_{1\cdot 4}\sigma_{3\cdot 4}, \ \ldots, \ \sigma_{1\cdot(n-1)} = \sigma_{1\cdot n}\sigma_{(n-1)\cdot n}, \ \sigma_{1\cdot n} = \sigma_{1\cdot 2}\sigma_{2\cdot n}$$

a opakovaným dosazením dostáváme

$$\sigma_{1\cdot 2} = \sigma_{1\cdot 2} \cdot \big(\sigma_{2\cdot 3}\sigma_{3\cdot 4} \ldots \sigma_{(n-1)\cdot n}\sigma_{2\cdot n}\big).$$

Tudíž, buďto $\sigma_{1\cdot 2} = 0$ (tj. $\xi_1 \perp\!\!\!\perp \xi_2$) nebo $|\sigma_{2\cdot 3}| = 1$ (tj. $\xi_1 \perp\!\!\!\perp \xi_3|\xi_2$), což je spor s tím, že $\boldsymbol{\xi}$ je g–reprezentace modelu J_n. $\square$

Lemma 18. *Pro libovolné $n \geq 3$ jsou nezávislostní modely*

$$\begin{aligned} I_n &= \{\langle 12|3\rangle, \langle 13|4\rangle, \ldots, \langle 1(n-1)|n\rangle\}, \\ I_n^* &= \{\langle 12|\emptyset\rangle\}, \quad I_n^{**} = \emptyset \end{aligned}$$

regulárně Gaussovsky reprezentovatelné nad množinou $N = \{1, \ldots, n\}$.

Důkaz. Nechť $\boldsymbol{\xi} = (\xi_a)_{a \in \{1,\ldots,n\}}$ je Gaussovsky rozdělený náhodný vektor s varianční maticí $\mathbf{\Sigma} = (\sigma_{a\cdot b})_{a,b \in \{1,\ldots,n\}}$ takovou, že

$$\begin{aligned} \forall a: \ \sigma_{a\cdot a} &= 1, \\ \forall a > 1: \ \sigma_{1\cdot a} &= \sigma_{a\cdot 1} = \epsilon^{n-a+1}, \\ \forall a, b > 1, a \neq b: \ \sigma_{a\cdot b} &= \sigma_{b\cdot a} = \epsilon, \end{aligned}$$

kde $\epsilon > 0$ je libovolně zvolené dostatečně malé číslo. Zjevně, $\mathbf{\Sigma} > 0$, neboť $\forall A \subseteq \{1,\ldots,n\} : \; |\mathbf{\Sigma}_{A\cdot A}| = 1 + \epsilon(\cdots) > 0$, kde „$\ldots$" zastupuje polynom v proměnné ϵ.

Nyní použijeme obdobných argumentů v kombinaci s lemmatem 5 pro důkaz, že $\mathcal{I}(\boldsymbol{\xi}) = I_n$:

i) Pro $a, b \in \{1, 2, \ldots, n\}$ je $\sigma_{a\cdot b} \neq 0$, tudíž zjevně neplatí $\xi_a \perp\!\!\!\perp \xi_b$.

ii) Pro $a, b \in \{2, 3, \ldots, n\}$ a množinu C, $\emptyset \neq C \subseteq \{1, \ldots, n\} \setminus ab$, determinant

$$|\Sigma_{aC\cdot bC}| = \left| \begin{pmatrix} \epsilon & \epsilon & \epsilon & \epsilon & \cdots \\ \epsilon & 1 & \epsilon & \epsilon & \cdots \\ \epsilon & \epsilon & 1 & \epsilon & \cdots \\ \epsilon & \epsilon & \epsilon & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \right| = \epsilon + \epsilon^2(\ldots) \neq 0$$

a tudíž tvrzení $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_C$ nemůže býti platným (pro ϵ dostatečně malé).

iii) Pro $a \in \{2, 3, \ldots n\}$ a neprázdnou množinu $C = \{c_1, c_2, \ldots, c_k\} \subseteq \{2, 3, \ldots, a-1\}$, determinant

$$|\Sigma_{aC\cdot 1C}| = \left| \begin{pmatrix} \epsilon^{n-a+1} & \epsilon & \epsilon & \epsilon & \cdots \\ \epsilon^{n-c_1+1} & 1 & \epsilon & \epsilon & \cdots \\ \epsilon^{n-c_2+1} & \epsilon & 1 & \epsilon & \cdots \\ \epsilon^{n-c_3+1} & \epsilon & \epsilon & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \right| = \epsilon^{n-a+1} + \epsilon^{n-a+2}(\ldots) \neq 0.$$

Proto ani tvrzení $\xi_a \perp\!\!\!\perp \xi_1 | \boldsymbol{\xi}_C$ nemůže býti platným.

iv) Analogicky, pro $a \in \{2, 3, \ldots n\}$ a $C = \{c_1, \ldots, c_k\} \subseteq \{2, 3, \ldots, n\} \setminus a$ takovou, že $m := \max_{c \in C} c > a + 1$, příslušný determinant má tvar

$$|\Sigma_{aC\cdot 1C}| = \left| \begin{pmatrix} \epsilon^{n-a+1} & \epsilon & \epsilon & \epsilon & \cdots \\ \epsilon^{n-c_1+1} & 1 & \epsilon & \epsilon & \cdots \\ \epsilon^{n-c_2+1} & \epsilon & 1 & \epsilon & \cdots \\ \epsilon^{n-c_3+1} & \epsilon & \epsilon & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \right| = \pm\epsilon^{n-m+2} + \epsilon^{n-m+3}(\ldots) \neq 0,$$

Proto ani v tomto případě tvrzení $\xi_a \perp\!\!\!\perp \xi_1 | \boldsymbol{\xi}_C$ nemůže býti platným.

v) Zbývá vyšetřit případ $a \in \{2, 3, \ldots n-1\}$ a $C = \{(a+1), c_2, \ldots, c_k\}$, $\emptyset \neq C \subseteq (a+1) \cup \{2, 3, \ldots, a-1\}$. Pro C dvou či víceprvkovou má determinant $|\Sigma_{aC\cdot 1C}|$ tvar

$$\left| \begin{pmatrix} \epsilon^{n-a+1} & \epsilon & \epsilon & \epsilon & \cdots \\ \epsilon^{n-a} & 1 & \epsilon & \epsilon & \cdots \\ \epsilon^{n-c_2+1} & \epsilon & 1 & \epsilon & \cdots \\ \epsilon^{n-c_3+1} & \epsilon & \epsilon & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \right| = \epsilon^{n-a+1} - \epsilon^{n-a+1} \pm \epsilon^{n-a+2} + \epsilon^{n-a+3}(\ldots).$$

Tudíž $|\Sigma_{1C\cdot aC}| \neq 0$ a tvrzení $\xi_1 \perp\!\!\!\perp \xi_a | \boldsymbol{\xi}_C$ neplatí.

Avšak pokud je naopak $C = \{a+1\}$ jednoprvková, potom

$$|\Sigma_{a(a+1)\cdot 1(a+1)}| = \left| \begin{pmatrix} \epsilon^{n-a+1} & \epsilon \\ \epsilon^{n-a} & 1 \end{pmatrix} \right| = 1\epsilon^{n-a+1} - \epsilon\epsilon^{n-a} = 0.$$

Skutečně tedy platí $\xi_1 \perp\!\!\!\perp \xi_a | \xi_{a+1}$.

Zcela analogickými argumenty lze ukázat, že modely I_n^* a I_n^{**} jsou g^+–reprezentovány Gaussovskými náhodnými vektory s variančními maticemi

$$\boldsymbol{\Sigma}^* = \begin{pmatrix} 1 & 0 & \epsilon & \epsilon & \cdots \\ 0 & 1 & \epsilon & \epsilon & \cdots \\ \epsilon & \epsilon & 1 & \epsilon & \cdots \\ \epsilon & \epsilon & \epsilon & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad \boldsymbol{\Sigma}^{**} = \begin{pmatrix} 1 & \epsilon & \epsilon & \epsilon & \cdots \\ \epsilon & 1 & \epsilon & \epsilon & \cdots \\ \epsilon & \epsilon & 1 & \epsilon & \cdots \\ \epsilon & \epsilon & \epsilon & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Dopracování důkazu je ponecháno pozornému čtenáři. □

S využitím lemmatu 9 dostáváme jako důsledek lemmat 17 a 18 následující závěrečné tvrzení:

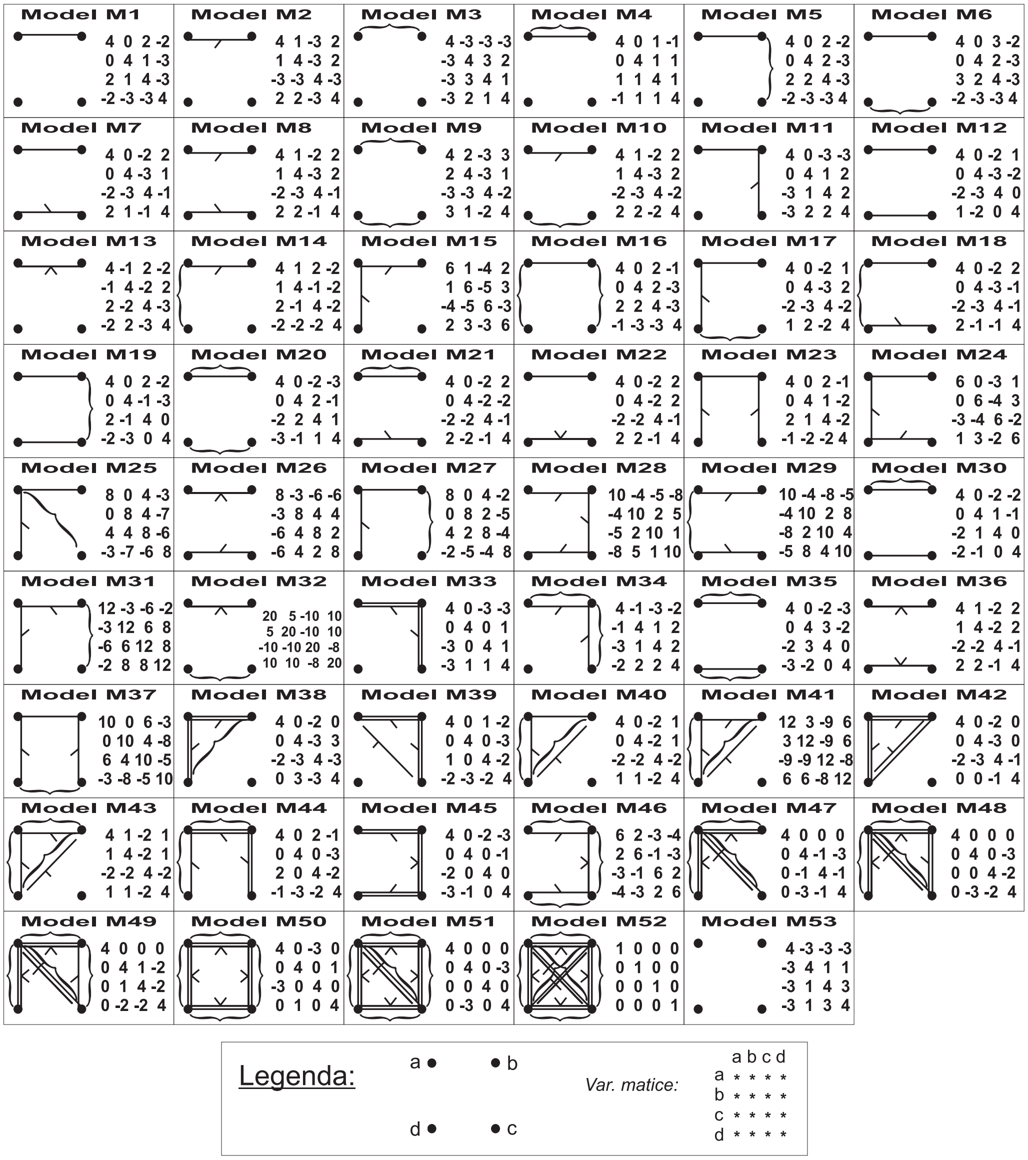
Věta 3. *Třídy (obecně) Gaussovsky a regulárně Gaussovsky reprezentovatelných nezávislostních modelů nejsou konečně charakterizovatelné.*

2.5 Závěr kapitoly

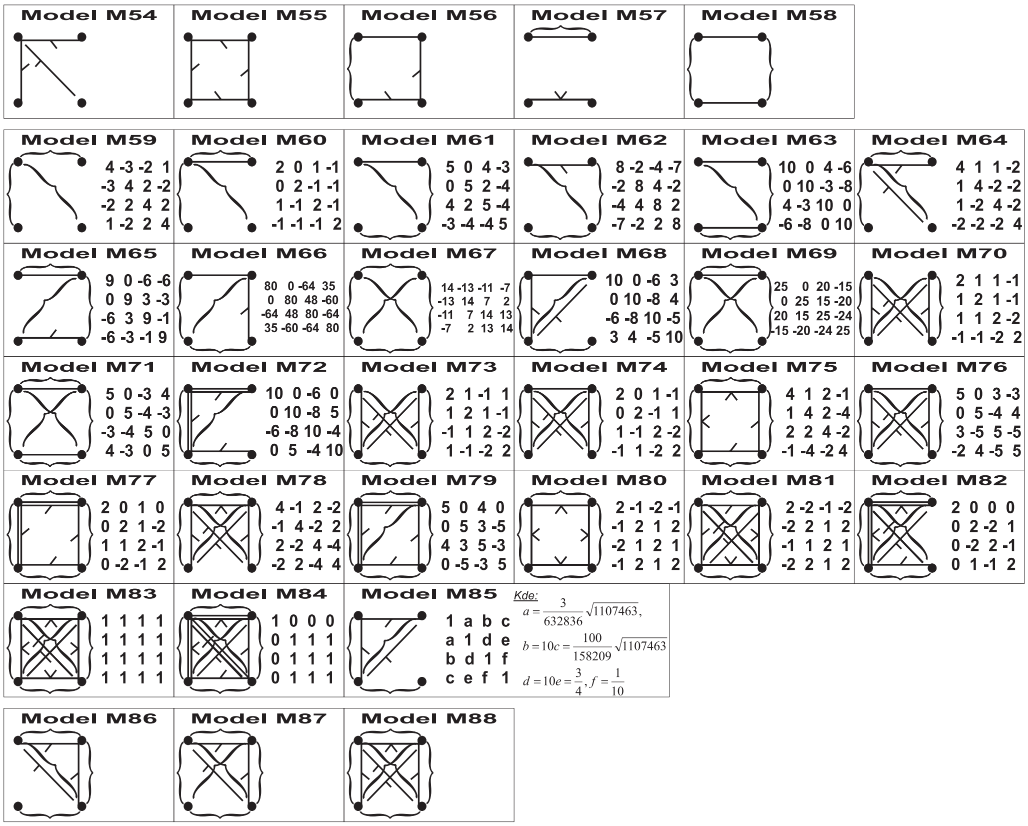
V této kapitole jsme dokázali, že neexistuje konečná charakterizace Gaussovsky reprezentovatelných nezávislostních modelů pomocí zakázaných minorů a nalezli všechny Gaussovsky reprezentovatelné modely nad $N = \{1, 2, 3, 4\}$. Vyvstává přirozená otázka, zda by nebylo možné pokračovat dále a charakterizovat Gaussovsky reprezentovatelné modely pro $|N| = 5$. Lze ukázat, že existuje 366177 typů nezávislostních modelů s g^+–reprezentovatelnými 4–minory. Bohužel, u většiny těchto typů nejsme zatím schopni najít g^+–reprezentaci, ani dokázat, že neexistuje.

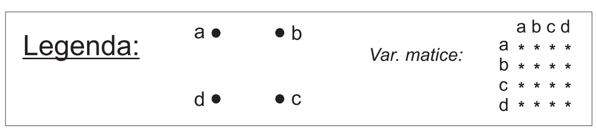
Zmiňme ještě několik dalších otevřených problémů. Všechny matice rozptylu na obr. 2.3 a obr. 2.4 vyjma matice příslušné modelu M85 jsou celočíselné. Otevřeným problémem zůstává, zda lze lze celočíselnou (resp. racionální) reprezentaci nalézt pro každý g^+–reprezentovatelný (resp. g–reprezentovatelný) model. Tento problém úzce souvisí s hypotézou vyslovenou F. Matúšem na konci [Mat99], že pro každý d–reprezentovatelný model existuje jeho racionální reprezentace.

Poslední bod se již týká obsahu příštích dvou kapitol. Je snadné zkonstruovat model, jenž je d–reprezentovatelný nebo b–reprezentovatelný, ale není g–reprezentovatelný. M. Studený však formuloval otázku, zda existuje g–reprezentovatelný model, jenž by nebyl d–reprezentovatelný. Odpověď na tuto otázku dává model M71 na obr. 2.4, jenž protiřečí pravidlům odvozeným z „nezávislostní implikace“ a může být tedy reprezentován jen rozdělením, které nemá konečnou multiinformaci. Nadále však zůstává otevřeným problémem, zda existuje g^+–reprezentovatelný model, který by nebyl b^+–reprezentovatelný, resp. d^+–reprezentovatelný. Tento problém bude podrobněji studován na konci následující kapitoly.



Obrázek 2.3: g^{+}–reprezentovatelné typy s varianční maticí reprezentace.



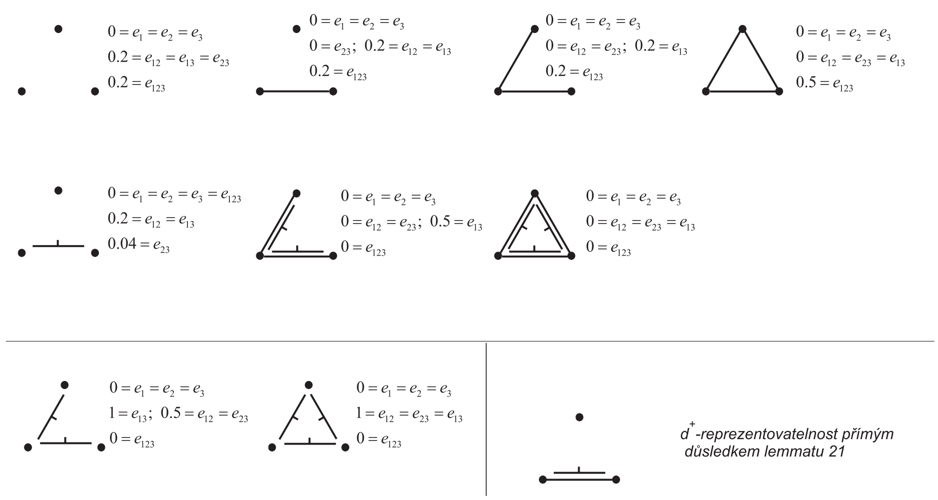


Obrázek 2.4: Po řadě od shora g^+–nereprezentovatelné typy s g^+–reprezentovatelnými minory, zbylé g–reprezentovatelné typy, g–nereprezentovatelné typy s g–reprezentovatelnými minory splňující vlastnosti z lemmatu 13.

Kapitola 3

Diskrétně a binárně reprezentovatelné nezávislostní modely

V této kapitole nalezneme všechny diskrétně (resp. binárně) reprezentovatelné nezávislostní modely pro $|N| \leq 3$. Dále uvedeme známý výsledek pro diskrétní reprezentovatelnost v případě $|N| = 4$ doplněný částečnými výsledky pro pozitivní diskrétní a binární reprezentovatelnost.



Obrázek 3.1: Po řadě od shora a zleva: b^+/d^+–reprezentovatelné, b/d–reprezentovatelné a zbylý d^+–reprezentovatelný typ pro $|N| = 3$.

Pro $|N| \leq 2$ jsou všechny možné nezávislostní modely zjevně (pozitivně binárně) reprezentovatelné. S rozborem tedy začneme pro $|N| = 3$. Modely, jež vyhovují podmínce (1.9) z lemmatu 8 jsou **10** typů, které jsou znázorněny na obr. 3.1 na str. 41.

Podmínce pro kladně reprezentovatelné modely (1.10) vyhovuje **8** typů, z nichž všechny jsou d^+–reprezentovatelné, ovšem jeden z nich není b^+–reprezentovatelný (viz níže, lemma 19). Zbylé **2** typy jsou (obecně) binárně a tedy i diskrétně reprezentovatelné.

