ÚVOD DO LOGLINEÁRNÍHO MODELOVÁNÍ

- historie
  - až do 60. let se k analýze kontingenčních tabulek používal pouze chí-kvadrát test, byl schopen odpovědět na otázku, zdali existuje mezi proměnnými vztah
  - při vícerozměrných tabulkách se počítal chí-kvadrát pro každou dvojrozměrnou sub-tabulku zvlášť, aby se ukázalo, kde vztah existuje a neexistuje
  - v 70. letech se situace dramaticky mění s texty Leo Goodmana a se dvěmi učebnicemi zaměřený na analýzu kontingenčních tabulek (Bishop, Finberg, Holland, 1975; Habermann 1975)
  - v průběhu 80. a 90. let se log-lineární modelování stává součástí standardních statistických znalostí

ÚVOD DO LOGLINEÁRNÍHO MODELOVÁNÍ (POKR.)

- log-lineární modely jsou navrhnuty pro modelování kontingenčních tabulek, používají se tedy k analýze vztahů mezi proměnnými v kontingenčních tabulkách,
- Goodman (1981) dělí kontingenční tabulky podle vztahů mezi proměnnými na 3 typy:
  - sdružená distribuce dvou vysvětlujících proměnných (např. váha a výška)
  - kauzální vztah mezi vysvětlovanou a vysvětlujícími proměnnými (např. kouření a rakovina)
  - asociace mezi dvěma vysvětlovanými proměnnými (např. postoj k interupcím a postoj k předmanželskému sexu)

ÚVOD DO LOGLINEÁRNÍHO MODELOVÁNÍ

- závisle proměnná je v tomto případě počet případů v jednotlivých polích tabulky (poměrová proměnná), aplikace je tedy možná pouze na agregovaná data, při interpretaci pak nerozlišujeme závisle a nezávisle proměnnou jako v logistické regresi
- log-lineární modely jsou schopny ukázat pouze asociaci mezi proměnnými, v tomto smyslu je tato analýza analogická korelační analýze zaměřené na vzorec a sílu vztahu mezi spojitými proměnnými
- název log-lineární analýza je odvozen od transformace, při níž jsou frekvence v polích tabulky převedeny do přirozených logaritmů (v jazyce GLM se jedná o LOG link) a tyto hodnoty jsou modelovány jako lineární funkce sady parametrů

HIERARCHICKÉ LOGLINEÁRNÍ MODELY

- v hierarchických loglineárních modelech je přirozený logaritmus tabulkových četností modelován jako suma efektů, platí přitom princip hierarchie: každá složitější interakce (komplikovanější parametr) obsahuje vždy všechny jednodušší interakce (jednodušší parametry)
- aditivní (lineární) rovnice saturovaného modelu ($f_{ij} = F_{ij}$) pro trojrozměrnou tabulku je:

$$\log(F_{ijk}^{MWY}) = \lambda + \lambda_i^M + \lambda_j^W + \lambda_k^Y + \lambda_{ik}^{MY} + \lambda_{jk}^{WY} + \lambda_{ij}^{MW} + \lambda_{ijk}^{MWY}$$

kde $\log(F_{ijk}^{MWY})$ je přirozený logaritmus očekávané (modelované) četnosti, pro *i*-tý řádek (*M*), *j*-tý sloupec (*W*) a *k*-tou vrstvu (*Y*), přičemž λ je hlavním průměrem, λ_i^M, λ_j^W, λ_k^Y jsou marginálními efekty proměnných *M,W,Y*, λ_{ij}^{MW}, λ_{ik}^{MY}, λ_{jk}^{WY} jsou dvojrozměrnými interakcemi (asociacemi) mezi proměnnými *M,W,Y* a λ_{ijk}^{MWY} je trojrozměrnou interakcí mezi proměnnými *M*, *W*, a *Y*

- upravená rovnice:

$$F_{ijk}^{MWY} = e^{(\lambda + \lambda_i^M + \lambda_j^W + \lambda_k^Y + \lambda_{ik}^{MY} + \lambda_{jk}^{WY} + \lambda_{ij}^{MW} + \lambda_{ijk}^{MWY})}$$

$$F_{ijk}^{MWY} = e^{\lambda} (e^{\lambda_i})^M (e^{\lambda_j})^W (e^{\lambda_k})^Y (e^{\lambda_{ik}})^{MY} (e^{\lambda_{jk}})^{WY} (e^{\lambda_{ij}})^{MW} (e^{\lambda_{ijk}})^{MWY}$$

- multiplikativní rovnice saturovaného modelu tedy je:

$$F_{ijk}^{MWY} = \tau \tau_i^M \tau_j^W \tau_k^Y \tau_{ik}^{MY} \tau_{jk}^{WY} \tau_{ij}^{MW} \tau_{ijk}^{MWY}$$

HIERARCHICKÉ LOGLINEÁRNÍ MODELY

- v nesaturovaných modelech jsou některé z parametrů vynechány, znamená to, že předpokládáme, že jejich efekt odpovídá nule
- hierarchických nesaturovaných modelů přitom platí pravidlo, že pokud vynecháme některý z jednodušších parametrů, tak musíme vynechat rovněž všechny vyšší parametry/interakce, které tento parametr tvoří a naopak
- např. pokud model obsahuje parametr λ_{ij}^{MW}, tak musí obsahovat rovněž jednodušší parametry λ_i^M, λ_j^W
- nebo pokud je dvojrozměrná interakce λ_{ij}^{MW} z modelu vynechána, tak z modelu musí být vynechána rovněž i složitější interakce λ_{ijk}^{MWY}, která dvojrozměrnou interakci obsahuje, a rovněž musejí být rovny nule všechny další vyšší interakce

PARAMETRY LOGLINEÁRNÍCH MODELŮ

- parametry loglineárního modelu ukazují přirozené logaritmy očekávaných/modelových četností jako důsledek tzv. „efektů“, což znamená, že tyto parametry lze interpretovat jako velikost efektu, jímž poznamenávají:
  - distribuci marginálií (jednoduché parametry λ_i, λ_j, λ_k)
  - asociaci nebo parciální asociaci mezi dvěmi proměnnými (dvojrozměrné interakční parametry λ_{ij}, λ_{ik}, λ_{jk})
  - vícerozměrnou asociaci (vícerozměrné interakční parametry λ_{ijk})
- příklady modelů a interpretace parametrů pro trojrozměrnou tabulku M a W podle Y.
  - jednoduchý model $\log(F_{ijk}^{MWY}) = \lambda$ se označuje jako „grand mean model“ a předpokládá, že všechny přirozené logaritmy tabulkových četností jsou si rovné
    - jedná se o „equiprobability model“ a λ je zde chápána jako konstanta (intercept)
  - složitější model $\log(F_{ijk}^{MWY}) = \lambda + \lambda_i^M + \lambda_j^W + \lambda_k^Y$ předpokládá, že přirozené logaritmy tabulkových četností jsou navíc ještě funkcí marginálních četností jednotlivých proměnných
    - jedná se o model nezávislosti (nepředpokládáme výskyt interakcí)
    - parametry λ_i^M λ_j^W λ_j^Y ukazují relativní počet případů v jednotlivých variantách proměnných *M*, *W* a *Y*.