Lemma 19. *Nezávislostní model* $\{\langle ab|c\rangle, \langle ab|\emptyset\rangle\}$ *není* b*–reprezentovatelný.*

Důkaz. Jedná se o známý fakt, viz např. [Spo94]. V námi zvolené parametrizaci se provede důkaz následovně: rovnice příslušné nezávislostním vztahům $\xi_a \perp\!\!\!\perp \xi_b$ a $\xi_a \perp\!\!\!\perp \xi_b | \xi_c$ jsou

$$
\begin{aligned}
e_{ab} &= e_a e_b \\
e_{abc} &= e_a e_{bc} + e_b e_{ac} - e_{ab} e_c \\
e_{ab} + e_{abc} e_c &= e_{ac} e_{bc} + e_a e_b
\end{aligned}
$$

Pokud dosadíme první dvě rovnice do třetí, dostaneme po úpravě

$$(e_{ac} - e_a e_c) \cdot (e_{bc} - e_b e_c) = 0,$$

a proto nutně $\xi_a \perp\!\!\!\perp \xi_c$ nebo $\xi_b \perp\!\!\!\perp \xi_c$. □

3.1 Pozitivní binární reprezentovatelnost

Poněkud netradičně začneme u rozboru reprezentovatelnosti nad čtyřmi náhodnými veličinami s pozitivní binární reprezentovatelností, neboť tyto výsledky (byť značně neúplné) nám budou později užitečné při charakterizaci pozitivně diskrétně reprezentovatelných modelů.

Připomeňme, že třídy b–reprezentovatelných a b^+–reprezentovatelných modelů nejsou uzavřené na minory: kupříkladu model $I = \{\langle ab|cd\rangle, \langle ab|d\rangle\}$ je b^+–reprezentovatelný[1], zatímco model $I\,[^d$ podle lemmatu 19 b–reprezentovatelný není.

[1] Reprezentaci snadno získáme jako směsové rozdělení, kdy za podmínky $\xi_d = -1$ zvolíme za rozdělení $\boldsymbol{\xi}_{abc}$ nějakou reprezentaci modelu $\{\langle ab|c\rangle, \langle ab|\emptyset\rangle, \langle ac|b\rangle, \langle ac|\emptyset\rangle\}$ a za podmínky $\xi_d = 1$ zvolíme za rozdělení $\boldsymbol{\xi}_{abc}$ reprezentaci modelu $\{\langle ab|c\rangle, \langle ab|\emptyset\rangle, \langle bc|a\rangle, \langle bc|\emptyset\rangle\}$.

Použili jsme podobný postup jako v případě Gaussovské reprezentovatelnosti. Jako horní odhad počtu b^+–reprezentovatelných modelů můžeme vzít všechny nezávislostní modely I nad $N = \{1, 2, 3, 4\}$ takové, že jejich 3–minory neprotiřečí podmínce (1.10) z lemmatu 8 (tzn. jsou jednoho z osmi typů z obr. 3.1), a zároveň je splněna podmínka z lemmatu 19 (tzn. nenastane $I\,[_d = \{\langle ab|c\rangle, \langle ab|\emptyset\rangle\}$). Počet typů modelů splňujících výše uvedené je **240**.

Následně jsme náhodně generovali soubor racionálních čísel z intervalu $[-1, 1]$ indexovaný podmnožinami N a testovali, zda-li pokud vezmeme tento soubor jako soubor momentů $(e_A)_{A\subseteq N}$ dostaneme parametrizaci nějakého binárního rozdělení a případně jakého typu je model tomuto rozdělení příslušný. Pro generování jsme využili několika algoritmů, kdy pro zmenšení prohledávaného prostoru jsme například volili $e_1 = e_2 = e_3 = e_4 = 0$ nebo (na základě poznámky na konci pododdílu 1.2.3) se snažili odvodit binární reprezentaci na základě reprezentace Gaussovské. Takto jsme dostali kladnou binární reprezentaci **139** typů.

Ač je pochopitelně možné zvláště horní mez počtu b^+–reprezentovatelných typů dále vylepšovat (např. za pomoci dále v textu zmíněných vlastností diskrétně a pozitivně diskrétně reprezentovatelných modelů lze rozhodnout o reprezentovatelnosti dalších zhruba 40 typů), je zjevné, že zbývá velké množství typů, u nichž nejsme schopni rozhodnout, zda jsou či nejsou b^+–reprezentovatelné .

Uveďme však několik nových vlastností, specifických pro b^+–reprezentovatelné modely, jež jsme na základě našeho zkoumání získali.

Lemma 20. *Nechť I je b^+–reprezentovatelný nezávislostní model nad N a a, b, c, d jsou různé prvky N. Potom platí:*

i) $\{\langle ab|cd\rangle, \langle ab|c\rangle, \langle ad|\emptyset\rangle, \langle cd|\emptyset\rangle\} \subseteq I \implies$
$$\{\langle ad|c\rangle, \langle cd|a\rangle\} \subseteq I \;\vee\; \langle bd|\emptyset\rangle \in I$$

ii) $\{\langle ab|cd\rangle, \langle ab|c\rangle, \langle ad|\emptyset\rangle, \langle bd|\emptyset\rangle\} \subseteq I \implies$
$$\{\langle ad|b\rangle, \langle bd|a\rangle\} \subseteq I \;\vee\; \langle cd|\emptyset\rangle \in I$$

iii) $\{\langle ab|cd\rangle, \langle ab|\emptyset\rangle, \langle bc|\emptyset\rangle, \langle cd|\emptyset\rangle, \langle ad|\emptyset\rangle\} \subseteq I \implies$
$$\{\langle ad|c\rangle, \langle cd|a\rangle\} \subseteq I \;\vee\; \{\langle bd|c\rangle, \langle cd|b\rangle\} \subseteq I$$

iv) $\{\langle ab|cd\rangle, \langle ad|b\rangle, \langle cd|b\rangle\} \subseteq I \implies$
$$\{\langle bc|\emptyset\rangle, \langle cd|\emptyset\rangle, \langle bc|d\rangle\} \subseteq I \;\vee\; \langle ab|c\rangle \in I$$

v) $\{\langle ab|cd\rangle, \langle ab|\emptyset\rangle, \langle bc|\emptyset\rangle, \langle bd|\emptyset\rangle, \langle ab|c\rangle, \langle bc|a\rangle\} \subseteq I \Longrightarrow \ \langle ad|\emptyset\rangle \in I \ \vee$ $\langle ac|\emptyset\rangle \in I \ \vee \ \langle cd|\emptyset\rangle \in I \ \vee \ \langle ad|c\rangle \in I \ \vee \ \{\langle ab|d\rangle, \langle bd|a\rangle\} \subseteq I$

vi) $\{\langle ab|cd\rangle, \langle ab|d\rangle, \langle bc|a\rangle\} \subseteq I \Longrightarrow$ $\langle ac|\emptyset\rangle \in I \ \vee \ \langle ad|\emptyset\rangle \in I \ \vee \ \langle cd|\emptyset\rangle \in I \ \vee \ \langle cd|a\rangle \in I \ \vee \ \langle ac|d\rangle \in I$

Důkaz. Celý důkaz provedeme v parametrizaci s centrovanými momenty. Při výpočtech je vhodné použít software umožňující práci s algebraickými vzorci (Maple, Mathematica, ...), neboť jinak jsou příliš pracné.

i) Pokud předpokládáme, že $\xi_a \perp\!\!\!\perp \xi_d$ a $\xi_c \perp\!\!\!\perp \xi_d$, potom $f_{ad} = f_{cd} = 0$. Pokud navíc $\xi_a \perp\!\!\!\perp \xi_b | \xi_c$, potom $f_{ab} = \frac{f_{ac} f_{bc}}{1-e_c^2}$ a $f_{abc} = -2e_c f_{ab}$. Dosadíme-li toto do první rovnice příslušné vztahu $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, viz str. 17, dostaneme

$$f_{abcd} = -2e_c f_{abd} + f_{ac} f_{bd} \tag{3.1}$$

a dosadíme-li (3.1) do druhé rovnice příslušné vztahu $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, získáme po úpravě

$$0 = f_{bd} f_{acd}.$$

Za uvedených předpokladů tedy buďto $\xi_b \perp\!\!\!\perp \xi_d$ ($f_{bd} = 0$) nebo $\xi_a \perp\!\!\!\perp \xi_d | \xi_c$ a zároveň $\xi_c \perp\!\!\!\perp \xi_d | \xi_a$ ($f_{acd} = 0$).

ii) Pokud dosadíme předpoklady $f_{ad} = f_{bd} = 0$, $f_{ab} = \frac{f_{ac} f_{bc}}{1-e_c^2}$ a $f_{abc} = -2e_c f_{ab}$ do první rovnice příslušné vztahu $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, viz str. 17, dostaneme

$$f_{abcd} = \frac{-f_{ac} f_{bc} f_{cd} - 2e_c f_{abd} + 2e_c^3 f_{abd}}{1 - e_k^2}$$

a dosadíme-li toto do druhé rovnice příslušné vztahu $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, získáme

$$0 = f_{cd} f_{abd}.$$

Za uvedených předpokladů tedy buďto $\xi_c \perp\!\!\!\perp \xi_d$ nebo $\xi_a \perp\!\!\!\perp \xi_d | \xi_b$ a zároveň $\xi_b \perp\!\!\!\perp \xi_d | \xi_a$.

iii) Pokud dosadíme předpoklady $f_{ab} = f_{bc} = f_{cd} = f_{ad} = 0$ do první rovnice příslušné vztahu $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, dostaneme

$$f_{abcd} = -2e_c f_{abd} - 2e_d f_{abc} + f_{ac} f_{bd}. \tag{3.2}$$

Dosadíme-li (3.2) do druhé rovnice příslušné vztahu $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, získáme

$$f_{bd} = \frac{f_{abc}(1 - e_d^2)}{f_{acd}}. \tag{3.3}$$

Dosadíme-li dále (3.2) a (3.3) do třetí rovnice příslušné $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, získáme

$$f_{ac} = \frac{f_{abd}(1 - e_c^2)}{f_{bcd}}. \tag{3.4}$$

Nakonec, pokud (3.2), (3.3) a (3.4) dosadíme do čtvrté rovnice příslušné vztahu $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, dostaneme

$$0 = f_{acd} f_{bcd}.$$

Nutně tedy buďto $\xi_d \perp\!\!\!\perp \boldsymbol{\xi}_{ac}$ ($f_{acd} = 0$) nebo $\xi_c \perp\!\!\!\perp \boldsymbol{\xi}_{bd}$ ($f_{bcd} = 0$).

iv) Postupujeme opět analogicky jako v předchozích případech. Z rovnic příslušných nezávislostním vztahům $\xi_a \perp\!\!\!\perp \xi_d | \xi_b$ a $\xi_c \perp\!\!\!\perp \xi_d | \xi_b$ vyjádříme f_{ad}, f_{abd}, f_{cd} a f_{bcd} a z rovnic příslušných $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$ vyjádříme f_{abcd}, f_{abc} a f_{acd}. Dosazením do poslední rovnice příslušné $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$ zjistíme, že buďto $f_{bc} = 0$ ($\xi_b \perp\!\!\!\perp \xi_c | \emptyset$) nebo

$$0 = \begin{array}{l} (f_{bd} + 1 + e_b + e_b e_d + e_d) \cdot (f_{bd} - 1 - e_b + e_b e_d + e_d) \cdot \\ (f_{bd} - 1 + e_b + e_b e_d - e_d) \cdot (f_{bd} + 1 - e_b + e_b e_d - e_d) \cdot \\ (f_{ac} f_{bc} + e_c^2 f_{ab} - f_{ab}). \end{array} \tag{3.5}$$

Ovšem za předpokladu kladného rozdělení $\boldsymbol{\xi}_{bd}$ může být jediný člen na pravé straně (3.5) nulový a to $f_{ac} f_{bc} + e_c^2 f_{ab} - f_{ab}$. Dosadíme-li toto do třetí rovnice příslušné $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, dostaneme $f_{abc} + 2e_c f_{ac} f_{bc} - f_{abc} e_c^2 = 0$ ($\xi_a \perp\!\!\!\perp \xi_b | \xi_c$).

v) Analogicky, za předpokladu $f_{ab} = f_{bc} = f_{abc} = f_{bd} = 0$ použijeme první rovnici příslušnou $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{cd}$, abychom vyjádřili f_{abcd}, druhou, abychom vyjádřili f_{bcd}, a z třetí dostáváme, že buďto $f_{ad} = 0$ ($\xi_a \perp\!\!\!\perp \xi_d | \emptyset$) nebo $f_{abd} = 0$ ($\xi_a \perp\!\!\!\perp \xi_b | \xi_d$, $\xi_b \perp\!\!\!\perp \xi_d | \xi_a$) nebo $f_{ad} = \frac{f_{ac} f_{cd}}{1 - e_c^2}$. Pokud za f_{ad} do předchozích rovnic dosadíme $\frac{f_{ac} f_{cd}}{1 - e_c^2}$, zjistíme, že buďto $f_{cd} = 0$ ($\xi_c \perp\!\!\!\perp \xi_d | \emptyset$) nebo $f_{ac} = 0$ ($\xi_a \perp\!\!\!\perp \xi_c | \emptyset$) nebo $f_{acd} - f_{acd} e_c^2 + 2e_c f_{ac} f_{cd} = 0$ ($\xi_a \perp\!\!\!\perp \xi_d | \xi_c$).

vi) Obdobně, lze za předpokladů ukázat, že buďto je f_{ac} nebo f_{ad} nebo f_{cd} rovno nule nebo

$$0 = \begin{array}{l} (f_{cd} e_a^2 - f_{cd} + f_{ad} f_{ac}) \cdot (f_{ad} + 1 + e_d + e_a e_d + e_a) \cdot \\ (f_{ad} + 1 - e_d + e_a e_d - e_a) \cdot (f_{ad} - 1 - e_d + e_a e_d + e_a) \cdot \\ (f_{ad} - 1 + e_d + e_a e_d - e_a) \cdot (f_{ad} f_{cd} - f_{ac} + f_{ac} e_d^2), \end{array}$$

odkud za předpokladu kladně rozděleného $\boldsymbol{\xi}_{ad}$ plyne, že buďto $f_{cd}e_a^2 - f_{cd} + f_{ad}f_{ac} = 0$ nebo $f_{ad}f_{cd} - f_{ac} + f_{ac}e_d^2 = 0$. A dosazením za f_{cd} (resp. f_{ac}) ověříme, že skutečně $\xi_c \perp\!\!\!\perp \xi_d | \xi_a$ (resp. $\xi_a \perp\!\!\!\perp \xi_d | \xi_c$).

□

3.2 Obecná diskrétní reprezentovatelnost

Problém d–reprezentovatelnosti pro $|N| = 4$ byl vyřešen F. Matúšem (s přispěním M. Studeného) v sérii článků [MS95], [Mat95] a [Mat99]. Závěry byly v čitelnější a k čtenáři vlídnější formě následně prezentovány v článcích [SB94] a [Mat97][2].

Klíčovým nástrojem pro popis třídy diskrétně reprezentovatelných modelů je následující lemma.

Lemma 21. *Nechť $I = \mathcal{I}(\boldsymbol{\xi})$ a $I' = \mathcal{I}(\boldsymbol{\xi}')$ jsou d–reprezentovatelné nezávislostní modely nad množinou N. Potom $I \cap I'$ je taktéž d–reprezentovatelný. Navíc, pokud jsou $\boldsymbol{\xi}$ a $\boldsymbol{\xi}'$ pozitivní reprezentace, potom je i $I \cap I'$ pozitivně diskrétně reprezentovatelný.*

Důkaz. Nechť $\boldsymbol{X} = \prod X_a$ a $\boldsymbol{X}' = \prod X'_a$ jsou po řadě stavové prostory $\boldsymbol{\xi}$ a $\boldsymbol{\xi}'$. Definujme požadovanou reprezentaci $\hat{\boldsymbol{\xi}}$ modelu $\mathcal{I}(\hat{\boldsymbol{\xi}}) = I \cap I'$ na prostoru

$$\widehat{\boldsymbol{X}} = \prod_{a \in N} (X_a \times X'_a)$$

s následujícím rozdělením

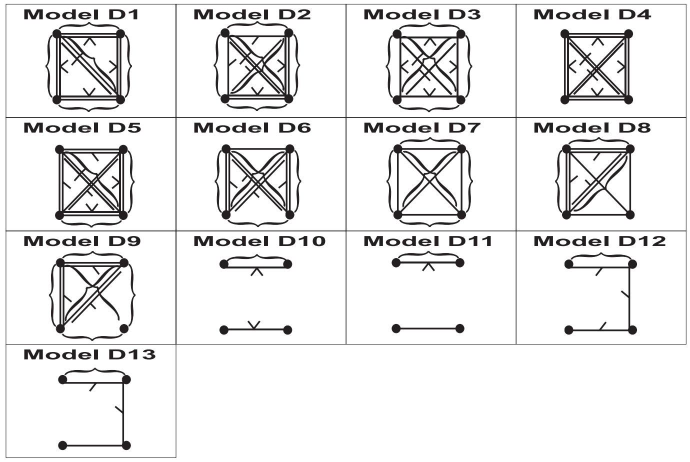
$$P(\hat{\boldsymbol{\xi}} = (x_a, x'_a)_{a \in N}) = P(\boldsymbol{\xi} = (x_a)_{a \in N}) \cdot P(\boldsymbol{\xi}' = (x'_a)_{a \in N}).$$

Tudíž zjevně $\hat{\xi}_a \perp\!\!\!\perp \hat{\xi}_b | \hat{\boldsymbol{\xi}}_C$ právě tehdy, když $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_C$ a zároveň $\xi'_a \perp\!\!\!\perp \xi'_b | \boldsymbol{\xi}'_C$. □

Jako důsledek lemmatu 21 odvodíme, že minory diskrétně reprezentovatelných modelů jsou taktéž diskrétně reprezentovatelné.

Lemma 22. *Třídy d–reprezentovatelných a d^+–reprezentovatelných modelů jsou uzavřené na minory.*

[2] Abychom předešli nedorozumění, dovolil bych si upozornit na drobný přetisk v jinak bezchybném článku [Mat97] na str. 21, obr. 14. Nad horní čarou v prvních dvou diagramech má být $\emptyset$ namísto $*$.



Obrázek 3.2: Generátory třídy d–reprezentovatelných modelů.

Důkaz. Nechť I je obecně (resp. pozitivně) diskrétně reprezentovatelný model a $\boldsymbol{\xi}$ příslušná reprezentace. Potom snadno nahlédneme, že pro $E \subseteq N$ platí $I\,[^{\emptyset}_{E} = \mathcal{I}(\boldsymbol{\xi}_{N \setminus E})$.

Dále, z definice podmíněné nezávislosti pro diskrétní rozdělení je pro $F \subseteq N$ minor $I\,[^{F}_{\emptyset}$ průnikem nezávislostních modelů příslušných rozdělení náhodného vektoru $\boldsymbol{\xi}_{N \setminus F}$ podmíněnému $\boldsymbol{\xi}_F = \boldsymbol{x}_F$ přes všechny hodnoty $\boldsymbol{x}_F$, již $\boldsymbol{\xi}_F$ nabývá s nenulovou pravděpodobností. Platí tedy, že minor $I\,[^{F}_{\emptyset}$ je průnikem konečně mnoha diskrétně reprezentovatelných modelů a tudíž musí být diskrétně reprezentovatelný.

Pro libovolný minor $I\,[^{F}_{E}$ dostáváme tvrzení kombinací výše uvedených dvou kroků za pomoci identity (1.11). □

Z lemmatu 21 plyne, že průnik konečného počtu d–reprezentovatelných modelů je opět d–reprezentovatelný. Tudíž přirozenou cestou, jak popsat d–reprezentovatelné modely, je nalezení množiny (d–reprezentovatelných) modelů–generátorů takových, že každý netriviální[3] d–reprezentovatelný model můžeme zapsat jako průnik těchto generátorů. Ve výše zmíněné literatuře je ukázáno, že modely–generátory pro $|N| = 4$ jsou **13** typů a to právě těch zobrazených na obr. 3.2.

Poněkud pikantní je otázka celkového počtu d–reprezentovatelných modelů (a typů) pro $|N| = 4$. V řadě pramenů můžeme nalézt, že počet d–reprezentovatelných nezávislostních modelů nad $N = \{1, 2, 3, 4\}$ je 18300,

[3]tj. neprázdný, neúplný

viz např. [Stu02], [LR02], [RSSW03], [Šim06b]... To však není pravda a skutečný počet těchto modelů je **18478** a tyto se rozpadají na **1098** typů. Uvedená čísla byla překontrolována několika počítačovými programy včetně programu SG POKUS P. Bočka. Seznam d–reprezentovatelných modelů a typů lze nalézt na přiloženém CD.