PARAMETRY LOGLINEÁRNÍCH MODELŮ

- ještě složitější model $\log(F_{ij}^{MWY}) = \lambda + \lambda_i^M + \lambda_j^W + \lambda_k^Y + \lambda_{ik}^{MY} + \lambda_{jk}^{WY}$ předpokládá, že přirozené logaritmy tabulkových četností jsou navíc ještě funkcí dvojrozměrných interakcí *MY* a *WY*
    - jedná se o model podmíněné nezávislosti (vztah mezi *MW* je modelován tak, aby zavedením třetí proměnné *Y* zmizel, nepředpokládáme tedy interakci *MW*)
    - parametry λ_{ik}^{MY}, λ_{jk}^{MY} ukazují velikost parciální asociace mezi *M a Y, M a Y*
- ještě složitější model $\log(F_{ij}^{MWY}) = \lambda + \lambda_i^M + \lambda_j^W + \lambda_k^Y + \lambda_{ij}^{MW} + \lambda_{ik}^{MY} + \lambda_{jk}^{WY}$ předpokládá, že přirozené logaritmy tabulkových četností jsou ještě navíc funkcí dvojrozměrných interakcí *MW*, *MY* a *WY*
    - jedná se o model konstantní asociace (vztah mezi *MW* je modelován jako neměnný podle třetí proměnné *Y*)
    - parametry λ_{ij}^{MW}, λ_{ik}^{MY}, λ_{jk}^{MY} ukazují velikost parciální asociace mezi *M a W, M a Y, M a Y*
- nejsložitější model $\log(F_{ijk}^{MWY}) = \lambda + \lambda_i^M + \lambda_j^W + \lambda_k^Y + \lambda_{ij}^{MW} + \lambda_{ik}^{MY} + \lambda_{jk}^{WY} + \lambda_{ijk}^{MWY}$ předpokládá, že přirozené logaritmy tabulkových četností jsou ještě navíc funkcí trojrozměrné interakce *MWY*
    - jedná se o saturovaný model (všechny parametry, jež ovlivňují strukturu dat)
    - parametr λ_{ij}^{MWY} v tomto případě ukazuje, jak se jednotlivé dvojrozměrné interakce liší jedna od druhé v rámci kategorií třetí proměnné, neboli popisuje rozdíl mezi parciální a podmíněnou asociací

PARAMETRY LOGLINEÁRNÍCH MODELŮ A INTERPRETACE

- při celkové interpretaci odlišujeme substantivně méně významné parametry (jednoduché parametry v rovnici, obvykle sedí na data přesně, protože jsou odhadnuty přesně) a substantivně významné parametry (obvykle interakční parametry)
- obecně platí, že parametry marginálií absorbují marginální distribuce a interakční parametry odkazují k asociaci.
- dvojrozměrné interakční parametry tedy přímo korespondují s přirozeným logaritmem poměru šancí v tabulkách:

$$\log(OR) = \log\frac{F_{ij}F_{i'j'}}{F_{ij'}F_{i'j}} = \log F_{ij} + \log F_{i'j'} - \log F_{ij'} - \log F_{i'j} = (\lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB})$$

$$+(\lambda + \lambda_{i'}^A + \lambda_{j'}^B + \lambda_{i'j'}^{AB}) - (\lambda + \lambda_i^A + \lambda_{j'}^B + \lambda_{ij'}^{AB}) - (\lambda + \lambda_{i'}^A + \lambda_j^B + \lambda_{i'j}^{AB})$$

$$= \lambda_{ij}^{AB} + \lambda_{i'j'}^{AB} - \lambda_{ij'}^{AB} - \lambda_{i'j}^{AB}$$

- odhadnuté parametry v rovnici musejí být normalizovány, aby mohly být identifikovány, rozlišujeme:
  - ANOVA typ normalizace, neboli effect coding
  - dummy coding

EFFECT CODING & DUMMY CODING

- effect coding: součet každé sady parametrů u každého *i* nebo *j* se rovná 0 (nebo 1)

$$\log(F_{ij}^{AB}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \qquad F_{ij}^{AB} = \tau\tau_i^A\tau_j^B\tau_{ij}^{AB}$$

$$\sum_i \lambda_i^A = \sum_j \lambda_j^B = \sum_i \lambda_{ij}^{AB} = \sum_j \lambda_{ij}^{AB} = 0 \qquad \prod_i \lambda_i^A = \prod_j \lambda_j^B = \prod_i \lambda_{ij}^{AB} = \prod_j \lambda_{ij}^{AB} = 1$$

  parametry interpretujeme na základě vztahu k průměru, tedy k modelu, který předpokládá, že distribuce četností v tabulce jsou totožné a odpovídají 0 (nebo $e^0=1$), jedná se modelování všechny parametrů jako odchylek od „equiprobability“ modelu

- dummy coding: interpretace parametrů na základě vztahu ke zvoleným (obvykle prvním) kategoriím, *i'* a *j'* jsou referenční kategorie (=0), a proto λ_{ij}^{AB} přímo vyjadřuje přirozený logaritmus poměru (nebo poměr) šancí mezi variantami proměnných vzhledem k referenčním kategoriím

$$\lambda_1^A = \lambda_1^B = \lambda_{1j}^{AB} = \lambda_{i1}^{AB} = 0 \qquad \tau_1^A = \tau_1^B = \tau_{1j}^{AB} = \tau_{i1}^{AB} = 1$$

- normalizace je otázkou konvence a i když jsou velikosti parametrů podle typu normalizace odlišné, rozdíly mezi nimi a výsledná interpretace modelu se neliší, stejně jako se neliší očekávané (modelové) četnosti (LEM preferuje effect coding, nicméně lze zvolit dummy coding, Stata pracuje s dummy coding)

PŘÍKLAD NORMALIZACE PARAMETRŮ

- effect coding: parametry dvojrozměrné interakce (výstup z programu LEM)

```
                     beta     std err   z-value     exp(beta)
 1 1               1.8352    0.0358    51.302       6.2663
 1 2               0.2354    0.0252     9.338       1.2655
 1 3              -0.6977    0.0252   -27.707       0.4978
 1 4              -1.3730                           0.2534
 2 1               0.1875    0.0211     8.896       1.2063
 2 2               1.0131    0.0145    69.950       2.7542
 2 3              -0.0843    0.0132    -6.401       0.9192
 2 4              -1.1164                           0.3275
 3 1              -0.5475    0.0227   -24.119       0.5784
 3 2              -0.1104    0.0146    -7.588       0.8955
 3 3               0.4229    0.0125    33.857       1.5264
 3 4               0.2350                           1.2650
 4 1              -1.4752                           0.2287
 4 2              -1.1382                           0.3204
 4 3               0.3590                           1.4319
 4 4               2.2543                           9.5290 13776.07   9 0.000
```

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

Proveď důkaz, že parametry ukazují jednu a tu samou věc a že tedy typ normalizace je otázkou konvence!!!

- dummy coding: parametry dvojrozměrné interakce (výstup z programu STATA)

```
---------------------------------------------------------------------------
Variable      Label                             estimate       s.e.   p-value
---------------------------------------------------------------------------
rc_fi2        Full interaction: level 2           2.4254     0.0281    0.0000
rc_fi3        Full interaction: level 3           2.2611     0.0371    0.0000
rc_fi4        Full interaction: level 4           1.9044     0.1116    0.0000
rc_fi5        Full interaction: level 5           2.0370     0.0392    0.0000
rc_fi6        Full interaction: level 6           3.5034     0.0564    0.0000
rc_fi7        Full interaction: level 7           3.9910     0.1241    0.0000
rc_fi8        Full interaction: level 8           1.9369     0.0933    0.0000
rc_fi9        Full interaction: level 9           4.3672     0.1043    0.0000
rc_fi10       Full interaction: level 10          6.9380     0.1613    0.0000
---------------------------------------------------------------------------
```

0	0	0	0
0	fi2	fi3	fi4
0	fi5	fi6	fi7
0	fi8	fi9	fi10

POSTUP PŘI LOGLINEÁRNÍM MODELOVÁNÍ

- kontingenční tabulky obvykle zachycují vzorec, který není běžnému pozorovateli zřejmý (v sociální realitě existuje struktura, která je součástí každodenního jednání, která je ovšem pozorovatelná při dostatečném počtu pozorování - zákon velkých čísel - „z ptačí perspektivy“ právě tabulkovým uspořádáním jednotlivých proměnných
- v případě jednoduché (dvojrozměrné) tabulky se tento vzorec v sociálních vědách obvykle interpretuje na základě podílových vyjádření a koeficientů asociace
- v případě vícerozměrné tabulky musíme pro data odhadnou log-lineárních model, v němž strukturu dat specifikujeme (modelujeme ji na základě určitých předpokladů) a v případě sednutí modelu na data pak můžeme interpretovat vztahy, které se v tabulce vyskytují
- specifikace loglineárních modelů a rozhodnutí o jejich „sednutí“ na data je tedy proces, v němž se rozhodujeme, která z asociací a interakcí se signifikantně neliší od 0, tyto parametry pak z modelu eliminujeme a dostáváme se tak k parametrům, které významně přispívají ke vztahům v tabulce