Pro úplnost si ještě uveďme seznam inferenčních pravidel vyplynuvších z dosažené charakterizace diskrétně reprezentovatelných modelů nad čtyřprvkovou množinou, jenž byl převzat z [SB94].

Lemma 23. *Nechť N je konečná množina a I je d–reprezentovatelný nezávislostní model nad N. Potom pro a, b, c, d různé prvky N a množiny $E \subseteq N \setminus abcd$ a $F \subseteq N \setminus abcd$ platí:*

i) $\{\langle ab|E\rangle, \langle ac|bE\rangle\} \subseteq I \iff \{\langle ac|E\rangle, \langle ab|cE\rangle\} \subseteq I$

ii) $\{\langle ab|cdF\rangle, \langle cd|aF\rangle, \langle cd|bF\rangle, \langle ab|F\rangle\} \subseteq I \iff$
$\iff \{\langle cd|abF\rangle, \langle ab|cF\rangle, \langle ab|dF\rangle, \langle cd|F\rangle\} \subseteq I$

iii) $\{\langle ab|cdF\rangle, \langle ad|bF\rangle, \langle cd|aF\rangle, \langle bc|F\rangle\} \subseteq I \iff$
$\iff \{\langle ad|bcF\rangle, \langle ab|dF\rangle, \langle bc|aF\rangle, \langle cd|F\rangle\} \subseteq I$

iv) $\{\langle ac|dF\rangle, \langle bd|cF\rangle, \langle bc|aF\rangle, \langle ad|bF\rangle\} \subseteq I \iff$
$\iff \{\langle ad|cF\rangle, \langle bc|dF\rangle, \langle bd|aF\rangle, \langle ac|bF\rangle\} \subseteq I$

v) $\{\langle ab|cF\rangle, \langle ac|dF\rangle, \langle ad|bF\rangle\} \subseteq I \iff \{\langle ac|bF\rangle, \langle ad|cF\rangle, \langle ab|dF\rangle\} \subseteq I$

vi) $\{\langle ab|cdF\rangle, \langle cd|abF\rangle, \langle ac|F\rangle, \langle bd|F\rangle\} \subseteq I \iff$
$\iff \{\langle ac|bdF\rangle, \langle bd|acF\rangle, \langle ab|F\rangle, \langle cd|F\rangle\} \subseteq I$

vii) $\{\langle ab|cF\rangle, \langle ab|dF\rangle, \langle bc|aF\rangle, \langle cd|F\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

viii) $\{\langle ab|cF\rangle, \langle ac|dF\rangle, \langle bc|aF\rangle, \langle bd|F\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

ix) $\{\langle ab|cF\rangle, \langle ac|dF\rangle, \langle cd|aF\rangle, \langle bd|F\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

x) $\{\langle ab|cF\rangle, \langle ac|dF\rangle, \langle ad|cF\rangle, \langle bd|F\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

xi) $\{\langle ab|dF\rangle, \langle ac|bF\rangle, \langle bd|cF\rangle, \langle ad|bF\rangle\} \subseteq I \implies \langle ab|cF\rangle \in I$

xii) $\{\langle ab|dF\rangle, \langle ac|bF\rangle, \langle bd|cF\rangle, \langle bd|aF\rangle\} \subseteq I \implies \langle ab|cF\rangle \in I$

xiii) $\{\langle ab|dF\rangle, \langle ac|bF\rangle, \langle bd|cF\rangle, \langle bc|dF\rangle\} \subseteq I \implies \langle ab|cF\rangle \in I$

xiv) $\{\langle ab|dF\rangle, \langle ac|bF\rangle, \langle bd|cF\rangle, \langle cd|bF\rangle\} \subseteq I \implies \langle ab|cF\rangle \in I$

xv) $\{\langle ab|dF\rangle, \langle ac|bF\rangle, \langle bd|acF\rangle, \langle ad|bF\rangle\} \subseteq I \implies \langle ab|cdF\rangle \in I$

xvi) $\{\langle ab|dF\rangle, \langle ac|bF\rangle, \langle bd|acF\rangle, \langle bd|aF\rangle\} \subseteq I \implies \langle ab|cdF\rangle \in I$

xvii) $\{\langle ab|cF\rangle, \langle ab|dF\rangle, \langle ac|dF\rangle, \langle cd|abF\rangle\} \subseteq I \implies \langle ab|cdF\rangle \in I$

xviii) $\{\langle ac|dF\rangle, \langle ab|dF\rangle, \langle bc|aF\rangle, \langle ad|bcF\rangle\} \subseteq I \implies \langle ab|cdF\rangle \in I$

xix) $\{\langle ab|dF\rangle, \langle ac|bF\rangle, \langle ac|dF\rangle, \langle bd|cF\rangle\} \subseteq I \implies \langle ab|cF\rangle \in I$

xx) $\{\langle ab|cF\rangle, \langle ab|dF\rangle, \langle ab|F\rangle, \langle cd|abF\rangle\} \subseteq I \implies \langle ab|cdF\rangle \in I$

xxi) $\{\langle ab|cF\rangle, \langle ab|dF\rangle, \langle cd|aF\rangle, \langle cd|F\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

xxii) $\{\langle ab|cF\rangle, \langle ac|dF\rangle, \langle bd|cF\rangle, \langle bd|F\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

xxiii) $\{\langle ab|cF\rangle, \langle ac|dF\rangle, \langle bd|aF\rangle, \langle bd|F\rangle\} \subseteq I \implies \langle ab|F\rangle \in I$

xxiv) $\{\langle ab|cF\rangle, \langle ad|bF\rangle, \langle bc|adF\rangle, \langle ad|F\rangle\} \subseteq I \implies \langle ab|cdF\rangle \in I$

xxv) $\{\langle ab|cF\rangle, \langle ad|bF\rangle, \langle bc|adF\rangle, \langle cd|bF\rangle\} \subseteq I \implies \langle ab|cdF\rangle \in I$

3.3 Pozitivní diskrétní reprezentovatelnost

Je s podivem, že ač se v praxi s konceptem kladného diskrétního rozdělením setkáváme poměrně často, nebyla dosud reprezentovatelnost nezávislostních modelů za pomoci pozitivního diskrétního rozdělení podrobena důkladnějšímu zkoumání a výsledky jsou zatím spíše skromné.

Je zjevné, že množina d^+–reprezentovatelných modelů je podmnožinou d–reprezentovatelných modelů. Jsou známy dvě podmínky, jež musí navíc d^+–reprezentovatelný model splňovat: podmínku (1.10) z druhé části lemmatu 8 a podmínku ze [Spo94], která je v mírně pozměněné podobě uvedena v následujícím lemmatu.

Lemma 24. *Nechť a, b, c, d jsou různé prvky množiny N. Jestliže I je d^+–reprezentovatelný nezávislostní model nad N takový, že*

$$\{\langle ab|cd\rangle, \langle cd|ab\rangle, \langle cd|a\rangle\} \subseteq I, \tag{3.6}$$

potom

$$\langle cd|b\rangle \in I \iff \langle cd|\emptyset\rangle \in I.$$

Důkaz. Nechť $\boldsymbol{\xi}$ je d^+–reprezentace modelu I. Za předpokladu (3.6) lze výraz $P(\boldsymbol{\xi}_{abcd} = \boldsymbol{x}_{abcd})$ rozepsat jako $\frac{P(\boldsymbol{\xi}_{abc}=\boldsymbol{x}_{abc})P(\boldsymbol{\xi}_{abd}=\boldsymbol{x}_{abd})}{P(\boldsymbol{\xi}_{ab}=\boldsymbol{x}_{ab})}$ nebo také jako $\frac{P(\boldsymbol{\xi}_{ac}=\boldsymbol{x}_{ac})P(\boldsymbol{\xi}_{ad}=\boldsymbol{x}_{ad})P(\boldsymbol{\xi}_{bcd}=\boldsymbol{x}_{bcd})}{P(\xi_a=x_a)P(\boldsymbol{\xi}_{cd}=\boldsymbol{x}_{cd})}$. Platí tedy identita

$$\frac{P(\boldsymbol{\xi}_{abc} = \boldsymbol{x}_{abc})P(\boldsymbol{\xi}_{abd} = \boldsymbol{x}_{abd})}{P(\boldsymbol{\xi}_{ab} = \boldsymbol{x}_{ab})} = \frac{P(\boldsymbol{\xi}_{ac} = \boldsymbol{x}_{ac})P(\boldsymbol{\xi}_{ad} = \boldsymbol{x}_{ad})P(\boldsymbol{\xi}_{bcd} = \boldsymbol{x}_{bcd})}{P(\xi_a = x_a)P(\boldsymbol{\xi}_{cd} = \boldsymbol{x}_{cd})}.$$

Zvolíme-li x_a libovolné, pevné, můžeme předchozí rovnost přepsat jako

$$P(\boldsymbol{\xi}_{bcd} = \boldsymbol{x}_{bcd}) = P(\boldsymbol{\xi}_{cd} = \boldsymbol{x}_{cd})g(\boldsymbol{x}_{bc})h(\boldsymbol{x}_{bd}), \tag{3.7}$$

kde g a h jsou nějaké funkce s oborem hodnot $(0, \infty)$. Pokud dále předpokládáme $\xi_c \perp\!\!\!\perp \xi_d$ neboli $P(\boldsymbol{\xi}_{cd} = \boldsymbol{x}_{cd}) = P(\xi_c = x_c)P(\xi_d = x_d)$, dostáváme ze (3.7) dle lemmatu 1, části vi), že i $\xi_c \perp\!\!\!\perp \xi_d | \xi_b$. A naopak, pokud předpokládáme $\xi_c \perp\!\!\!\perp \xi_d | \xi_b$ a zafixujeme v (3.7) hodnotu x_b, dostáváme $\xi_c \perp\!\!\!\perp \xi_d$. □

Nad $N = \{1, 2, 3, 4\}$ existuje **5547** d–reprezentovatelných modelů (**356** typů), které navíc splňují podmínku (1.10) z lemmatu 8 a podmínku z lemmatu 24 a určují nám tedy odhad shora pro počet pozitivně reprezentovatelných modelů nad N. Na základě lemmatu 21 můžeme tuto nadmnožinu množiny d^+–reprezentovatelných modelů opět popsat množinou modelů–generátorů, která ji za pomoci operace průniku generuje. Seznam modelů–generátorů (resp. jejich 23 typů) lze nalézt na přiloženém CD.

K nalezení odhadu zdola jsme využili 139 b^+–reprezentovatelných typů nalezených v oddíle 3.1. Jelikož z lemmatu 21 víme, že množina d^+–reprezentovatelných modelů je uzavřená na operaci průniku, dostáváme aplikací uzávěru na průnik množinu **4555** modelů (**299** typů), jež jsou nutně d^+–reprezentovatelné. Zbývalo tedy rozhodnout o d^+–reprezentovatelnosti **57** typů P1–P57, které jsou znázorněny na obr. 3.3.

Další vlastnosti charakteristické pro d^+–reprezentovatelnost se však autorovi nalézt nepodařilo. Pokud by existovaly, musel by jejich důkaz být založen na jiných „principech“ než je tomu u grafoidové či Spohnovy vlastnosti.

Se značným usilím se byla ověřena d^+–reprezentovatelnost P54 a P49 (a tudíž díky lemmatu 21 i P52, P50 a P12), pro reprezentace viz tab. 3.1. Pracovní (zatím nepotvrzená, ani nevyvrácená) **hypotéza** zní, že všechny typy na obr. 3.3 jsou reprezentovatelné. K jejímu důkazu by bylo třeba nalézt reprezentace pro P57, P46, P45 a P37, P41 a P32, P39 a P30, P24, P19, P3,

$\boldsymbol{x}_{abcd}$	P54	P49
0 0 0 0	$\frac{1}{4} \cdot \frac{1 \cdot 1 \cdot 622}{3 \cdot 3 \cdot 83325}$	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 33}{3 \cdot 7 \cdot 112}$
0 0 0 1	$\frac{1}{4} \cdot \frac{1 \cdot 1 \cdot 82703}{3 \cdot 3 \cdot 83325}$	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 79}{3 \cdot 7 \cdot 112}$
0 0 1 0	$\frac{1}{4} \cdot \frac{1 \cdot 2 \cdot 622}{3 \cdot 3 \cdot 83325}$	$\frac{1}{3} \cdot \frac{1 \cdot 6 \cdot 33}{3 \cdot 7 \cdot 112}$
0 0 1 1	$\frac{1}{4} \cdot \frac{1 \cdot 2 \cdot 82703}{3 \cdot 3 \cdot 83325}$	$\frac{1}{3} \cdot \frac{1 \cdot 6 \cdot 79}{3 \cdot 7 \cdot 112}$
0 1 0 0	$\frac{1}{4} \cdot \frac{2 \cdot 1 \cdot 622}{3 \cdot 3 \cdot 83325}$	$\frac{1}{3} \cdot \frac{2 \cdot 1 \cdot 33}{3 \cdot 7 \cdot 112}$
0 1 0 1	$\frac{1}{4} \cdot \frac{2 \cdot 1 \cdot 82703}{3 \cdot 3 \cdot 83325}$	$\frac{1}{3} \cdot \frac{2 \cdot 1 \cdot 79}{3 \cdot 7 \cdot 112}$
0 1 1 0	$\frac{1}{4} \cdot \frac{2 \cdot 2 \cdot 622}{3 \cdot 3 \cdot 83325}$	$\frac{1}{3} \cdot \frac{2 \cdot 6 \cdot 33}{3 \cdot 7 \cdot 112}$
0 1 1 1	$\frac{1}{4} \cdot \frac{2 \cdot 2 \cdot 82703}{3 \cdot 3 \cdot 83325}$	$\frac{1}{3} \cdot \frac{2 \cdot 6 \cdot 79}{3 \cdot 7 \cdot 112}$
1 0 0 0	$\frac{1}{4} \cdot \frac{1 \cdot 1 \cdot 299}{4 \cdot 4 \cdot 27775}$	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 1}{2 \cdot 5 \cdot 8}$
1 0 0 1	$\frac{1}{4} \cdot \frac{1 \cdot 1 \cdot 27476}{4 \cdot 4 \cdot 27775}$	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 7}{2 \cdot 5 \cdot 8}$
1 0 1 0	$\frac{1}{4} \cdot \frac{1 \cdot 3 \cdot 299}{4 \cdot 4 \cdot 27775}$	$\frac{1}{3} \cdot \frac{1 \cdot 4 \cdot 1}{2 \cdot 5 \cdot 8}$
1 0 1 1	$\frac{1}{4} \cdot \frac{1 \cdot 3 \cdot 27476}{4 \cdot 4 \cdot 27775}$	$\frac{1}{3} \cdot \frac{1 \cdot 4 \cdot 7}{2 \cdot 5 \cdot 8}$
1 1 0 0	$\frac{1}{4} \cdot \frac{3 \cdot 1 \cdot 299}{4 \cdot 4 \cdot 27775}$	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 1}{2 \cdot 5 \cdot 8}$
1 1 0 1	$\frac{1}{4} \cdot \frac{3 \cdot 1 \cdot 27476}{4 \cdot 4 \cdot 27775}$	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 7}{2 \cdot 5 \cdot 8}$
1 1 1 0	$\frac{1}{4} \cdot \frac{3 \cdot 3 \cdot 299}{4 \cdot 4 \cdot 27775}$	$\frac{1}{3} \cdot \frac{1 \cdot 4 \cdot 1}{2 \cdot 5 \cdot 8}$
1 1 1 1	$\frac{1}{4} \cdot \frac{3 \cdot 3 \cdot 27476}{4 \cdot 4 \cdot 27775}$	$\frac{1}{3} \cdot \frac{1 \cdot 4 \cdot 7}{2 \cdot 5 \cdot 8}$

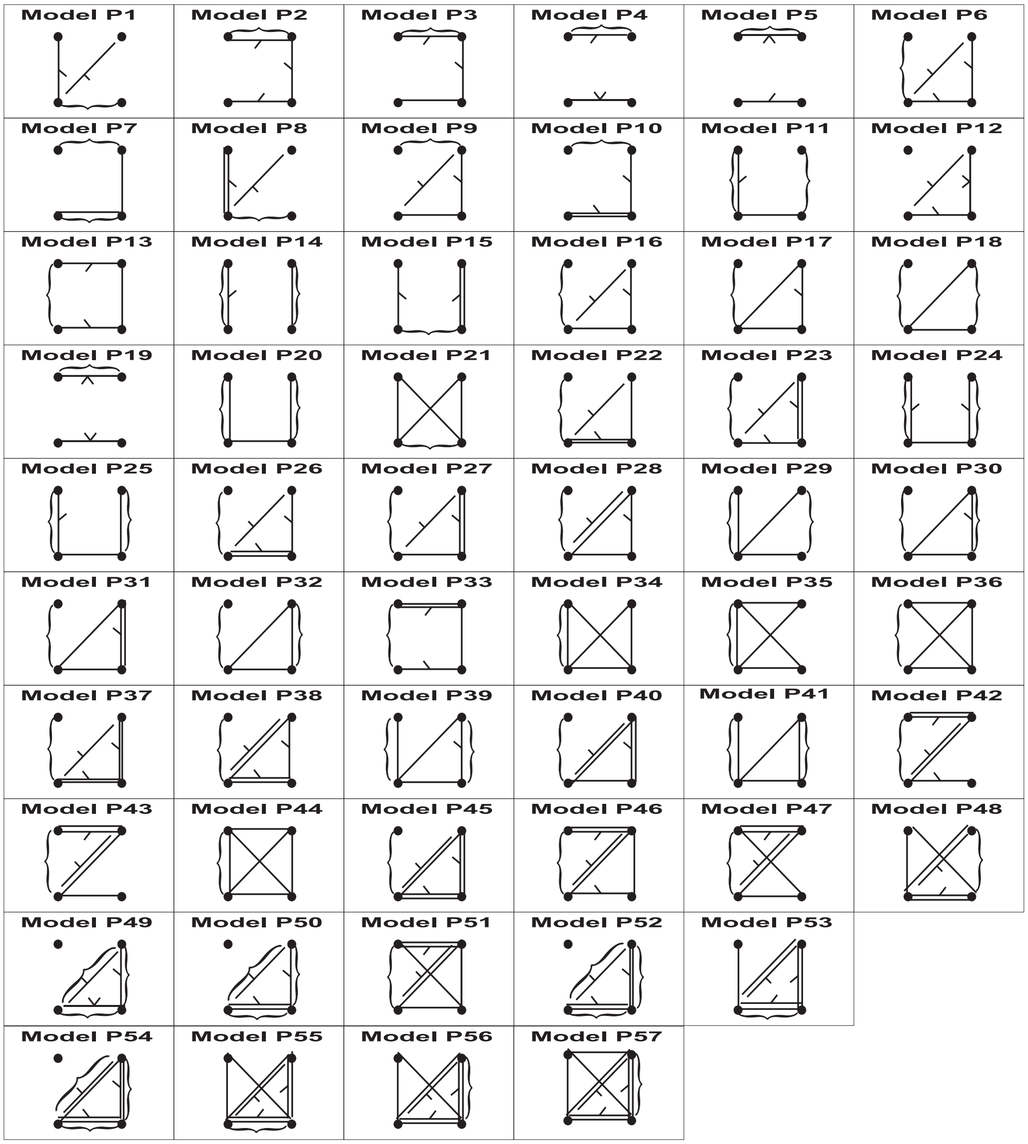
$\boldsymbol{x}_{abcd}$	P54	P49
2 0 0 0	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 1}{2 \cdot 3 \cdot 100}$	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 1}{6 \cdot 3 \cdot 7}$
2 0 0 1	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 99}{2 \cdot 3 \cdot 100}$	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 6}{6 \cdot 3 \cdot 7}$
2 0 1 0	$\frac{1}{3} \cdot \frac{1 \cdot 2 \cdot 1}{2 \cdot 3 \cdot 100}$	$\frac{1}{3} \cdot \frac{1 \cdot 2 \cdot 1}{6 \cdot 3 \cdot 7}$
2 0 1 1	$\frac{1}{3} \cdot \frac{1 \cdot 2 \cdot 99}{2 \cdot 3 \cdot 100}$	$\frac{1}{3} \cdot \frac{1 \cdot 2 \cdot 6}{6 \cdot 3 \cdot 7}$
2 1 0 0	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 1}{2 \cdot 3 \cdot 100}$	$\frac{1}{3} \cdot \frac{5 \cdot 1 \cdot 1}{6 \cdot 3 \cdot 7}$
2 1 0 1	$\frac{1}{3} \cdot \frac{1 \cdot 1 \cdot 99}{2 \cdot 3 \cdot 100}$	$\frac{1}{3} \cdot \frac{5 \cdot 1 \cdot 6}{6 \cdot 3 \cdot 7}$
2 1 1 0	$\frac{1}{3} \cdot \frac{1 \cdot 2 \cdot 1}{2 \cdot 3 \cdot 100}$	$\frac{1}{3} \cdot \frac{5 \cdot 2 \cdot 1}{6 \cdot 3 \cdot 7}$
2 1 1 1	$\frac{1}{3} \cdot \frac{1 \cdot 2 \cdot 99}{2 \cdot 3 \cdot 100}$	$\frac{1}{3} \cdot \frac{5 \cdot 2 \cdot 6}{6 \cdot 3 \cdot 7}$
3 0 0 0	$\frac{1}{6} \cdot \frac{1 \cdot 91 \cdot 1}{3 \cdot 120 \cdot 101}$	–
3 0 0 1	$\frac{1}{6} \cdot \frac{1 \cdot 91 \cdot 100}{3 \cdot 120 \cdot 101}$	–
3 0 1 0	$\frac{1}{6} \cdot \frac{1 \cdot 29 \cdot 1}{3 \cdot 120 \cdot 101}$	–
3 0 1 1	$\frac{1}{6} \cdot \frac{1 \cdot 29 \cdot 100}{3 \cdot 120 \cdot 101}$	–
3 1 0 0	$\frac{1}{6} \cdot \frac{2 \cdot 91 \cdot 1}{3 \cdot 120 \cdot 101}$	–
3 1 0 1	$\frac{1}{6} \cdot \frac{2 \cdot 91 \cdot 100}{3 \cdot 120 \cdot 101}$	–
3 1 1 0	$\frac{1}{6} \cdot \frac{2 \cdot 29 \cdot 1}{3 \cdot 120 \cdot 101}$	–
3 1 1 1	$\frac{1}{6} \cdot \frac{2 \cdot 29 \cdot 100}{3 \cdot 120 \cdot 101}$	–

Tabulka 3.1: Pravděpodobnostní distribuce reprezentací modelů P54 a P49

P2, celkem tedy **12** reprezentací, neboť za předpokladu, že tyto jsou reprezentovatelné dostaneme zbylé průsekem s již známými d^+–reprezentovatelnými modely. Naopak pokud výše uvedená hypotéza neplatí, musí existovat vlastnost, která vyloučí reprezentovatelnost alespoň jednoho ze zmíněných 12 typů.

Shňme tedy dosažené poznatky do závěrečné věty:

Věta 4. *Nad množinou* $N = \{1, 2, 3, 4\}$ *existuje* 18478 *diskrétně reprezentovatelných nezávislostních modelů* 1098 *typů. Kladně reprezentovatelných je 4607–5547 z nich (304–356 typů).*



Obrázek 3.3: Typy, jejíž reprezentace musí být nalezena pro ověření výše uvedené **hypotézy** na str. 50.

3.4 Neexistence konečné charakterizace

V tomto oddílu dokážeme, že diskrétně, ani binárně reprezentovatelné nezávislostní modely nelze charakterizovat konečným seznamem zakázaných minorů. Postupovat budeme obdobně jako v případě Gaussovsky reprezentovatelných modelů: ukážeme, že pro libovolné $n \geq 4$ není model

$$J_n = \{\langle 12|3\rangle, \langle 13|4\rangle, \dots, \langle 1(n-1)|n\rangle, \langle 1n|2\rangle\}$$

d–reprezentovatelný, zatímco jeho $(n-1)$–minory jsou b^+–reprezentovatelné, a následně využijeme lemmatu 9 z oddílu 1.4.

Uvědomme si, že třídy b–reprezentovatelných a b^+–reprezentovatelných modelů konečnou charakterizaci seznamem zakázaných minorů mít nemohou už proto, že nejsou uzavřené na minory. Z níže uvedených argumentů však vyplývá, že konečná charakterizace neexistuje pro libovolnou třídu diskrétních rozdělení, která je uzavřená na minory a zahrnuje b^+–reprezentovatelné modely. Příkladem je třída d^+–reprezentovatelných modelů.

Lemma 25. *Pro libovolné $n \geq 4$ je nezávislostní model*

$$J_n = \{\langle 12|3\rangle, \langle 13|4\rangle, \dots, \langle 1(n-1)|n\rangle, \langle 1n|2\rangle\} \tag{3.8}$$

zakázaným minorem pro třídu diskrétně reprezentovatelných modelů.

Důkaz. Předpokládejme, že by existoval d–reprezentovatelný nezávislostní model I nad množinou $\{1, \dots, n\} \cup EF$ takový, že $I[^F_E = J_n$. Potom by dle lemmatu 22 byl d–reprezentovatelný i model J_n. Označme příslušnou d–reprezentaci $\boldsymbol{\xi}$ a dále pokračujme Studeného důkazem převzatým z [Stu92], str. 6.

Z druhé části lemmatu 2 pro nezávislostní vztahy $\xi_1 \perp\!\!\!\perp \xi_2 | \xi_3$, $\xi_1 \perp\!\!\!\perp \xi_3 | \xi_4$, $\dots$,$\xi_1 \perp\!\!\!\perp \xi_{n-1} | \xi_n$, $\xi_1 \perp\!\!\!\perp \xi_n | \xi_2$ vyplývá, že

$$\begin{aligned}
0 \;=\; & \big[MI(\boldsymbol{\xi}_{123}) + MI(\boldsymbol{\xi}_3) - MI(\boldsymbol{\xi}_{23}) - MI(\boldsymbol{\xi}_{13})\big] + \\
& +\big[MI(\boldsymbol{\xi}_{134}) + MI(\boldsymbol{\xi}_4) - MI(\boldsymbol{\xi}_{34}) - MI(\boldsymbol{\xi}_{14})\big] + \\
& + \cdots + \\
& +\big[MI(\boldsymbol{\xi}_{12n}) + MI(\boldsymbol{\xi}_2) - MI(\boldsymbol{\xi}_{2n}) - MI(\boldsymbol{\xi}_{12})\big] \\
=\; & \big[MI(\boldsymbol{\xi}_{123}) + MI(\boldsymbol{\xi}_2) - MI(\boldsymbol{\xi}_{23}) - MI(\boldsymbol{\xi}_{12})\big] + \\
& +\big[MI(\boldsymbol{\xi}_{134}) + MI(\boldsymbol{\xi}_3) - MI(\boldsymbol{\xi}_{34}) - MI(\boldsymbol{\xi}_{13})\big] + \\
& + \cdots + \\
& + \big[MI(\boldsymbol{\xi}_{12n}) + MI(\boldsymbol{\xi}_n) - MI(\boldsymbol{\xi}_{2n}) - MI(\boldsymbol{\xi}_{1n})\big],
\end{aligned}$$

přičemž členy v závorkách musí být dle lemmatu 2 nezáporné a tedy nulové. To však znamená, že nutně $\xi_1 \perp\!\!\!\perp \xi_3|\xi_2$, $\xi_1 \perp\!\!\!\perp \xi_4|\xi_3$, $\dots$, $\xi_1 \perp\!\!\!\perp \xi_2|\xi_n$, což je spor s předpokladem $\mathcal{I}(\boldsymbol{\xi}) = J_n$. $\square$

Lemma 26. *Pro libovolné $n \geq 3$ jsou nezávislostní modely*

$$\begin{aligned} I_n &= \{\langle 12|3\rangle, \langle 13|4\rangle, \dots, \langle 1(n-1)|n\rangle\}, \\ I_n^* &= \{\langle 12|\emptyset\rangle\}, \qquad I_n^{**} = \emptyset \end{aligned}$$

kladně binárně reprezentovatelné nad množinou $N = \{1, \dots, n\}$.

Důkaz. Důkaz je analogický k důkazu lemmatu 18. Uvažme binární rozdělení $\boldsymbol{\xi}$ takové, že

$$\begin{aligned} \forall a: \ e_a &= 0, \\ \forall a > 1: \ e_{1a} &= \epsilon^{n-a+1}, \\ \forall a, b > 1, a \neq b: \ e_{ab} &= \epsilon, \\ \forall A \subseteq N, |A| > 2: \ e_A &= 0. \end{aligned}$$

Pro $\epsilon > 0$ dostatečně blízko 0 se zjevně jedná o kladné binární rozdělení a níže ukážeme, že $\mathcal{I}(\boldsymbol{\xi}) = I_n$. Podmínku pro platnost $\xi_a \perp\!\!\!\perp \xi_b|\boldsymbol{\xi}_C$ z lemmatu 7 lze v našem případě přeformulovat jako platnost tvrzení[4], že je výraz

$$e_{ab} \cdot \left(1 + \sum_{D \subseteq C, |D|=2} e_D \prod_{d \in D} x_d\right) - \left(\sum_{c \in C} e_{ac} x_c\right) \cdot \left(\sum_{c \in C} e_{bc} x_c\right) \tag{3.9}$$

roven nule pro každé $\boldsymbol{x}_C \in \{-1, 1\}^C$.

i) Pro $a, b \in \{1, 2, \dots, n\}$ je $e_{ab} \neq 0$, tudíž zjevně neplatí $\xi_a \perp\!\!\!\perp \xi_b$.

ii) Pro $a, b \in \{2, 3, \dots, n\}$ a množinu C, $\emptyset \neq C \subseteq \{1, \dots, n\} \setminus ab$, dostáváme dosazením do vztahu (3.9) pro $\xi_a \perp\!\!\!\perp \xi_b|\boldsymbol{\xi}_C$

$$\epsilon(1 + \epsilon(\dots)) - \epsilon^2(\dots), \tag{3.10}$$

kde „…“ je zástupný symbol pro polynom v ϵ. Výraz (3.10) je řádově ϵ, neboť krom jediného jsou všechny členy řádu ϵ^2 či vyššího a tudíž pro ϵ dostatečně malé je zjevně nenulový. Tudíž ani tvrzení $\xi_a \perp\!\!\!\perp \xi_b|\boldsymbol{\xi}_C$ nemůže býti platným.

[4]Povšimněme si, že se na rozdíl od Gaussovského případu jedná o větší počet rovnic.

iii) Pro $a \in \{2, 3, \dots n\}$ a neprázdnou množinu $C = \{c_1, c_2, \dots, c_k\} \subseteq \{2, 3, \dots, a-1\}$, nabývá výraz (3.9) pro $\xi_a \perp\!\!\!\perp \xi_1 | \boldsymbol{\xi}_C$ hodnoty

$$\epsilon^{n-a+1}(1 + \epsilon(\dots)) - \epsilon^{n-a+3}(\dots) \neq 0.$$

Proto ani tvrzení $\xi_a \perp\!\!\!\perp \xi_1 | \boldsymbol{\xi}_C$ nemůže býti (pro ϵ dostatečně malé) platným.

iv) Analogicky, pro $a \in \{2, 3, \dots n\}$ a $C = \{c_1, \dots, c_k\} \subseteq \{2, 3, \dots, n\} \setminus a$ takovou, že $m := \max_{c \in C} c > a + 1$, nabývá výraz (3.9) pro $\xi_a \perp\!\!\!\perp \xi_1 | \boldsymbol{\xi}_C$ a $\boldsymbol{x}_C = \{1\}^C$ nenulové hodnoty

$$\epsilon^{n-a+1}(\dots) - |C|\epsilon^{n-m+2} - \epsilon^{n-m+3}(\dots) \neq 0.$$

Proto ani v tomto případě tvrzení $\xi_a \perp\!\!\!\perp \xi_1 | \boldsymbol{\xi}_C$ nemůže býti platným.

v) Zbývá vyšetřit případ $a \in \{2, 3, \dots n-1\}$ a $C = \{(a+1), c_2, \dots, c_k\}$, $\emptyset \neq C \subseteq (a+1) \cup \{2, 3, \dots, a-1\}$. Pro C dvou či víceprvkovou a $\boldsymbol{x}_C = \{1\}^C$ je výraz (3.9) roven

$$\epsilon^{n-a+1}(1+\epsilon(\dots)) - |C|\epsilon^{n-a+1} - \epsilon^{n-a+2}(\dots) = (1-|C|)\epsilon^{n-a+1} + \epsilon^{n-a+2}(\dots),$$

což (pro ϵ dostatečně malé) je nenulové a tudíž ani tvrzení $\xi_1 \perp\!\!\!\perp \xi_a | \boldsymbol{\xi}_C$ neplatí.

Avšak pokud je naopak $C = \{a+1\}$ jednoprvková, potom pro libovolné $x_{(a+1)} \in \{-1, 1\}$ platí

$$1\epsilon^{n-a+1} - x_{(a+1)}^2 \epsilon\epsilon^{n-a} = 0.$$

Skutečně tedy (dle lemmatu 7) platí $\xi_1 \perp\!\!\!\perp \xi_a | \xi_{a+1}$.

Obdobně lze ukázat, že pokud definujeme binární rozdělení $\boldsymbol{\xi}^*$ jako

$$\begin{aligned} \forall a: \; e_a^* &= 0, \\ e_{12}^* &= 0, \\ \forall a, b, a \neq b, a + b > 3: \; e_{ab}^* &= \epsilon, \\ \forall A \subseteq N, |A| > 2: \; e_A^* &= 0, \end{aligned}$$

dostaneme kladnou binární reprezentaci I_n^* a pro $\boldsymbol{\xi}^{**}$

$$\begin{aligned} \forall a: \; e_a^* &= 0, \\ \forall a, b, a \neq b: \; e_{ab}^* &= \epsilon, \\ \forall A \subseteq N, |A| > 2: \; e_A^* &= 0, \end{aligned}$$

kladnou binární reprezentaci I_n^{**}. $\square$

S využitím lemmatu 9 dostáváme jako důsledek lemmat 25 a 26 následující závěrečné tvrzení:

Věta 5. *Třídy (obecně) diskrétně, kladně diskrétně, (obecně) binárně, ani kladně binárně reprezentovatelných nezávislostních modelů nejsou konečně charakterizovatelné.*

3.5 Závěr kapitoly

V této kapitole jsme diskutovali možná rozšíření známého výsledku o diskrétní reprezentovatelnosti nad čtyřmi veličinami na pozitivní, případně binární pozitivní reprezentovatelnost. Uvedli jsme si několik charakteristik binárně pozitivně reprezentovatelných modelů a vyslovili hypotézu o pozitivně diskrétně reprezentovatelných modelech nad $N = \{1, 2, 3, 4\}$. Pro její ověření nebo vyvrácení je třeba rozhodnout o d^+–reprezentovatelnosti modelů P2, P3, P19, P24, P30, P32, P39, P37, P41, P45, P46 a P57 z obr. 3.3.

Připomeňme otevřený problém zmíněný již v závěru 2. kapitoly a to vztah tříd g^+–reprezentovatelných a b^+–reprezentovatelných modelů. Z 53 g^+–reprezentovatelných typů byla nalezena b^+–reprezentace pro 51 z nich, nerozhodnuté zůstávají M25 a M37 z obr. 2.3. Binární reprezentaci modelu lze z Gaussovské odvodit přímo, jestliže jsou prvky varianční matice mimo diagonálu blízko nule, na základě poznámky na konci pododdílu 1.2.3. Takovou g^+–reprezentaci však nemáme pro všechny typy k dispozici.

Kapitola 4

Grafické nezávislostní modely

V této kapitole se budeme věnovat odhadům v nezávislostních modelech. Každý nezávislostní model lze totiž chápat při daném distribučním rámci jako „statistický model“, tj. určitou rodinu pravděpodobnostních rozdělení. Určení tohoto statistického modelu je proces skládající se ze tří částí: identifikace samotného nezávislostního modelu, parametrizaci modelu vzhledem k omezením na podmíněnou nezávislost a odhadu těchto parametrů.

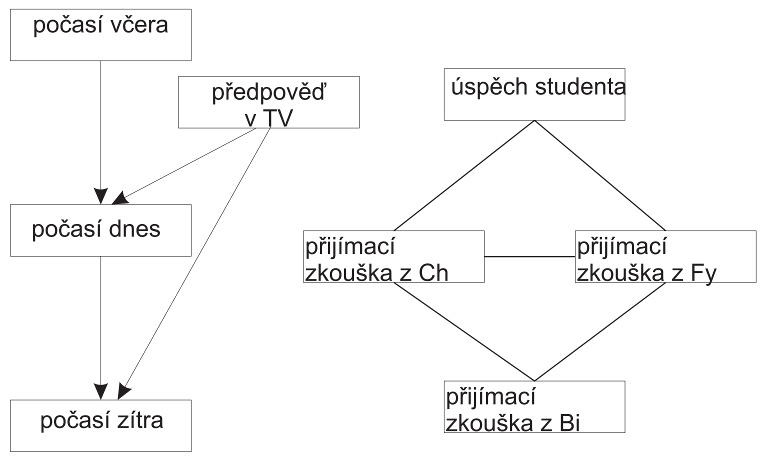
V první oddíle se budeme věnovat teorii tzv. „grafických modelů“ s důrazem zvláště na třídu tzv. rozložitelných modelů (tj. těch modelů, jejichž nezávislostní strukturu lze reprezentovat jak neorientovaným, tak orientovaným acyklickým grafem), která má v praxi nejvíce aplikací. V druhé části se bude zabývat statistickými modely pro nezávislostní, zvláště pak grafické modely.

A konečně poslední oddíl bude věnován hledání nezávislostního modelu na základě dat. Za pomoci simulace se pokusíme zodpovědět otázku, jak velké chyby se dopouštíme, pokud se omezujeme ze všech g^+–reprezentovatelných nezávislostních modelů pouze na modely grafické.

Na začátek si neodpustím připojit několik autorových poznámek k dané problematice. Pro nově příchozího do této oblasti může být tato diskuze obtížně stravitelná a nezbývá než mu doporučit přejít přímo na oddíl 4.1 (a případně se k ní později vrátit).

Grafické nezávislostní modely[1] jsou specifickou podtřídou nezávilostních modelů, kde seznam tripletů příslušných modelu lze získat za pomoci něja-

[1] V literatuře se obvykle hovoří o „grafických modelech“ a pod tímto pojmem je myšlena jak nezávislostní struktura, tak i přímo pravděpodobnostní rozdělení této struktuře odpovídající. Tedy celý statistický a nikoli pouze nezávislostní model.



Obrázek 4.1: Dvě praktické ukázky aplikace grafických modelů.

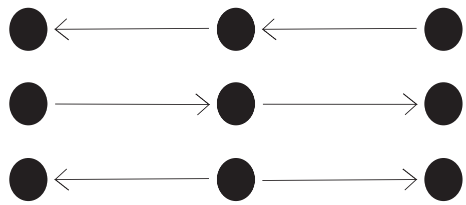
kého (separačního) kritéria z neorientovaného či orientovaného acyklického grafu.

Příklad 1. *Dva intuitivně pochopitelné příklady lze nalézt na obr. 4.1. Mezi počasím včera a zítra zjevně existuje jistá korelace, která však vymizí, jestliže již známe dnešní počasí. Analogicky, přestože je úspěch studenta pozitivně korelován s body z přijímacích zkoušek z biologie, neznamená to ještě nutně, že je kvalita tohoto testu dostačující. Pokud nezávislostní schéma vypadá jako na obr. 4.1, je za předpokladu znalosti počtu bodů z chemie a fyziky již znalost výsledku přijímaček z biologie pro predikci úspěchu studenta neužitečná. Dané dvě veličiny jsou totiž podmíněně nezávislé, neboli – vezmeme-li úspěch jako závislou veličinu – je regresní[2] koeficient příslušný bodům z biologie nulový. Studenti by tedy měli být přijímáni pouze na základě váženého průměru bodů z chemie a fyziky, biologická část by se neměla započítávat (nebo by se měl způsob testování změnit). Jedná se mimochodem o reálnou studii na 1. LF UK, viz [ŠŠ06].*

Použití grafů má jak interpretační výhodu, totiž že neorientované hrany odpovídají intuitivně pochopitelnému vztahu „míti souvislost“, orientované „býti příčinou“, tak je výhodný i z teoretického hlediska. Obecně totiž není možné u zvoleného nezávislostního modelu identifikovat parametry pravděpodobnostního rozdělení (v daném distribučním rámci), jež maximalizuje

[2]Protože úspěch studenta byl měřen jako schopnost dokončit studium (binární veličina), jednalo by se v našem případě spíše o logistickou regresi.

pro daná data věrohodnostní funkci. Taktéž operace marginalizace či podmínění jsou v grafických modelech obvykle snáze proveditelné, viz [CLDS99], [Nea03], [CGH96], [Bou95], [Cas02], [Drt04] a [Koč01].