POSTUP PŘI LOGLINEÁRNÍM MODELOVÁNÍ

- při modelování postupujeme stejným způsobem jako v logistické regresi nebo regresní analýze, hledáme co
  - co nejpřesnější model (rozdíl mezi modelovými četnostmi a měřenými četnostmi je malý, neboli velikost residuálů je malá a jejich rozložení + a - je stejné)
  - a zároveň co nejúspornější model (nízký počet parametrů, tedy vysoký počet *df*)
- nejpřesnější model (*df*=0) je saturovaný model - jedná se o parametrizaci pozorovaných četností bez odpovědi na otázku, který z parametrů substantivně přispívá ke struktuře dat v tabulce
- úsporný model má vysoký počet *df*, obvykle se jedná o model nezávislosti
- loglineární modelování je pak hledání modelu, který se nachází někde mezi modelem nezávislosti (pokud tento nesedí na data) a saturovaným modelem

ODHAD HIERARCHICKÉHO LOGLINEÁRNÍHO MODELU

- přímý odhad: vypočet očekávaných četností z distribucí marginálních četností, lze použit pouze v případě testování nulové hypotézy, tedy modelu nezávislosti
  - pro dvojrozměrnou tabulku očekávané četnosti pak vypočítáme podle vzorce:

$$F_{ij} = f_{i+} f_{+j} / f_{++}$$

- nepřímý odhad: výpočet očekávaných četností na základě iterací, používá se v případě testování jiných hypotéz než je nulová hypotéza (pro nulovou hypotézu je výsledek odhadu na základě této metody totožný s výsledkem na základě přímého odhadu)...

TESTY SEDNUTÍ MODELU NA DATA

- Pearsonův chí-kvadrát se stupni volnosti (I - 1)(J - 1), nazývá se také jako reziduální statistika, protože ukazuje rozdíl mezi pozorovanými a očekávanými četnostmi, měří tedy velikost nepadnutí modelu na data
  - pro trojrozměrnou tabulku je jeho vzorec:

$$X^2 = \sum_{ijk} \frac{\left(f_{ijk} - F_{ijk}\right)^2}{F_{ijk}}$$

- věrohodnostní poměr,označuje se jako L^2 (někdy také jako G^2)
  - pro trojrozměrnou tabulku je jeho vzorec: $L^2 = 2\sum_{ijk} f_{ijk} \log\left(\frac{f_{ijk}}{F_{ijk}}\right)$

- stupně volnost (*df*) odkazují k počtu parametrů, které zůstávají při výpočtu „volné“
  - *df* = 0 (saturovaný model), df=(M-1)(W-1)Y (model nezávislosti), df=(M-1)(W-1)(Y-1) (model podmíněné nezávislosti)
  - výpočet df: *df* = počet polí v tabulce - počet parametrů

TESTY SEDNUTÍ MODELU NA DATA (POKRAČ.)

- doplňující statistiky
  - Cressie-Read statistika - podobné jako X^2 a L^2
  - Δ (delta), procentuální vyjádření počtu případů nezařazených do modelu
  - *rG2*, vysvětlující síla modelu, podíl L^2 odhadnutého modelu a L^2 modelu nezávislosti

KOMPARACE MODELŮ

- informační kritéria, jejich cílem není určit, který model je pravdivější, ale který model podává bohatší informaci o reálném světě
  - BIC (Bayesovské informační kritérium)
  - AIC (Akaikeovské informační kritérium)
    - tyto statistiky upřednostňují úspornost před přesností

Lekce 9:

Nehierarchické a topologické log-lineární modely

NEHIERARCHICKÉ LOG-LINEÁRNÍ MODELY

- hierarchické modely nejsou pro specifikaci vztahů v kontingenční tabulce obvykle nejvhodnější
- parametry lze omezovat pouze tak, že je do modelu zařadíme nebo nikoliv
- konkrétnější specifikace parametrů není u hierarchických log-lineárních modelů možná
- k přesnějšímu a zároveň úspornějšímu log-lineárnímu modelu se lze dostat cestou specifikace jednotlivých parametrů
- znamená to, že navrhujeme matici, v níž jednotlivé parametry specifikujeme a na základě této specifikace je pak také modelujeme
- v navržené matici, která obsahuje pole x_{ij}, sloupce označují parametry a řádky tabulková pole
- existují dva způsoby jak specifikovat parametry (souvisejí s typy normalizace parametrů)
  - effect coding: první varianta proměnné v poli matice x_{ij} podle umístění v tabulce odpovídá 1 a poslední odpovídá -1, zbytek odpovídá 0
  - dummy coding: první varianta proměnné v poli matice x_{ij} podle umístění v tabulce odpovídá 1 a zbytek odpovídá 0
    - interakce dostaneme vynásobením parametrů vyjadřujících marginální distribuce, tedy čísel ve sloupcích