Obrázek 4.2: Tři DAGy reprezentující tentýž nezávislostní model.

Zmiňme také alespoň tři nejčastější problémy spojené s analýzou dat za pomoci grafických modelů. Za prvé, v případě orientovaného grafu často nedokážeme rozlišit směr „kauzality“. Tento problém je principiální a nelze jej vyřešit, pokud data pocházejí z jediného časového okamžiku: grafům na obr. 4.2 odpovídá totiž tentýž nazávislostní model. Z praktického hlediska nás ale „směr kauzality“ nesmírně zajímá, neboť odpovídá na otázku, co se stane, když nějakou proměnnou zásahem zvenčí pevně nastavíme na danou hodnotu, viz [Lau01].

Dalším problémem je obrovské množství grafických nezávislostních modelů a marná snaha o identifikaci jednoho „správného“ modelu při rostoucím množství ($\approx$ desítky) náhodných veličin a reálném množství dat ($\approx$ tisíce pozorování). Přestože se teoreticky pravděpodobnost výběru správného modelu asymptoticky blíží jedné, v reálných případech se obvykle věrohodnost několika nejlepších nezávislostních modelů liší pouze nepatrně, přestože jejich příčinná interpretace může být zcela jiná. Proto je velice užitečné ubezpečit se proto o důvěryhodnosti dosažených výsledků za pomoci simulační studie, viz např. [Šim04].

V neposlední řadě je třeba mít na paměti, že grafických je pouze nepatrný díl ze všech nezávislostních modelů. Přesněji řečeno, logaritmus počtu grafických nezávislostních modelů roste polynomiálně v závislosti na počtu náhodných veličin, zatímco logaritmus počtu všech nezávislostních modelů roste exponenciálně (detaily v [Stu05]).

Existují i jiné způsoby reprezentace nezávislostního modelu, v prvé řadě tzv. řetězcové grafy (viz [Lau96]), používající v grafu zároveň orientované i neorientované hrany, a různá jejich rozšíření o další typy hran, viz např. [RS02] a [CW96]. Jinou možnou volbou jsou tzv. „imsety“, nezávislostní model reprezentující celočíselné vektory, viz [Stu05], s řadou zajímavých teoretických problémů (viz např. [ŠS04]).

4.1 Teorie grafů

4.1.1 Neorientované grafy

Neorientovaným grafem (zkráceně[3] UG) $G = (N, E)$ budeme rozumět konečnou množinu vrcholů N a množinu neorientovaných hran E,

$$E \subseteq \binom{N}{2} = \{ab;\ a, b \in N, a \neq b\}.$$

Řekneme, že graf (N', E') je **podgrafem** (N, E), pokud $N' \subseteq N$ a $E' \subseteq \binom{N'}{2} \cap E$. Jestliže $E' = \binom{N'}{2} \cap E$, budeme mluvit o **indukovaném** podgrafu (značíme $G_{N'}$). Graf je **úplný**, pokud $E = \binom{N}{2}$. Podmnožinu množiny vrcholů $A \subseteq N$ nazveme úplnou, jestliže indukuje úplný podgraf. Množinu $A \subseteq N$ nazveme **klikou**, pokud je úplná a maximální vzhledem k relaci „$\subseteq$“ (neexistuje úplná vlastní nadmnožina A).

Cesta mezi vrcholy a a b je posloupnost $k \geq 1$ vrcholů $a = c_1, \ldots, c_k = b$ taková, že $c_i c_{i+1} \in E$, $i = 1, 2, \ldots, k-1$. Vrcholy a a b nazveme **spojené** (značíme $a \sim b$) pokud existuje cesta je spojující. Množiny vrcholů A a B jsou spojené, pokud existují vrcholy $a \in A$ a $b \in B$, které jsou spojené. Povšimněme si, že relace spojitosti mezi vrcholy grafu je zjevně ekvivalencí, její třídy budeme nazývat **komponentami** grafu.

Množinou sousedů množiny $A \subseteq N$ rozumíme

$$ne(A) = \{b;\ a \in A, b \in N \setminus A, ab \in E\}.$$

Nechť A, B, C jsou po dvou disjunktní podmnožiny N. Řekneme, že C **odděluje** A a B (značíme $A \perp B | C$), jestliže každá cesta mezi A a B obsahuje vrchol z C neboli v grafu $G_{N \setminus C}$ nejsou A a B spojené. Pro $a, b \in N$ nazveme množinu $C \subseteq N$ minimálním ab–**separátorem**, pokud $a \perp b | C$ v grafu G a je minimální taková vzhledem k relaci „$\subseteq$“ (neexistuje vlastní podmnožina C oddělující a a b).

Nezávislostní model odvozený **globálním** separačním kritériem z neorientovanému grafu G je nezávislostní model nad N definovaný následovně

$$\mathcal{I}^g_{UG}(G) = \{\langle ab|C \rangle;\ a \perp b | C \text{ v grafu } G\}.$$

[3] z anglického „undirected graph“

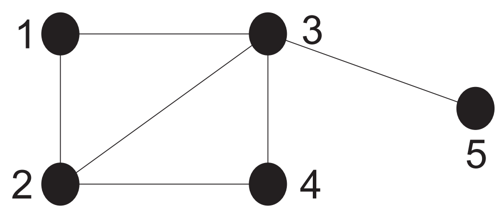
Dále definujeme nezávislostní modely odvozené z **lokálního** a **párového** kritéria

$$\begin{aligned}\mathcal{I}^l_{UG}(G) &= \{\langle ab|C\rangle;\ ne(a) \subseteq C, ab \notin E\}\,,\\ \mathcal{I}^p_{UG}(G) &= \{\langle ab|N \setminus ab\rangle;\ ab \notin E\}\,.\end{aligned}$$

Povšimněme si, že pro libovolný graf G zjevně platí

$$\mathcal{I}^p_{UG}(G) \subseteq \mathcal{I}^l_{UG}(G) \subseteq \mathcal{I}^g_{UG}(G). \tag{4.1}$$

Rozbor, kdy jsou inkluze v 4.1 ostré, lze nalézt v [Mat92].



Obrázek 4.3: Neorientovaný graf G.

Příklad 2. *Ilustrujme zavedené pojmy na grafu G z obr. 4.3. Povšimněme si, že hrana mezi dvěmi vrcholy je znázorněna čárou tyto vrcholy spojující. G_{123} je úplný a $\{1,2,3\}$ je klika G. Ovšem $\{1,2\}$ klikou G není, přestože G_{12} je úplný. Libovolné dva vrcholy grafu jsou spojené cestou. Množina sousedů vrcholu 4 je $ne(4) = \{2,3\}$. Platí $1 \perp 4|23$ a $\{2,3\}$ je minimální 14–separátor. Taktéž platí $1 \perp 4|235$, ovšem $\{2,3,5\}$ není minimálním ab–separátorem pro žádnou volbu a, b.*

4.1.2 Orientované acyklické grafy

V následujících dvou oddílech je mnoho důkazů vynecháno či nahrazeno odkazem s poznámkou „snadné cvičení“, aby čtenář nebyl ošizen o radost dokázat si daná tvrzení sám.

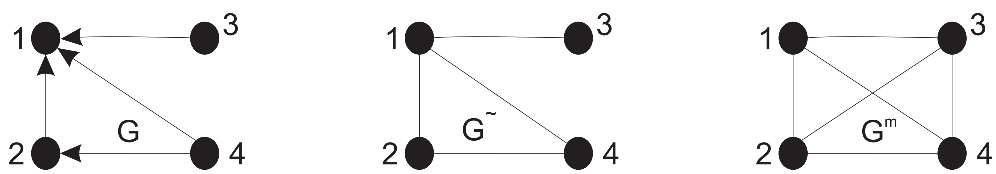
Orientovaným acyklickým grafem (zkráceně[4] DAG) $G = (N, E)$ budeme rozumět konečnou množinu vrcholů N a množinu orientovaných hran E,

$$E \subseteq \{(a,b);\ a, b \in N, a \neq b\}$$

[4]z anglického „directed acyclic graph“

splňující podmínku, že v grafu neexistuje **orientovaný cyklus**, t.j. posloupnost $k \geq 2$ vrcholů $c_1, \ldots, c_k$ taková, že $(c_i, c_{i+1}) \in E$, $i = 1, 2, \ldots, k-1$ a navíc $(c_k, c_1) \in E$. Podgraf a indukovaný podgraf definujeme analogicky jako v případě neorientovaného grafu. Kostrou orientovaného grafu $G = (N, E)$ rozumíme neorientovaný graf

$$G^{\sim} = (N, \{ab; (a,b) \in E \ \vee \ (b,a) \in E\}).$$



Obrázek 4.4: Graf G, jeho kostra $G^{\sim}$ a zmoralizovaný graf G^m.

Orientovanou cestou mezi vrcholy a a b (značíme $a \rightsquigarrow b$) nazveme posloupnost $k \geq 2$ vrcholů $a = c_1, \ldots, c_k = b$ takovou, že $(c_i, c_{i+1}) \in E$, $i = 1, 2, \ldots, k-1$. Množinou rodičů, dětí, předků a následníků[5] množiny $A \subseteq N$ rozumíme po řadě množiny

$$\begin{aligned}
pa(A) &= \{b;\ a \in A, b \in N \setminus A, (b,a) \in E\}, \\
ch(A) &= \{b;\ a \in A, b \in N \setminus A, (a,b) \in E\}, \\
an(A) &= \{b;\ a \in A, b \in N \setminus A, b \rightsquigarrow a\}, \\
de(A) &= \{b;\ a \in A, b \in N \setminus A, a \rightsquigarrow b\}.
\end{aligned}$$

Nechť A, B, C jsou po dvou disjunktní podmnožiny N. Řekneme, že C **odděluje** A a B (značíme $A \perp B | C$) v acyklickém orientovaném grafu G, jestliže mezi A a B není hrana ($\forall a \in A, b \in B : (a,b) \notin E, (b,a) \notin E$) a pro každou (neorientovanou) cestu v $G^{\sim}$ o $k > 2$ vrcholech $A \ni a = d_1, \ldots, d_k = b \in B$ existuje vrchol d_i, jež není na kraji ($1 < i < k$), pro nějž nastává jedna z následujících dvou možností

- $(d_{i-1}, d_i) \in E \ \wedge \ (d_{i+1}, d_i) \in E$ a zároveň $\big(d_i \cup de(d_i)\big) \cap C = \emptyset$
- $\big[(d_{i-1}, d_i) \notin E \ \vee \ (d_{i+1}, d_i) \notin E\big]$ a zároveň $d_i \in C$.

[5]Většina autorů užívá mírně odlišnou definici, kdy vrchol je sám sobě předkem a následníkem.

Existuje ekvivalentní způsob, jak definovat oddělitelnost v orientovaném grafu $G = (N, E)$, a to za pomoci **moralizačního grafu**[6], tj. neorientovaného grafu

$$G^m = \big(N, \{ab;\ a, b \in N, a \neq b, (ab \text{ je hrana v } G^{\sim}) \ \vee \ (ch(a) \cap ch(b) \neq \emptyset)\}\big).$$

Příklad lze nalézt na obr. 4.4.

Lemma 27. *Nechť $G = (N, E)$ je DAG a A, B, C jsou po dvou disjunktní podmnožiny N. Potom $A \perp B | C$ vzhledem ke G právě tehdy, když $A \perp B | C$ vzhledem ke $(G_{ABC \cup an(ABC)})^m$.*

Důkaz. [Lau96], str. 48–49 (nepříliš těžké cvičení) □

Nezávislostní model příslušný orientovanému acyklickému grafu G definujeme analogicky jako pro UG, tzn. jako nezávislostní model nad N

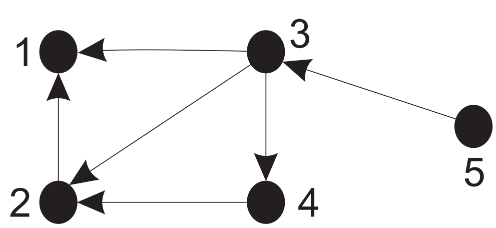
$$\mathcal{I}_{DAG}(G) = \{\langle ab|C\rangle;\ a \perp b | C \text{ v grafu } G\}\,.$$

Jak ukazuje obr. 4.2, různým orientovaným acyklickým grafům může být příslušný tentýž nezávislostní model. Následující lemma (převzaté z [CGH96], věta 6.16) charakterizuje takové DAGy.

Lemma 28. *Nechť $G_1 = (N, E_1)$ a $G_2 = (N, E_2)$ jsou DAGy. Potom platí*

$$\mathcal{I}_{DAG}(G_1) = \mathcal{I}_{DAG}(G_2)$$

právě tehdy, když kostry grafů jsou totožné ${G_1}^{\sim} = {G_2}^{\sim}$ a v obou grafech se na týchž místech vyskytují tzv. v–konfigurace, tj. indukované podgrafy typu $i \to j \leftarrow k$.



Obrázek 4.5: Orientovaný acyklický graf G.

[6] „Moralizace“ znamená, že se „hranou sezdají“ vrcholy, jež mají společné dítě.

Příklad 3. *Ilustrujme zavedené pojmy na grafu G z obr. 4.5. Hranu (a, b) je znázorňujeme jako šipku z vrcholu a do vrcholu b. Graf je acyklický, k porušení tohoto předpokladu by došlo například přidáním hrany $(4, 5)$. Platí $pa(5) = \emptyset$, $ch(3) = \{1, 2\}$, $an(\{1, 2\}) = \{3, 4, 5\}$, $de(4) = \{1, 2\}$.*

Nezávislostní model příslušný grafu G je $\mathcal{I}_{DAG}(G) = \big\{\langle 14|23\rangle, \langle 14|235\rangle, \langle 15|3\rangle, \langle 15|34\rangle, \langle 15|23\rangle, \langle 15|234\rangle, \langle 25|3\rangle, \langle 25|34\rangle, \langle 25|13\rangle, \langle 25|134\rangle, \langle 45|3\rangle, \langle 45|13\rangle, \langle 45|23\rangle, \langle 45|123\rangle\big\}$.

Povšimněme si, že tentýž nezávislostní model lze odvodit i z neorientovaného grafu na obr. 4.3.

4.1.3 Rozložitelné grafy

Rozložitelný graf je speciální typ UG s množstvím překvapivých vlastností. Uvedeme rekurzivní definici podle počtu vrcholů grafu:

Neorientovaný graf $G = (N, E)$ nazveme **rozložitelný**, jestliže je buďto úplný nebo existuje rozklad množiny vrcholů $N = ABC$ na tři po dvou disjunktní množiny $A \neq \emptyset$, $B \neq \emptyset$, C a platí, že $A \perp B | C$ v grafu G, G_C je úplný a G_{AC}, G_{BC} jsou rozložitelné.

Lemma 29. *Nechť $G = (N, E)$ je neorientovaný graf. Potom jsou následující tvrzení ekvivalentní.*

i) Graf G je rozložitelný.

*ii) Graf G je **chordální**, tzn. neobsahuje jako indukovaný podgraf kružnici*

$$C_n = \big(\{c_1, \ldots, c_n\},\ \{c_1c_2, c_2c_3, \ldots, c_{n-1}c_n, c_nc_1\}\big)$$

o $n \geq 4$ vrcholech.

iii) Všechny minimální ab–separátory jsou úplné.

*iv) Existuje uspořádaní klik grafu $C_1, \ldots, C_k$ takové, že tvoří **perfektní posloupnost**, tj. platí*

$$\forall i, 1 < i \leq k\ \exists j < i :\ C_i \cap \bigcup_{k=1}^{i-1} C_k \subseteq C_j. \tag{4.2}$$

v) Existuje uspořádání vrcholů grafu $N = \{c_1, \ldots, c_n\}$ takové, že

$$\big(c_i \cup ne(c_i)\big) \cap \{c_1, \ldots, c_i\}, \quad i = 1, \ldots, n \tag{4.3}$$

tvoří posloupnost úplných množin.

Důkaz. [Lau96], $(i) \Leftrightarrow (ii) \Leftrightarrow (iii)$, str. 9, $(i) \Leftrightarrow (iv) \Leftrightarrow (v)$, str. 19. $\square$

Příklad rozložitelného grafu lze nalézt na obr. 4.3. Algoritmy, jak nalézt uspořádání klik (4.2) a vrcholů (4.3), lze nalézt např. v dodatku k [Šim03].

Nezávislostní modely odvoditelné z rozložitelných grafů jsou právě ty, jež lze odvodit jak z UG, tak z DAG.

Lemma 30. *Nechť I je nezávislostní model nad N. Potom jsou následující dvě tvrzení ekvivalentní*

i) Existuje rozložitelný graf G takový, že $I = \mathcal{I}^g_{UG}(G)$.

ii) Existují UG G' a DAG G'' takové, že $I = \mathcal{I}^g_{UG}(G') = \mathcal{I}_{DAG}(G'')$.

Důkaz. Pro přehled provázanosti těchto i jiných tříd modelů, viz [Cas02], str. 36, nebo [CGH96], str. 261. $\square$

K druhému tvrzení z předchozího lemmatu ještě poznamenejme, že příslušný UG G' a DAG G'' jsou navzájem provázané následujícím způsobem: $G' = (G'')^{\sim} = (G'')^m$ a naopak G'' získáme z G', pokud zorientujeme jeho hrany ve směru uspořádání (4.3) z lemmatu 29.

4.2 Pravděpodobnostní rozdělení nad grafickými nezávislostními modely

Definice 12. *Nechť I je nezávislostní model. Potom rodinu takových náhodných rozdělení vektoru $\boldsymbol{\xi}$, že platí $I \subseteq \mathcal{I}(\boldsymbol{\xi})$, budeme značit $\mathcal{P}(I)$.*

Obvykle nás budou zajímat rozdělení v daném distribučním rámci, který bude naznačen příslušným dolním indexem. Např. $\mathcal{P}_g(\{\langle 12|\emptyset\rangle\})$ je třída všech Gaussovských rozdělení náhodného vektoru $\boldsymbol{\xi}$ takových, že $\xi_1 \perp\!\!\!\perp \xi_2$ neboli prvek varianční matice $\sigma_{1\cdot 2}$ je nulový.

Pro pozitivní reprezentace rodiny rozdělení odvozené z UG párovým, lokálním a globálním kritériem splývají. Z tohoto důvodu budeme později v textu indexy „gl“, „loc“ a „par“ u pozitivních rodin rozdělení vynechávat.

Lemma 31. *Označíme-li $\mathcal{P}_+(I)$ průnik $\mathcal{P}(I)$ s třídou všech pozitivních rozdělení, potom pro libovolný neorientovaný graf G platí*

$$\mathcal{P}_+(\mathcal{I}^{par}_{UG}(G)) = \mathcal{P}_+(\mathcal{I}^{loc}_{UG}(G)) = \mathcal{P}_+(\mathcal{I}^{gl}_{UG}(G)).$$

Důkaz. Cvičení na použití vlastnosti (1.10) z lemmatu 8, viz [Lau96], str. 34. □

Pro acyklický orientovaný graf G lze za pomoci moralizačního kritéria nahlédnout, že pro všechna rozdělení z $\mathcal{P}(\mathcal{I}_{DAG}(G))$ platí, že jejich hustotu lze za pomoci lemmatu 1 faktorizovat následujícím způsobem

$$f(\boldsymbol{x}) = \prod_{a \in N} f_{a|pa(a)}(x_a|\boldsymbol{x}_{pa(a)}) = \prod_{a \in N} \frac{f_{a \cup pa(a)}(\boldsymbol{x}_{a \cup pa(a)})}{f_{pa(a)}(\boldsymbol{x}_{pa(a)})}$$

za konvence $\frac{0}{0} = 0$.

4.2.1 Faktorizace pravděpodobnostních rozdělení nad rozložitelnými grafy

Nyní se podrobněji podíváme, jak tyto rodiny rozdělení vypadají, pokud je graf G rozložitelný.

Lemma 32. *Nechť G je rozložitelný graf a $C_1, C_2, \ldots C_m$ perfektně uspořádaná posloupnost jeho klik. Označíme-li*[7] *dále*

$$S_i = C_i \setminus \bigcup_{j=1}^{i-1} C_j, \qquad i = 2, \ldots, m, \tag{4.4}$$

potom pro hustotu rozdělení z $\mathcal{P}(\mathcal{I}_{DAG}(G))$ platí

$$f(\boldsymbol{x}) = \frac{\prod_{i=1}^m f_{C_i}(\boldsymbol{x}_{C_i})}{\prod_{i=2}^m f_{S_i}(\boldsymbol{x}_{S_i})} \tag{4.5}$$

za konvence $\frac{0}{0} = 0$. Konkrétně pro diskrétní rozdělení můžeme (4.5) přepsat na

$$P(\boldsymbol{\xi} = \boldsymbol{x}) = \frac{\prod_{i=1}^m P(\boldsymbol{\xi}_{C_i} = \boldsymbol{x}_{C_i})}{\prod_{i=2}^m P(\boldsymbol{\xi}_{S_i} = \boldsymbol{x}_{S_i})} \tag{4.6}$$

a pro regulární Gaussovské rozdělení s varianční maticí $\boldsymbol{\Sigma}$ platí

$$\boldsymbol{\Sigma}^{-1} = \sum_{i=1}^m [(\boldsymbol{\Sigma}_{C_i \cdot C_i})^{-1}]_N - \sum_{i=2}^m [(\boldsymbol{\Sigma}_{S_i \cdot S_i})^{-1}]_N, \tag{4.7}$$

[7] Jak dokázal F. Matúš, systém $\{S_i,\ i = 2, \ldots, m\}$ je zároveň systémem všech minimálních ab–separátorů grafu G, více viz definice ab–separátoru a jeho indexu v [Šim05]. Tento fakt však nebude v této práci dále využit.

kde operací $[\cdot]_N$ myslíme zvětšení marginální matice na původní rozměr, přičemž prvky marginální matice jsou umístěny na odpovídající místa a na ta zbylá jsou doplněny nuly.

Důkaz. Snadné cvičení (opakované použití lemmatu 1 postupně odzadu na seznam klik), viz [Lau96], str. 90 a str. 145. $\square$

Jako ilustraci použití perfektního uspořádání klik a lemmatu 32 uvažme následující problém z [Šim05]: Mějme rozložitelný graf $G = (N, E)$ a označme $\mathcal{P}_b^p(\mathcal{I}_{UG}(G))$ všechna pravděpodobnostní rozdělení z $\mathcal{P}_b(\mathcal{I}_{UG}(G))$, jež navíc splňují

$$\max_{a \in N} P(\xi_a = 1) \le p,$$

kde $p \in [0, 1]$.

Našim cílem bude pro dané p minimalizovat výraz

$$P(\bigcap_{a \in N} \xi_a = -1) \tag{4.8}$$

přes všechna rozdělení z $\mathcal{P}_b^p(\mathcal{I}_{UG}^{gl}(G))$.

Lemma 33. *Nechť G je rozložitelný graf a $p \in [0, 1]$. Potom*

$$\min_{P \in \mathcal{P}_b^p(\mathcal{I}_{UG}^{gl}(G))} P(\bigcap_{a \in N} \xi_a = -1) = \begin{cases} \frac{\prod_{a=1}^m (1 - |C_a| p)}{\prod_{a=2}^m (1 - |S_a| p)}, & p \le \frac{1}{\max_{a=1,\ldots,m} |C_a|} \\ 0 & p > \frac{1}{\max_{a=1,\ldots,m} |C_a|} \end{cases}$$

Důkaz. Uspořádejme kliky G do perfektní posloupnosti $C_1, \ldots, C_m$ a definujme $S_2, \ldots, S_m$ podle vztahu (4.4). Jestliže $p \le \frac{1}{\max_{a=1,\ldots,m} |C_a|}$, potom můžeme definovat pravděpodobnostní rozdělení P^* pro každou kliku C_a jako

$$\forall b \in C_a : \qquad \begin{array}{l} P^*(\bigcap_{c \in C_a} \xi_c = -1) = 1 - |C_a| p \\ P^*(\xi_b = 1, \bigcap_{c \in C_a \setminus b} \xi_c = -1) = p \\ P^*(\ldots) = 0 \quad \textit{jinak} \end{array}$$

a rozdělení P^* je dále jednoznačně určeno lemmatem 32.

Zjevně $P^* \in \mathcal{P}_b^p(\mathcal{I}_{UG}^{gl}(G))$, stačí tedy dokázat

$$\forall P \in \mathcal{P}_b^p(\mathcal{I}_{UG}^{gl}(G)) : \ P(\bigcap_{a \in N} \xi_a = -1) \ge \frac{\prod_{a=1}^m (1 - |C_a| p)}{\prod_{a=2}^m (1 - |S_a| p)}.$$

To lze provést indukcí na počtu klik m. Pro jednu kliku je výsledek zřejmý. Jinak označme $T = S_m = C_m \cap (\cup_{a=1,\dots,m-1} C_a)$ a z (4.2) dostáváme, že existuje $a \in \{1,\dots,m-1\}$ takové, že $T \subset C_a$. Graf $G' = G_{(N\setminus C_m)\cup T}$ má oproti grafu G o jednu kliku méně, můžeme tedy použít indukční předpoklad pro pravděpodobnost $P(\bigcap_{a\in N}\xi_a = -1)$ zapsanou jako součin

$$\frac{P(\bigcap_{a\in T}\xi_a=-1,\ \bigcap_{a\in C_m\setminus T}\xi_a=-1)}{P(\bigcap_{a\in T}\xi_a=-1)} \cdot P(\bigcap_{a\in(N\setminus C_m)\cup T}\xi_a = -1) \geq$$
$$\geq \frac{1-\hat{p}-(|C_m|-|T|)p}{1-\hat{p}} \cdot \frac{\prod_{a=1}^{m-1}(1-|C_a|p)}{\prod_{a=2}^{m-1}(1-|S_a|p)}, \tag{4.9}$$

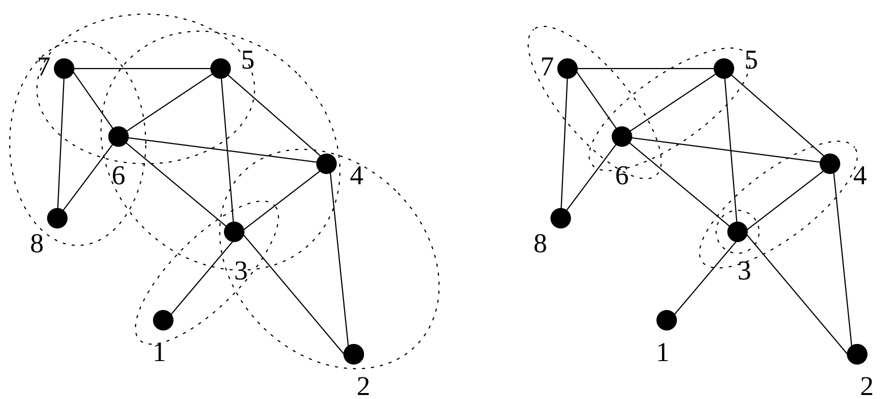
kde definujeme $\hat{p} \equiv 1 - P(\bigcap_{a\in T}\xi_a = -1) \leq |T|\cdot p$.

Zlomek v na levé straně (4.9) nabývá minima pro $\hat{p} = |T|.p$ a tudíž

$$P(\bigcap_{a\in N}\xi_a = -1) \geq \frac{1-|C_m|p}{1-|T|p} \cdot \frac{\prod_{a=1}^{m-1}(1-|C_a|p)}{\prod_{a=2}^{m-1}(1-|S_a|p)} = \frac{\prod_{a=1}^{m}(1-|C_a|p)}{\prod_{a=2}^{m}(1-|S_a|p)},$$

čímž je indukčních krok dokončen.

Zbývá rozebrat případy $p > \frac{1}{\max_{a\in N}|C_a|}$. Stačí si však uvědomit, že minimum (4.8) je zjevně nerostoucí funkce p a již pro $p = \frac{1}{\max_{a\in N}|C_a|}$ nabývá nuly. □



Obrázek 4.6: Graf G, jeho kliky C_a (vlevo) a separátory S_a (vpravo)

Příklad 4. *Nechť G je rozložitelný graf znázorněný na obr. 4.6. Jeho kliky jsou $\{1,3\}$,$\{2,3,4\}$,$\{3,4,5,6\}$,$\{5,6,7\}$,$\{6,7,8\}$. Tudíž dle lemmatu 33,*

$$\min_{P\in\mathcal{P}_b^p(\mathcal{I}_{UG}^{gl}(G))} P(\bigcap_{a\in N}\xi_a = -1) = \begin{cases} \frac{(1-2p)(1-3p)^3(1-4p)}{(1-p)(1-2p)^3} = \frac{(1-3p)^3(1-4p)}{(1-p)(1-2p)^2} & \text{pokud } p \leq \frac{1}{4}, \\ 0 & \text{pokud } p \geq \frac{1}{4}. \end{cases}$$

4.2.2 Maximálně věrohodné odhady

V tomto oddíle vysvětlíme, v čem je vyjímečná třída rozložitelných modelů. Umožňuje totiž elegantně vyřešit otázku identifikace parametrů modelu na základě dat.

Metoda **maximální věrohodnosti** (MV) je nejpoužívanějším způsobem, jak odhadnout parametry v dané rodině rozdělení. Za odhad zvolíme takovou hodnotu parametru, která maximalizuje **log–věrohodnostní funkci**. Za předpokladu, že rozdělení náhodné veličiny patří do exponenciální rodiny[8], takovýto odhad skoro jistě existuje a je jednoznačný. Čtenáře, který si potřebuje znalosti MV odhadů a jejich vlastností spolu s teorií exponenciálních rodin doplnit, si dovolujeme odkázat na [And78] a dodatky [Lau96].

Nechť $\boldsymbol{X}$ jsou **data** o k řádcích neboli matice, kde každé z k nezávislých pozorování náhodného vektoru $\boldsymbol{\xi}$, je zapsáno do jednoho řádku.

Jestliže $\boldsymbol{\xi}$ má diskrétní rozdělení, potom MV odhad jeho parametrů bez nezávislostních omezení je při označení hustoty vůči čítací míře $p(\boldsymbol{x}) = P(\boldsymbol{\xi} = \boldsymbol{x})$ roven

$$\hat{p}(\boldsymbol{x}) = \frac{1}{k} \sum_{a=1}^{k} I[\boldsymbol{X}_a = \boldsymbol{x}],$$

kde $I[A] = 1$ pokud jev A nastal a nula naopak, a $\boldsymbol{X}_a$ je a–té měření (řádek a matice X).

Pokud má $\boldsymbol{\xi}$ naopak regulární Gaussovské rozdělení, potom MV odhady jeho parametrů bez nezávislostních omezení vypadají následovně:

- $\hat{\mu}_a = \frac{1}{k} \sum_{b=1}^{k} X_{ba}$,
- $\hat{\sigma}_{a \cdot b} = \frac{1}{k} \sum_{c=1}^{k} (X_{ca} - \hat{\mu}_a)(X_{cb} - \hat{\mu}_b)$.

Níže zmíníme, jak situace vypadá v případě regulárního Gaussovského rozdělení nad UG a diskrétního (resp. binárního) či Gaussovského rozdělení nad rozložitelným grafem.

Není obtížné nahlédnout, že v obou případech se jedná o exponenciální rodiny (viz [Lau96]). Ovšem v případě UG je nám známo pouze iterační řešení, zatímco v druhém případě je znám exaktní vzorec.

[8] a několika dalších technických předpokladů jako že se skutečná hodnota nenachází na hranici prostoru parametrů…

Lemma 34. *Předpokládejme, že $\boldsymbol{\xi}$ je regulárně Gaussovsky rozdělený vektor se střední hodnotou $\boldsymbol{\mu}$ a s varianční maticí $\boldsymbol{\Sigma}$. Za předpokladu, že rozdělení $\boldsymbol{\xi}$ je prvkem rodiny rozdělení $\mathcal{P}_{g^+}(\mathcal{I}_{UG}(G))$, nalezneme maximálně věrohodné odhady $\hat{\boldsymbol{\mu}}$ a $\hat{\boldsymbol{\Sigma}}$ následovně:*

Nechť $\boldsymbol{X}$ jsou data o k řádcích. Potom

- $\hat{\mu}_a = \frac{1}{k}\sum_{b=1}^{k} X_{ba}$,

u inverze varianční matice platí

- $\langle ab|N \setminus ab\rangle \in \mathcal{I}_{UG}(G) \quad \Rightarrow \quad (\widehat{\boldsymbol{\Sigma}}^{-1})_{a\cdot b} = 0$

a zároveň ovšem

- $\hat{\sigma}_{a\cdot a} = \frac{1}{k}\sum_{b=1}^{k}(X_{ba} - \hat{\mu}_a)^2$
- $\langle ab|N \setminus ab\rangle \notin \mathcal{I}_{UG}(G) \quad \Rightarrow \quad \hat{\sigma}_{a\cdot b} = \frac{1}{k}\sum_{c=1}^{k}(X_{ca} - \hat{\mu}_a)(X_{cb} - \hat{\mu}_b)$

Důkaz. Viz [Lau96], str. 133. □

Povšimněme si dvou faktů. U neorientovaného grafu $G = (N, E)$ platí

$$\langle ab|N \setminus ab\rangle \in \mathcal{I}_{UG}(G) \quad \Longleftrightarrow \quad ab \notin E \tag{4.10}$$

a poslední dvě rovnice, jež musí prvky varianční matice MV odhadu splňovat, jsou za požadavku na nulovost některých členů $\boldsymbol{\Sigma}^{-1}$ řešitelné pouze iteračně.

Faktorizace rozdělení nad rozložitelnými grafy spolu s perfektním uspořádáním klik lze použít pro odvození následujícího lemmatu:

Předpokládejme, že G je rozložitelný graf a rozdělení $\boldsymbol{\xi}$ je prvkem třídy $\mathcal{P}(\mathcal{I}_{UG}(G))$. Potom jestliže dokážeme nalézt MV odhady pro rozdělení klik (viz začátek tohoto oddílu), můžeme nalézt MV odhad pro parametry $\boldsymbol{\xi}$ na základě lemmatu 32 (formální důkaz je v [Lau96], str. 84).

Lemma 35. *Nechť G je rozložitelný graf a rozdělení $\boldsymbol{\xi}$ je prvkem třídy $\mathcal{P}_d(\mathcal{I}_{UG}(G))$. Potom při perfektním uspořádáním klik $C_1, \ldots, C_m$, definici $S_2,\ldots,S_m$ podle vztahu (4.4) a označení hustoty vůči čítací míře $p(\boldsymbol{x}) = P(\boldsymbol{\xi} = \boldsymbol{x})$ platí, že*

$$\hat{p}(\boldsymbol{x}) = \frac{\prod_{a=1}^{m} \frac{1}{k}\sum_{b=1}^{k} I[\boldsymbol{X}_{b\cdot C_a} = \boldsymbol{x}_{C_a}]}{\prod_{a=2}^{m} \frac{1}{k}\sum_{b=1}^{k} I[\boldsymbol{X}_{b\cdot S_a} = \boldsymbol{x}_{S_a}]}.$$

Analogicky, pokud rozdělení $\boldsymbol{\xi}$ je prvkem $\mathcal{P}_{g^+}(\mathcal{I}_{UG}(G))$ a $\widehat{\boldsymbol{\Sigma}}_{C_a\cdot C_a}$ jsou MV odhady variančních matic rozdělení klik $\boldsymbol{\xi}_{C_a}$, potom nám vztah (4.7) udává MV odhad pro varianční matici $\boldsymbol{\xi}$.

Důkaz. Viz [Lau96], str. 91 a 145. □

4.3 Hledání modelu

V zásadě existují dva způsoby jak na základě dat vybrat nezávislostní model. V této kapitole si oba stručně představíme a uvedeme použití při výběru grafického modelu. Na závěr se je pokusíme aplikovat na Gaussovsky reprezentovatelné nezávislostní modely nad čtyřmi veličinami.

První způsob je založený na opakovaném testování podmíněné nezávislosti:

1. Pro všechna $\langle ab|C\rangle \in \mathcal{T}_N$ testuj, zda-li $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_C$.
2. Z daného prostoru nezávislostních modelů vyber takový, který odpovídá nalezeným nezávislostem z bodu 1.

Ve skutečnosti často v bodě 1 není třeba testovat všechny elementární podmíněné nezávislosti, ale můžeme využít vlastností prostoru modelů, z kterého vybíráme. Například díky vlastnosti (4.10) nezávislostních modelů příslušných UG lze předchozí algoritmus zjednodušit na

1. Pro všechna a, b různé prvky N testuj, zda-li $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{N\setminus ab}$.
2. Vyber graf G takový, že bude mít mezi a a b hranu právě tehdy, když test nezávislosti mezi $\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_{N\setminus ab}$ není signifikatní.

Analogická, byť složitější optimalizece kroku 1 pro DAG se nazývá PC algoritmus, viz např. [Nea03], str. 542.

Neodstranitelným problémem však zůstává, že po (např. Bonferroniho) korekci na mnohonásobné testování získáme při zachované hladině α velmi konzervativní testy jednotlivých elementárních nezávislostí. Pokud naopak korekci na mnohonásobné testování neprovedeme, není výběr modelu asymptoticky konzistentní[9] a je třeba být velmi opatrný s interpretací výsledků.

Druhou možností je nalézt model maximalizací Akaikeho (AIC), Bayesovského (BIC) či jiného informačního kritéria

$$
\begin{aligned}
AIC(\hat{\boldsymbol{\theta}}, \boldsymbol{X}) &= 2LL(\hat{\boldsymbol{\theta}}, \boldsymbol{X}) - 2D = 2\sum_{i=1}^{k} \log f_{\hat{\boldsymbol{\theta}}}(X_i) - 2D, \\
BIC(\hat{\boldsymbol{\theta}}, \boldsymbol{X}) &= 2LL(\hat{\boldsymbol{\theta}}, \boldsymbol{X}) - D\log k = 2\sum_{i=1}^{k} \log f_{\hat{\boldsymbol{\theta}}}(X_i) - D\log k,
\end{aligned}
$$

[9] Tzn. neplatí, že pokud jde počet pozorování k nekonečnu, blíží se pravděpodobnost výběru správného modelu k jedné.

kde $\boldsymbol{X}$ jsou data o k řádcích, LL log–věrohodnostní funkce, $\hat{\boldsymbol{\theta}}$ MV odhad parametrů rozdělení a D dimenze prostoru parametrů. Podle [Nea03] (podrobnosti v [Mee97]) je výběr modelu za pomoci AIC/BIC asymptoticky konzistentní.

Jak již bylo zmíněno, problémem je obrovský počet možných grafů (a tudíž i modelů). Je tedy třeba zvolit nějakou vyhledávací strategii. Ať již postupné odebírání (resp. přidávání) hran z úplného (resp. prázdného) grafu nebo střídavé přidávání, odebírání (a u DAG obracení směru hran) podle toho, co maximalizuje nejvíce nárůst hodnoty kritéria. Tyto algoritmy sice mohou být asymtoticky konzistentní, nicméně při větším množství veličin a reálném množství dat (viz úvod) je šance na určení správného modelu poměrně malá.

4.3.1 Gaussovky reprezentovatelné modely

Do této chvíle jsme předpokládali, že data skutečně pocházejí z rozdělení, jehož příslušný nezávislostní model je grafický. Jak však bylo zmíněno v úvodu tento předpoklad je sice nezbytný (už jenom kvůli superexponenciálně rostoucímu počtu všech nezávislostních modelů), v praxi často ani nezmiňovaný, nicméně na první pohled obtížně ospravedlnitelný.

V tomto oddíle si nejprve ukážeme, že třída UG grafických modelů je přirozenou volbou ve smyslu, že jsou to právě ty g^+–reprezentovatelné nezávislostní modely, které odpovídají exponenciálním rodinám. Následně se pokusíme naleznout algoritmus pro výběr modelu a identifikaci parametrů v případě Gaussovsky reprezentovatelných nezávislostních (ne nutně grafických) modelů. A konečně, v simulační studii srovnáme odhady založené pouze na grafických modelech a na všech g^+–reprezentovatelných nezávislostních modelech. Bez újmy na obecnosti se omezíme na regulárně reprezentovatelné modely, neboť u ostatních případů lze z varianční matice snadno určit funkční závislosti mezi veličinami a převést úlohu na výběr a hledání parametrů u příslušného regulárního podvektoru.