PŘÍKLAD SPECIFIKACE MODELU NA ZÁKLADĚ NAVRŽENÉ MATICE

- data 3x4x4 (Y-year, M-man's educational level, W-woman's educational level)

```
124 247 69 18
58 321 167 33
2 31 57 30
2 6 15 20

78 183 68 13
72 619 292 54
6 95 150 65
0 29 61 60

16 45 20 4
46 361 283 47
10 74 195 63
4 26 136 97
```

1) odhad modelu nezávislosti Y M C
počet parametrů = λ+(M-1)+(W-1)+(Y-1) = 9
(*df* = 48 - 9 = 39)

2) odhad modelu podmíněné nezávislosti YM YC
počet parametrů = λ+(M-1)+(W-1)+(Y-1)+[(M-1)(Y-1)
+(W-1)(Y-1)] = 21
(*df* = 48 - 21 = 27)

NAVRŽENÁ MATICE PRO MODEL (YM YC), EFFECT CODING

	λ - parametry																				
pozice		1y	2y	1m	2m	3m	1w	2w	3w	11ym	12ym	13ym	21ym	22ym	23ym	11yw	12yw	13yw	21yw	22yw	23yw
111	1	1	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0
112	1	1	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0
113	1	1	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0
114	1	1	0	1	0	0	-1	-1	-1	1	0	0	0	0	0	-1	-1	-1	0	0	0
121	1	1	0	0	1	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0
122	1	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0
123	1	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
124	1	1	0	0	1	0	-1	-1	-1	0	1	0	0	0	0	-1	-1	-1	0	0	0
131	1	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0
132	1	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0
133	1	1	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0
134	1	1	0	0	0	1	-1	-1	-1	0	0	1	0	0	0	-1	-1	-1	0	0	0
141	1	1	0	-1	-1	-1	1	0	0	-1	-1	-1	0	0	0	1	0	0	0	0	0
142	1	1	0	-1	-1	-1	0	1	0	-1	-1	-1	0	0	0	0	1	0	0	0	0
143	1	1	0	-1	-1	-1	0	0	1	-1	-1	-1	0	0	0	0	0	1	0	0	0
144	1	1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0	-1	-1	-1	0	0	0
211	1	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
212	1	0	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0
213	1	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1
214	1	0	1	1	0	0	-1	-1	-1	0	0	0	1	0	0	0	0	0	-1	-1	-1
221	1	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
222	1	0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0
223	1	0	1	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1
224	1	0	1	0	1	0	-1	-1	-1	0	0	0	0	1	0	0	0	0	-1	-1	-1
231	1	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
232	1	0	1	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0
233	1	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1
234	1	0	1	0	0	1	-1	-1	-1	0	0	0	0	0	1	0	0	0	-1	-1	-1
241	1	0	1	-1	-1	-1	1	0	0	0	0	0	-1	-1	-1	0	0	0	1	0	0
242	1	0	1	-1	-1	-1	0	1	0	0	0	0	-1	-1	-1	0	0	0	0	1	0
243	1	0	1	-1	-1	-1	0	0	1	0	0	0	-1	-1	-1	0	0	0	0	0	1
244	1	0	1	-1	-1	-1	-1	-1	-1	0	0	0	-1	-1	-1	0	0	0	-1	-1	-1
311	1	-1	-1	1	0	0	1	0	0	-1	0	0	-1	0	0	-1	0	0	-1	0	0
312	1	-1	-1	1	0	0	0	1	0	-1	0	0	-1	0	0	0	-1	0	0	-1	0
313	1	-1	-1	1	0	0	0	0	1	-1	0	0	-1	0	0	0	0	-1	0	0	-1
314	1	-1	-1	1	0	0	-1	-1	-1	-1	0	0	-1	0	0	1	1	1	1	1	1
321	1	-1	-1	0	1	0	1	0	0	0	-1	0	0	-1	0	-1	0	0	-1	0	0
322	1	-1	-1	0	1	0	0	1	0	0	-1	0	0	-1	0	0	-1	0	0	-1	0
323	1	-1	-1	0	1	0	0	0	1	0	-1	0	0	-1	0	0	0	-1	0	0	-1
324	1	-1	-1	0	1	0	-1	-1	-1	0	-1	0	0	-1	0	1	-1	1	1	-1	1
331	1	-1	-1	0	0	1	1	0	0	0	0	-1	0	0	-1	-1	0	0	-1	0	0
332	1	-1	-1	0	0	1	0	1	0	0	0	-1	0	0	-1	0	-1	0	0	-1	0
333	1	-1	-1	0	0	1	0	0	1	0	0	-1	0	0	-1	0	0	-1	0	0	-1
334	1	-1	-1	0	0	1	-1	-1	-1	0	0	-1	0	0	-1	1	1	1	1	1	1
341	1	-1	-1	-1	-1	-1	1	0	0	1	1	1	1	1	1	-1	0	0	-1	0	0
342	1	-1	-1	-1	-1	-1	0	1	0	1	1	1	1	1	1	0	-1	0	0	-1	0
343	1	-1	-1	-1	-1	-1	0	0	1	1	1	1	1	1	1	0	0	-1	0	0	-1
344	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	1	1	1	1	1	1