Pro každý neorientovaný graf G je $\mathcal{I}_{UG}(G)$ g^+–reprezentovatelný (důkaz v [Lně05]) a navíc, jak již bylo zmíněno v předchozím oddíle, je rodina rozdělení $\mathcal{P}_{g^+}(\mathcal{I}_{UG}(G))$ exponenciální. Toto tvrzení lze však i obrátit, jak ukazuje následující lemma:

Lemma 36. *Nechť I je g^+–reprezentovatelný nezávislostní model a $\mathcal{P}_{g^+}(I)$ je exponenciální rodina. Potom existuje UG G takový, že $I = \mathcal{I}_{UG}(G)$.*

Důkaz. Důkaz lze nalézt v [Šim06c]. Myšlenka však pochází od F. Matúše. □

Omezme se nyní až na g^+–reprezentovatelné modely nad $N = \{1, 2, 3, 4\}$. Tento předpoklad budeme do konce kapitoly pro zjednodušení vynechávat, protože s žádnými jinými než g^+–reprezentovatelnými modely nad čtyřmi veličinami pracovat nebudeme. Nejsnažší způsob, jak vybrat obecný nezávislostní model pro daná data $\boldsymbol{X}$ o k řádcích, je pochopitelně otestovat všech 24 elementárních nezávislostí.

Pokud ale chceme omezit chybu omylem přidané nezávislosti na hladinu $\alpha = 0{,}05$ musíme (při Bonferroniho korekci) jednotlivé testy provádět na hladině $\frac{\alpha}{24} \doteq 0{,}0021$. Sílu testu je možné optimalizovat, jestliže využijeme toho, jak jsou jednotlivé nezávislostní modely do sebe „vnořeny“, a zvolíme sekvenční strategii výběru modelu.

Druhou možností je výběr modelů a určení jeho parametrů na základě AIC/BIC. Odhad střední hodnoty je pro všechny modely stejný a to

$$\hat{\mu}_a = \frac{1}{k} \sum_{b=1}^{k} X_{ba}, \quad a = 1, \ldots, 4.$$

Varianční matici $\boldsymbol{\Sigma}$ parametrizujme jako

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & 0 & 0 & 0 \\ 0 & \sigma_{22} & 0 & 0 \\ 0 & 0 & \sigma_{33} & 0 \\ 0 & 0 & 0 & \sigma_{44} \end{pmatrix} \boldsymbol{\Sigma}_0 \begin{pmatrix} \sigma_{11} & 0 & 0 & 0 \\ 0 & \sigma_{22} & 0 & 0 \\ 0 & 0 & \sigma_{33} & 0 \\ 0 & 0 & 0 & \sigma_{44} \end{pmatrix},$$

kde

$$\boldsymbol{\Sigma}_0 = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{pmatrix}.$$

Odhady parametrů $\sigma_{11}, \ldots, \sigma_{44}$ získáme jako klasické MV odhady rozptylu jednotlivých veličin, tedy

$$\hat{\sigma}_{aa}^2 = \frac{1}{k} \sum_{b=1}^{k} (X_{ba} - \hat{\mu}_a)^2, \quad a = 1, \ldots, 4.$$

Zbývá nalézt MV odhad parametrů z $\boldsymbol{\Sigma}_0$ vhledem k nezávislostním omezením příslušného modelu.

V některých případech bude snadnější parametrizovat nikoli varianční matici, ale její inverzi $\mathbf{\Sigma}^{-1}$ při analogické triku s převodem na matici s jedničkami na diagonále jako výše, tedy označme $\mathbf{\Sigma}_0^{-1} = (\partial_{ij})_{i,j=1,\ldots,4}$.

Připomeňme, že

$$\xi_a \perp\!\!\!\perp \xi_b | \boldsymbol{\xi}_C \quad \Longleftrightarrow \quad |\mathbf{\Sigma}_{aC\cdot bC}| = 0 \quad \Longleftrightarrow \quad |(\mathbf{\Sigma}^{-1})_{a(N\setminus abC)\cdot b(N\setminus abC)}| = 0.$$

Níže jsou navrženy parametrizace pro modely typů M1–M53 z obr. 2.3, str. 39:

M1) $\rho_{12} = 0$

M2) $\rho_{12} = \rho_{14}\rho_{24}$

M3) $\rho_{12} = \frac{\rho_{14}\rho_{24}+\rho_{13}\rho_{23}-\rho_{13}\rho_{24}\rho_{34}-\rho_{14}\rho_{23}\rho_{34}}{1-\rho_{34}^2}$

M4) $\rho_{12} = 0,\ \rho_{13} = -\frac{\rho_{14}(\rho_{23}\rho_{34}-\rho_{24})}{\rho_{24}\rho_{34}-\rho_{23}}$

M5) $\rho_{12} = 0,\ \rho_{34} = \frac{\rho_{23}(1-\rho_{14}^2)+\rho_{13}\rho_{24}\rho_{14}}{\rho_{24}}$

M6) $\rho_{12} = 0,\ \rho_{34} = \rho_{13}\rho_{14} + \rho_{23}\rho_{24}$

M7) $\rho_{12} = 0,\ \rho_{34} = \rho_{14}\rho_{13}$

M8) $\rho_{12} = \rho_{14}\rho_{24},\ \rho_{34} = \rho_{14}\rho_{13}$

M9) $\partial_{12} = 0,\ \partial_{34} = 0$

M10) $\rho_{12} = \rho_{14}\rho_{24},\ \rho_{13} = \frac{\rho_{34}+\rho_{23}\rho_{24}\rho_{14}^2-\rho_{23}\rho_{24}-\rho_{14}^2\rho_{24}^2\rho_{34}}{\rho_{14}(1-\rho_{24}^2)}$

M11) $\rho_{12} = 0,\ \rho_{23} = \rho_{24}\rho_{34}$

M12) $\rho_{12} = 0,\ \rho_{34} = 0$

M13) $\rho_{12} = \rho_{14}\rho_{24},\ \rho_{13} = \frac{\rho_{14}\rho_{24}}{\rho_{23}}$

M14) $\rho_{12} = \rho_{14}\rho_{24},\ \rho_{13} = \frac{\rho_{14}(1-\rho_{23}^2-\rho_{24}^2+\rho_{23}\rho_{24}\rho_{34})}{\rho_{34}-\rho_{23}\rho_{24}}$

M15) $\rho_{14} = \rho_{13}\rho_{34},\ \rho_{12} = \rho_{24}\rho_{13}\rho_{34}$

M16) $\partial_{14} = 0,\ \partial_{23} = 0,\ \partial_{12} = -\frac{\partial_{13}\partial_{24}\partial_{34}}{1-\partial_{34}^2}$

M17) $\rho_{12} = 0,\ \rho_{14} = \rho_{13}\rho_{34},\ \rho_{23} = \frac{\rho_{34}(1-\rho_{13}^2)}{\rho_{24}}$

M18) $\rho_{12} = 0,\ \rho_{34} = \rho_{13}\rho_{14},\ \rho_{24} = -\frac{\rho_{14}(1-\rho_{13}^2-\rho_{23}^2)}{\rho_{13}\rho_{23}}$

M19) $\rho_{12} = 0,\ \rho_{34} = 0,\ \rho_{13} = -\frac{\rho_{23}(1-\rho_{14}^2)}{\rho_{14}\rho_{24}}$

M20) $\partial_{12} = 0,\ \partial_{34} = 0,\ \partial_{13} = -\frac{\partial_{14}\partial_{24}}{\partial_{23}}$

M21) $\rho_{12} = 0,\ \rho_{34} = \rho_{13}\rho_{14},\ \rho_{23} = -\frac{\rho_{24}\rho_{14}(1-\rho_{13}^2)}{\rho_{13}(1-\rho_{14}^2)}$

M22) $\rho_{12} = 0,\ \rho_{34} = \rho_{23}\rho_{24},\ \rho_{13} = \frac{\rho_{23}\rho_{24}}{\rho_{14}}$

M23) $\rho_{12} = 0,\ \rho_{23} = \rho_{24}\rho_{34},\ \rho_{14} = \rho_{13}\rho_{34}$

M24) $\rho_{12} = 0,\ \rho_{34} = \rho_{23}\rho_{24},\ \rho_{14} = \rho_{13}\rho_{23}\rho_{24}$

M25) $\rho_{12} = 0,\ \rho_{14} = \rho_{13}\rho_{34},\ \rho_{23} = -\frac{1-\rho_{24}^2-\rho_{34}^2}{\rho_{24}\rho_{34}}$

M26) $\rho_{12} = \rho_{14}\rho_{24},\ \rho_{13} = \frac{\rho_{14}\rho_{24}}{\rho_{23}},\ \rho_{34} = \rho_{23}\rho_{24}$

M27) $\rho_{12} = 0,\ \rho_{14} = \rho_{13}\rho_{34},\ \rho_{24} = \frac{\rho_{23}(1-\rho_{13}^2\rho_{34}^2)}{\rho_{34}(1-\rho_{13}^2)}$

M28) $\rho_{12} = \rho_{14}\rho_{24},\ \rho_{23} = \rho_{14}\rho_{24}\rho_{13},\ \rho_{34} = \rho_{14}\rho_{24}^2\rho_{13}$

M29) $\partial_{14} = 0,\ \partial_{12} = \partial_{13}\partial_{23},\ \partial_{34} = \partial_{23}\partial_{24}$

M30) $\rho_{12} = 0,\ \rho_{34} = 0,\ \rho_{13} = -\frac{\rho_{14}\rho_{24}}{\rho_{23}}$

M31) $\partial_{23} = 0,\ \partial_{14} = \partial_{13}\partial_{34},\ \partial_{12} = \partial_{13}\partial_{34}\partial_{24}$

M32) $\partial_{34} = 0,\ \partial_{12} = \partial_{14}\partial_{24},\ \partial_{13} = \frac{\partial_{14}\partial_{24}}{\partial_{23}}$

M33) $\rho_{12} = 0,\ \rho_{23} = 0$

M34) $\partial_{12} = 0,\ \partial_{23} = 0$

M35) $\rho_{12} = 0,\ \rho_{34} = 0,\ \rho_{13} = \pm\rho_{24},\ \rho_{14} = \mp\rho_{23}$

M36) $\rho_{34} = \rho_{23}\rho_{24},\ \rho_{12} = \pm\rho_{23}\rho_{24},\ \rho_{13} = \pm\rho_{24},\ \rho_{14} = \pm\rho_{23}$

M37) $\rho_{12} = 0,\ \rho_{23} = \rho_{24}\rho_{34},\ \rho_{13} = \pm\sqrt{1-\rho_{24}^2},\ \rho_{14} = \pm\rho_{34}\sqrt{1-\rho_{24}^2}$

M38) $\rho_{12} = 0,\ \rho_{14} = 0,\ \rho_{23} = \frac{\rho_{24}(1-\rho_{13}^2)}{\rho_{34}}$

M39) $\rho_{12} = 0,\ \rho_{23} = 0,\ \rho_{13} = \rho_{14}\rho_{34}$

M40) $\partial_{14} = 0,\ \partial_{24} = 0,\ \partial_{13} = \frac{\partial_{12}(1-\partial_{34}^2)}{\partial_{23}}$

M41) $\partial_{14} = 0,\ \partial_{24} = 0,\ \partial_{12} = \partial_{13}\partial_{23}$

M42) $\rho_{12} = 0,\ \rho_{13} = 0,\ \rho_{23} = 0$

M43) $\partial_{12} = 0,\ \partial_{13} = 0,\ \partial_{23} = 0$

M44) $\rho_{12} = 0,\ \rho_{23} = 0,\ \rho_{14} = \rho_{13}\rho_{34}$

M45) $\rho_{12} = 0,\ \rho_{23} = 0,\ \rho_{34} = 0$

M46) $\partial_{12} = 0,\ \partial_{23} = 0,\ \partial_{34} = 0$

M47) $\rho_{12} = 0,\ \rho_{13} = 0,\ \rho_{14} = 0$

M48) $\rho_{12} = 0,\ \rho_{23} = 0,\ \rho_{13} = 0,\ \rho_{14} = 0$

M49) $\partial_{12} = 0,\ \partial_{13} = 0,\ \partial_{23} = 0,\ \partial_{14} = 0$

M50) $\rho_{12} = 0,\ \rho_{14} = 0,\ \rho_{23} = 0,\ \rho_{34} = 0$

M51) $\rho_{12} = 0,\ \rho_{13} = 0,\ \rho_{14} = 0,\ \rho_{23} = 0,\ \rho_{34} = 0$

M52) $\rho_{12} = 0,\ \rho_{13} = 0,\ \rho_{14} = 0,\ \rho_{24} = 0,\ \rho_{23} = 0,\ \rho_{34} = 0$

M53) saturovaný model (žádná omezení)

Pro maximalizaci log–věrohodností funkce a následné určení hodnoty AIC/BIC je vhodné využít optimalizační algoritmus nějakého numerického softwaru (např. funkci `optim` v [RD05]). Dimenze prostoru parametrů D je 14 bez počtu nezávislostních omezení (to jest rovnic, nikoli vztahů) pro daný model. Například pro model M11 je $D = 14 - 2 = 12$, neboť máme omezení na σ_{12} a σ_{23}.

Otázkou však zůstává, zda optimalizací skutečně nalezneme maximum log–věrohodnostní funkce. Tak tomu jistě bude u UG modelů, neboť tyto (díky lemmatu 36) vedou na exponenciální rodiny, kdy je log–věrohodnostní funkce konkávní. Vyhovující algoritmus existuje i pro zakřivené exponenciální rodiny (viz [Lau96], str. 272). Nevíme však, zda jsou rodiny příslušné modelům M1–M53 skutečně zakřiveně exponenciální.

Na okraj zmiňme, že za pomoci článku [Drt06] lze zkonstruovat postačující podmínku pro ověření, že daná rodina je zakřiveně exponenciální. Teorie se opírá o algebraickou geometrii a výpočet je proveden v programu Singular, [GPS05]. Avšak pro modely typů M4, M13, M20, M21, M22, M25, M26, M30, M35, M36 a M37 nebyla tato podmínka splněna a přesvědčení o funkčnosti maximalizačních algoritmů pro tyto modely je založeno pouze na empirii.

Odhady varianční matice a algoritmus výběru modelu na základě BIC kritéria jsou implementovány v přiloženém balíčku pro statistické prostředí R. Výběr správného modelu trvá v prostředí Windows na notebooku Inspiron 1501 (AMD Sempron 3500+, 1800 GHz, 1 GB RAM) kolem 9 minut. Uvědomme si, že velikost dat není pro složitost výpočtu směrodatná, neboť postačující statistikou je běžný odhad parametrů pro saturovaný model (tj. bez nezávislostních omezení).

Validita algoritmu byla ověřena následující simulací:

1. Pro každý z 53 typů zvolíme jeden model, označme jej I, a náhodně nagenerujeme varianční matici s tímto modelem korespondující, označme ji V.

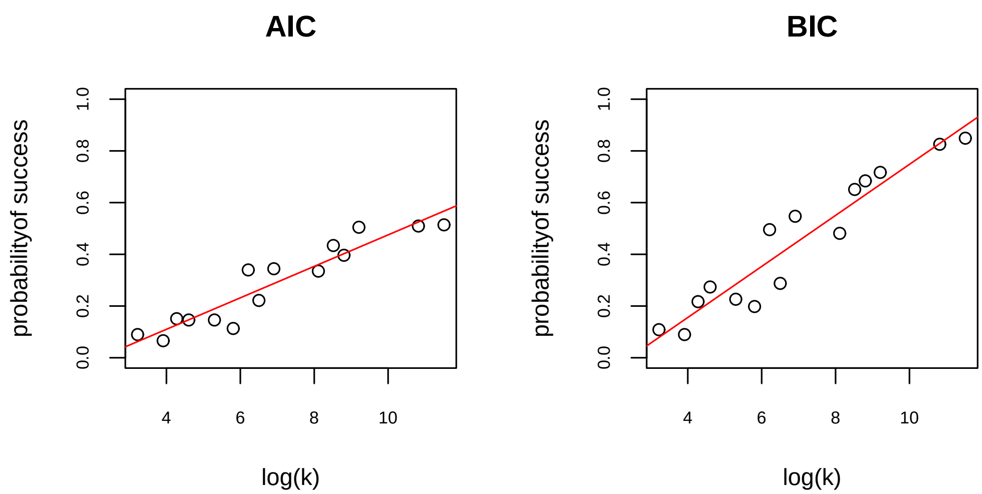
2. Nagenerujeme data z normálního rozdělení s varianční maticí V o k řádcích.

3. Na základě AIC/BIC kritéria vybereme jeden z 629 nezávislostních modelů, označme jej J.

4. Zjistíme, zda-li I a J jsou shodné. Pakliže ano, nazveme identifikaci modelu úspěšnou, ve všech ostatních případech neúspěšnou.

Zdůrazněme, že je lhostejné, jak při simulaci zvolíme parametr střední hodnoty, a že bez újmy na obecnosti můžeme brát varianční matici s jedničkami na diagonále. Kód pro R vypadá následovně:

```
I<-ind.identification(type=testovany.typ)$model
V<-ind.rgauss(model=I)
data<-generate.data(V,k)
J<-model.selection(data)$model
if (I==J) print('Uspech') else print ('Neuspech')
```



Obrázek 4.7: Úspěšnost identifikace modelu na základě AIC a BIC.

Celou simulaci jsme opakovali pro každý z 53 typů a každou z 15 různých velikostí dat k celkem čtyřikrát. Graf průměrné pravděpodobnosti úspěšné identifikace podle AIC/BIC v závislosti na $\log(k)$ můžete vidět na obr. 4.7.

Vidíme, že identifikace na základě BIC je podstatně úspěšnější než u AIC. Dále se tedy budeme zabývat pouze tímto kritériem. Pokud daty proložíme přímku, lze odvodit pomocné pravidlo

$$P(\text{úspěšné identifikace}) \approx 0{,}23 \cdot \log_{10}(k) - 0{,}24.$$

Chceme-li tedy mít pravděpodobnost úspěchu nad 80%, jsou k tomu třeba data o alespoň 40 tisíci řádek.

To, že nezvolíme správný model, nemusí nutně znamenat, že bychom dostali špatný odhad varianční matice. Pokud je např. korelace mezi dvěmi veličinami 0,000001 a my mezi nimi předpokládáme nezávislost, nedopouštíme se zjevně velké chyby. Vzdálenost mezi skutečným a odhadnutým rozdělením budeme měřit tzv. Kullback–Leiblerovy divergencí

$$DIV(\mu, V, \widehat{\mu}, \widehat{V}) = \frac{1}{2}\left(\log\left(\frac{|\widehat{V}|}{|V|}\right) + \operatorname{tr}\left(\widehat{V}^{-1}V\right) + (\widehat{\mu}-\mu)'\widehat{V}^{-1}(\widehat{\mu}-\mu) - n\right),$$

kde $\mu, V, \widehat{\mu}, \widehat{V}$ jsou po řadě parametry střední hodnoty a matice rozptylu prvního a druhého rozdělení, n je počet složek vektoru (tj. v našem případě $n = 4$), tr je stropa matice.

Odhad budeme považovat za tím „lepší“, čím nižší má průměrnou Kullback–Leiblerovu divergenci přes 53 možných typů modelů. Nyní tedy konečně můžeme srovnat odhad bez nezávislostních omezení, odhad založený na základě grafického modelu maximalizujícího BIC, odhad založený na základě nezávislostního modelu maximalizujícího BIC a odhad, kdy je nám skutečný nezávislostní model znám:

1. Pro každý z 53 typů zvolíme jeden model, označme jej I, a náhodně nagenerujeme varianční matici s tímto modelem korespondující, označme ji V.

2. Nagenerujeme data z normálního rozdělení s varianční maticí V o k řádcích.

3. Spočítáme odhad varianční matice v saturovaném modelu a označíme jej $\widehat{V}_0$. Na základě BIC kritéria vybereme nejlepší grafický model a příslušný odhad označíme $\widehat{V}_1$. Analogicky vybereme nezávislostní model a příslušný odhad označíme $\widehat{V}_2$. Odhad varianční matice pro nezávislostní model I označíme $\widehat{V}_3$.