LEM SYNTAX PRO MODEL (Y M C) POMOCÍ NAVRŽENÉ MATICE

- lem zadání pomocí cov( )

```
man 3
dim 3 4 4
lab Y M W
mod {cov(YMW,8)}
des [
1 1 1  1 1 1 1  1 1 1 1  1  1  1  1  1 0 0 0  0 0 0 0  0 0 0 0  0  0  0  0  0 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
0 0 0  0 0 0 0  0 0 0 0  0  0  0  0  0 1 1 1  1 1 1 1  1 1 1 1  1  1  1  1  1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
1 1 1  1 0 0 0  0 0 0 0  0 -1 -1 -1 -1 1 1 1  1 0 0 0  0 0 0 0  0 -1 -1 -1 -1  1  1  1  1  0  0  0  0  0  0  0  0 -1 -1 -1 -1
0 0 0  0 1 1 1  1 0 0 0  0 -1 -1 -1 -1 0 0 0  0 1 1 1  1 0 0 0  0 -1 -1 -1 -1  0  0  0  0  1  1  1  1  0  0  0  0 -1 -1 -1 -1
0 0 0  0 0 0 0  0 1 1 1  1 -1 -1 -1 -1 0 0 0  0 0 0 0  0 1 1 1  1 -1 -1 -1 -1  0  0  0  0  0  0  0  0  1  1  1  1 -1 -1 -1 -1
1 0 0 -1 1 0 0 -1 1 0 0 -1  1  0  0 -1 1 0 0 -1 1 0 0 -1 1 0 0 -1  1  0  0 -1  1  0  0 -1  1  0  0 -1  1  0  0 -1  1  0  0 -1
0 1 0 -1 0 1 0 -1 0 1 0 -1  0  1  0 -1 0 1 0 -1 0 1 0 -1 0 1 0 -1  0  1  0 -1  0  1  0 -1  0  1  0 -1  0  1  0 -1  0  1  0 -1
0 0 1 -1 0 0 1 -1 0 0 1 -1  0  0  1 -1 0 0 1 -1 0 0 1 -1 0 0 1 -1  0  0  1 -1  0  0  1 -1  0  0  1 -1  0  0  1 -1  0  0  1 -1]
dat [...]
```

- takto specifikovanou matici není nutné zadávat, mnohem jednodušší zadání je:

```
man 3
dim 3 4 4
lab Y M W
mod {Y M W}
dat [...]
```

TYPY OMEZENÍ PARAMETRŮ PŘI NAVRHOVÁNÍ MATICE

- navržená matice může být použita pro všechny druhy specifikace nehierarchických log-lineárních modelů, stejně jako pro specifikaci topologických log-lineárních modelů
- existují 3 typy omezení log-lineárních parametrů
  - parametr je specifikován jako roven nule (hierarchické a nehierarchické modely)
  - vybraný parametr specifikován jako roven jinému vybranému parametru (topologické modely)
  - vybraný parametr je specifikován jako daný poměr jiného vybraného parametru (asociativní modely u ordinálních proměnných)