4. Spočteme Kullback–Leiblerovu divergenci mezi V na jedné a $\widehat{V}_0$ až $\widehat{V}_3$ na druhé straně.

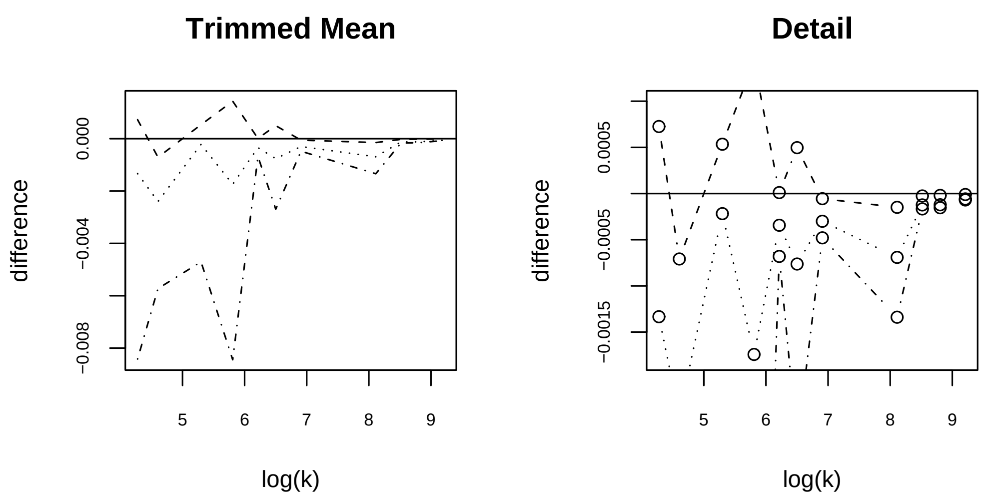
Kód pro R by vypadal následovně:

```
I<-ind.identification(type=testovany.typ)$model
V<-solve(ind.rgauss(model=I))
data<-generate.data(Sigma=V,n=k)

V0<-ind.mle(data,type=53)$Sigma
SEL<-model.selection(data)
V1<-SEL$gSigma
V2<-SEL$Sigma
V3<-ind.mle(data,type=I)$Sigma

DIV0<-kl.div(V,V0,inv=FALSE)
DIV1<-kl.div(V,V1,inv=FALSE)
DIV2<-kl.div(V,V2,inv=FALSE)
DIV3<-kl.div(V,V3,inv=FALSE)
```

Celou simulaci jsme opakovali pro každý z 53 typů a každou z 11 různých velikostí dat k celkem čtyřikrát. Výsledky je skutečně obtížné graficky znázornit. Vezměme odhad pro saturovaný model jako referenční a spočtěme, o kolik jsou jiné odhady lepší či horší. Průměrné[10] rozdíly divergencí v závislosti na $\log(k)$ můžete vidět na levé straně obr. 4.8, detail na témže obrázku vpravo.



Obrázek 4.8: Saturovaný model plná čára, grafický čárkovaná, nezávislostní tečkovaná, skutečný nezávislostní čerchovaná.

Pokusme se výsledky simulace interpretovat:

1. Odvozený odhad varianční matice odpovídající nezávislostnímu modelu maximalizujícímu BIC lze doporučit při rozsahu dat do 10 tisíc řádek. Nad touto mezí jsou si již všechny odhady podobné.

2. Odhad založený pouze na grafických modelech dává pro data o méně než tisíci řádkách velmi neuspokojivé výsledky. Ve všech případech je horší než odhad založený na všech nezávislostních modelech.

3. Odhad odpovídající skutečnému nezávislostnímu modelu dává (vcelku pochopitelně) nejlepší výsledky.

[10] Abychom odstranili extrémní případy, použijme ořezaný průměr s parametrem ořezu 20%.

4. Pro to aby se mělo cenu pokoušet o výběr z 629 nezávislostních modelů a hledání toho skutečného jsou potřeba data o desetitisících, ale spíše statisících pozorování.

S trochou kreativity lze uvedený postup výběru nezávislostního modelu na základě BIC aplikovat i v jiných případech. Jak jsme již zmínili, postačující statistikou pro příslušné odhady je běžný odhad varianční matice vzhledem k saturovanému modelu. Můžeme tedy odhadnout nejprve varianční matici nějakým jiným způsobem (např. použít některý z robustních odhadů) a výběr nezávislostního modelu aplikovat až ex-post. Výše uvedený postup lze tedy chápat jako jistou **korekci** ošetřující možné podmíněné nezávislosti.

Podobně jestliže je náhodných veličin více než čtyři a zapíšeme-li si indexovou množinu N jako sjednocení čtveřic

$$N = \bigcup_i \{a_1^i, a_2^i, a_3^i, a_4^i\},$$

takové, že každá dvojice různých prvků N se vyskytuje v alespoň jedné této čtveřici, potom na tyto čtveřice můžeme aplikovat standardní odhad na základě nezávislostního modelu maximalizujícího BIC. Každý prvek varianční matice je odhadnut díky předpokladu alespoň jednou, v případě více odhadů pro jeden prvek použijeme průměr. Nemáme však zaručeno, že dostaneme pozitivně definitní matici.

V některých případech může být výhodnější odhadovat nikoli varianční matici, ale její inverzi. Záleží na tom, zda se domníváme, že se nám budou v datech vyskytovat spíše nezávislosti typu $\xi_i \perp\!\!\!\perp \xi_j$,$\xi_i \perp\!\!\!\perp \xi_j | \xi_k$, $\xi_i \perp\!\!\!\perp \xi_j | \boldsymbol{\xi}_{kl}$ nebo $\xi_i \perp\!\!\!\perp \xi_j | \boldsymbol{\xi}_{N\setminus ij}$,$\xi_i \perp\!\!\!\perp \xi_j | \boldsymbol{\xi}_{N\setminus ijk}$, $\xi_i \perp\!\!\!\perp \xi_j | \boldsymbol{\xi}_{N\setminus ijkl}$. Výše uvedené postupy je třeba užívat s opatrností a o rozumném chování odhadů se důkladně ujistit za pomocí simulační studie či vyčleněné kontrolní skupiny.

Příklad 5. *Uvedeme si ryze praktický příklad. V chovu hospodářskýh zvířat je snaha o soustavné zlepšování efektivity. Hodnota zvířete pro daný selekční cíl se určuje pomocí tzv. selekčního indexu I. Ten je roven lineární kombinaci daných vlastností pro konkrétní zvíře $\boldsymbol{x} = (x_1, \ldots, x_n)$ s koeficienty $\boldsymbol{b} = (b_1, \ldots, b_n)$ závislými na ekonomické hodnotě a genetických parametrech těchto vlastností. Pro $\boldsymbol{b}$ platí vztah $\boldsymbol{P} \cdot \boldsymbol{b} = \boldsymbol{G} \cdot \boldsymbol{v}$, kde $\boldsymbol{P}$ je kovarianční matice pozorovaných vlastností (fenotypová varianční matice), $\boldsymbol{G}$ shrnuje* *genetické vlastnosti těchto znaků (genetická varianční matice) a* $\boldsymbol{v}$ *je vektor ekonomických hodnot.*

Určováním genetické varianční matice a ekonomických hodnot se zde nebudeme zabývat. Naším cílem bude odhadnout fenotypovou varianční matici. Tento odhad musí být velmi přesný, protože se od něj odvíjí kvalita celého šlechtitelského procesu. Základem musí být znalost hodnot sledovaných vlastností u co největšího reprezentativního vzorku jedinců a následně co nejlepší odhad varianční matice.

V našem příkladě se budeme zajímat o případ šlechění býků českého strakatého skotu. Základem jsou informace o býcích z let 1960 – 2006 z datové banky připravované Českomoravskou společností chovatelů a Plemdat s.r.o. Vlastnosti zahrnuté do selekčního indexu jsou průměrné množství nadojeného mléka za laktaci u dcer (v kg), množství tuku v tomto mléce (v kg), procento tuku v mléce, množství (v kg) a procento bílkovin v mléce a dále tělesné vlastnosti býka, tedy jeho výška v kohoutku, výška v kříži a délka trupu (v cm). Všechny tyto hodnoty jsou vyjádřeny jako odchylka od průměrné hodnoty v populaci. V databázi je všech těchto osm vlastností zaznamenáno u 2063 býků.

Data si rozdělíme na experimentální skupinu o 1800 býcích a kontrolní skupinu o 263 býcích. Indexovou množinu $N = \{1, 2, 3, 4, 5, 6, 7, 8\}$ *lze zapsat jako sjednocení 10 čtveřic:* $\{1, 2, 3, 4\}$, $\{5, 6, 7, 8\}$, $\{1, 2, 7, 8\}$, $\{3, 4, 5, 6\}$, $\{1, 3, 6, 8\}$, $\{2, 4, 5, 7\}$, $\{1, 4, 6, 7\}$, $\{2, 3, 5, 8\}$, $\{1, 2, 5, 6\}$, $\{3, 4, 7, 8\}$. *V experimentální i kontrolní skupině odhadneme parametr varianční matice a to jak v saturovaném modelu, tak za pomoci nezávislostních modelů, jak je popsáno výše. Odhady pro 1800 pozorování jsou si zjevně velice blízké, soustřeďme se tedy na odhady pro kontrolní skupinu. A zde s potěšením zjišťujeme, že odhad na základě nezávislostních modelů má nejen nižší divergenci k obdobnému odhadu v experimentální skupině, ale dokonce i k odhadu pro saturovaný model (byť o méně než* 0,0001*). Zdá se tedy, že odhad za pomoci nezávislostních modelů by mohl být pro tento případ vhodný.*

Pro praktickou konstrukci podobných odhadů by však bylo potřeba výrazně snížit časovou náročnost hledání modelu. Toho by snad bylo možné dosáhnout za pomoci přepsání programu do jazyka C, částečně algebraické namísto čistě numerické optimalizace (výpočet gradientu) a odstraněním opakovaných výpočtů a přenosů dat. U typů modelů, které odpovídají exponenciálním či alespoň zakřiveně exponenciálním rodinám, je zajisté ještě velká rezerva.

Závěr

Cílem této práce byl za prvé stručný úvod do problematiky podmíněné nezávislosti a grafických modelů a za druhé podrobný rozbor známých výsledků reprezentovatelnosti nezávislostních modelů nad čtyřmi (a méně) náhodnými veličinami a jejich další rozšíření. Nakolik byl tento úkol splněn necháváme na posouzení čtenáři.

Krom odpovědí nabízí tato práce i velké množství zajímavých otázek. Jednak se jedná o problémy racionální reprezentovatelnosti a vztahu mezi regulárně gaussovsky a kladně binárně (potažmo diskrétně) reprezentovatelnými nezávislostními modely zmíněnými na konci kapitoly 2. Dále o otevřenou otázku kladné diskrétní reprezentovatelnosti 12 modelů, jejíž zodpovězení je nutné k potvrzení či vyvrácení hypotézy z kapitoly 3. Následně o lepší porozumění, softwarovou implementaci a popularizaci Gaussovských rodin rozdělení z kapitoly 4. V neposlední řadě se zdá být lákavé koncept podmíněné nezávislosti (a tudíž i reprezentovatelnosti) rozšířit na obecné náhodné veličiny, resp. σ–algebry (viz [MR84] a [PS85]) nebo dokonce na obecné (ne nutně konečné) míry. Platí i v tomto případě, že reprezentovatelné modely nelze konečně charakterizovat?

Věřím, že předchozí část pojednávající o konstrukci odhadů založených na nezávislostních modelech by mohla být zajímavá i pro výzkumníky v oblasti aplikované statistiky. Pokud předpokládáme, že podmíněná nezávislost není jen teoretický pojem, ale i jev vyskytující se v reálném světě, potom pro tuto teorii existuje dle mého názoru široké uplatnění.

Pro snadnější využití dosažených výsledků jsou nalezené soubory nezávislostních modelů vzhledem k různým distribučním rámcům spolu s popisem příslušného formátu a procedurami pro snadnou manipulaci přiložené na CD. Zde také naleznete balíček pro prostředí R, v kterém jsou implementovány odhady parametrů regulárního Gaussova rozdělení popsané ve čtvrté kapitole. Dotazy a připomínky k této práci posílejte na adresu

`simecek@gmail.com`.

Literatura

[And78] J. Anděl. *Matematická statistika.* SNTL/ALFA, 1978.

[Bou95] R. R. Bouckaert. *Bayesian Belief Networks: from Construction to Inference.* PhD. práce, University of Utrecht, 1995.

[Cas02] R. Castelo. *The Discrete Acyclic Digraph Markov Model in Data Mining.* PhD. práce, University of Utrecht, 2002.

[CGH96] E. Castillo, J. M. Gutierrez a A. S. Hadi. *Expert Systems and Probabilistic Network Models.* Springer, 1996.

[CLDS99] R. G. Cowell, S. L. Lauritzen, A. P. Dawid a D. J. Spiegelhalter. *Probabilistic Networks and Expert Systems.* Springer, 1999.

[CW96] D.R. Cox a N. Wermuth. *Multivariate Dependencies: Models, Analysis and Interpretation.* Chapman & Hall, 1996.

[Dis00] M. Disman. *Jak se vyrábí sociologická znalost.* Karolinum, 2000.

[Drt04] M. Drton. *Maximum Likelihood Estimation in Gaussian AMP Chain Graph Models and Gaussian Ancestral Graph Models.* PhD. práce, University of Washington, 2004.

[Drt06] M. Drton. Algebraic techniques for Gaussian models. V M. Hušková a M. Janžura, editoři, *Proceedings of Prague Stochastics 2006*, str. 81–90, 2006.

[GPS05] G.-M. Greuel, G. Pfister a H. Schönemann. SINGULAR 3.0. A Computer Algebra System for Polynomial Computations. Centre for Computer Algebra, University of Kaiserslautern (2005). `http://www.singular.uni-kl.de`.

[Jir03] R. Jiroušek. On approximating multidimensional probability distributions by compositional models. V J.-M. Bernard, T. Seidenfeld a M. Zaffalon, editoři, *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, str. 305–320, Waterloo, 2003. Carleton Scientific.

[Jor04] M. I. Jordan. Graphical models. *Statistical Science*, 15:140–155, 2004.

[Koč01] T. Kočka. *Graphical models – learning and applications*. PhD. práce, VŠE Praha, 2001.

[Lac04] P. Lachout. *Teorie pravděpodobnosti*. Karolinum, 2004.

[Lau96] S. L. Lauritzen. *Graphical Models*. Oxford University Press, 1996.

[Lau01] S. L. Lauritzen. Causal inference from graphical models. V O.E. Barndorff-Nielsen, D.R. Cox a C. Klüppelberg, editoři, *Complex Stochastic Systems*, str. 63–107. Chapman & Hall/CRC, 2001.

[Lně05] R. Lněnička. On Gaussian conditional independence structures. Interní publikace DAR–ÚTIA 2005/14, ÚTIA AV ČR, Praha, 2005.

[LR02] S. L. Lauritzen a T. S. Richardson Chain graph models and their causal interpretations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64:321–361, 2002.

[Mat92] F. Matúš. On equivalence of Markov properties over undirected graphs. *Journal of Applied Probability*, 29:745–749, 1992.

[Mat95] F. Matúš. Conditional independences among four random variables II. *Combinatorics, Probability & Computing*, 4:407–417, 1995.

[Mat97] F. Matúš. Conditional independence structures examined via minors. *The Annals of Mathematics and Artificial Intelligence*, 21:99–128, 1997.

[Mat99] F. Matúš. Conditional independences among four random variables III. *Combinatorics, Probability & Computing*, 8:269–276, 1999.

[Mat05] F. Matúš. Conditional independences in Gaussian vectors and rings of polynomials. V G. Kern-Isberner, W. Rödder a F. Kulmann, editoři, *Conditionals, Information a Inference, International Workshop. WCII 2002*, str. 152–161. Springer, 2005.

[Mee97] C. Meek. *Graphical models: selecting causal and statistical models.* PhD. práce, Carnegie Mellon University, 1997.

[MR84] M. Mouchart a J. M. Rolin. A note on conditional independence with statistical application. *Statistica*, 44:557–584, 1984.

[MS95] F. Matúš a M. Studený. Conditional independences among four random variables I. *Combinatorics, Probability & Computing*, 4:269–278, 1995.

[Nea03] R.E. Neapolitan. *Learning Bayesian Networks.* Chapman & Hall/CRC, 2003.

[Pea98] J. Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann, 1998.

[PS85] C. van Putten a J. H. van Schuppen. Invariance properties of conditional independence relation. *The Annals of Probability*, 13:934–945, 1985.

[RD05] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2005. ISBN 3-900051-07-0.

[Rao73] R. C. Rao. *Linear Statistical Inference and Its Applications.* John Wiley & Sons, 1973.

[RS02] T. Richardson a P. Spirtes. Ancestral graph Markov models. *Annals of Statistics*, 30:962–1030, 2002.

[RSSW03] J. M. Robins, R. Scheines, P. Spirtes a L. Wasserman. Uniform consistency in causal inference. *Biometrika*, 90:491–515, 2003.

[SB94] M. Studený a P. Boček. Ci-models arising among 4 random variables. V *Proceedings of the 3rd workshop WUPES*, str. 268–282, 1994.

[Spo94] W. Spohn. On the properties of conditional independence. V P. Humphreys, editor, *Patrick Suppes: Scientific Philosopher*, volume 1, str. 173–194. Kluwer, 1994.

[Stu89] M. Studený. Multiinformation and the problem of characterization of conditional independence relations. *Problems of Control and Information Theory*, 18:3–16, 1989.

[Stu92] M. Studený. Conditional independence relations have no finite complete characterization. V S. Kubík a J.A. Víšek, editoři, *Information Theory, Statistical Decision Functions and Random Processes. Transactions of the 11th Prague Conference*, volume B, str. 377–396. Kluwer, 1992.

[Stu97] M. Studený. Semigraphoids and structures of probabilistic conditional independence. *Annals of Mathematics and Artificial Intelligence*, 21:71–98, 1997.

[Stu02] M. Studený. On stochastic conditional independence: the problems of characterization and description. *Annals of Mathematics and Artificial Intelligence*, 35:241–323, 2002.

[Stu05] M. Studený. *Probabilistic Conditional Independence Structures.* Springer, 2005.

[SV98] M. Studený a J. Vejnarová. The multiinformation function as a tool for measuring stochastic dependence. V M. I. Jordan, editor, *Learning in Graphical Models*, str. 261–298. Kluwer, 1998.

[Šim03] P. Šimeček. On the minimal probability of intersection of dependent events. diplomová práce, Karlova Univerzita v Praze, 2003.

[Šim04] P. Šimeček. A short note on structure learning. V Jana Šafránková, editor, *Proceedings of the 14th Annual Conference of Doctoral Students - WDS 2004*, str. 84–87, Prague, 2004. Matfyzpress.

[Šim05] P. Šimeček. On the minimal probability of intersection of conditionally independent events. Submitted to Kybernetika, 2005.

[Šim06a] P. Šimeček. Classes of Gaussian, discrete and binary representable independence models have no finite characterization. V M. Hušková a M. Janžura, editoři, *Proceedings of Prague Stochastics 2006*, CD volume, str. 622–631. MATFYZPRESS, 2006.

[Šim06b] P. Šimeček. Gaussian representation of independence models over four random variables. V A. Rizzi, editor, *Proceedings of COMPSTAT 2006*, CD volume, str. 1405–1412. Springer, 2006.

[Šim06c] P. Šimeček. Independence models. V J. Vejnarová, editor, *Proceedings of Workshop on Uncertainty Processing 2006*, str. 151–161. University of Economics, 2006.

[Šim06d] P. Šimeček. A short note on discrete representability of independence models. V M. Studený a J. Vomlel, editoři, *Proceedings of Workshop on Probabilistic Graphical Models 2006*, str. 287–292. Action M Agency, 2006.

[ŠS04] P. Šimeček a M. Studený. Využití Hilbertovy báze k ověřování shodnosti strukturálních a kombinatorických imsetů. V Jaromír Antoch a Gejza Dohnal, editoři, *Sborník prací 13. letní školy JČMF ROBUST*, str. 395–402. Jednota českých matematiků a fyziků, 2004.

[ŠŠ06] Č. Štuka a P. Šimeček. Studium souvislostí mezi úspěšností studia medicíny, známkami na střední škole a výsledky přijímacích zkoušek. V *Sborník konference MEDSOFT 2006*. Action M Agency, 2006.

[ŠZTJ04] P. Šimeček, J. Zvárová, M. Tomečková a R. Jiroušek. Application of compositional models to cardiology. V M. Fleschl, editor, *Abstracts of the 11th World Congress on Medical Informatics*, CD volume, str. 1863. IMIA, 2004.