ordinalita kategorií proměnných. Skóry pro řádky a sloupce tabulky jsou neznámé parametry μ_i a μ_j a jsou odhadovány dohromady s parametrem β, který indikuje tabulkovou asociaci. Z tohoto důvodu se tento model nazývá log-multiplikativní (neboli *RC* model, někdy také model *RC II*). Rovnice pro tento model je následující:

$$F_{ij}^{HM} = \eta \tau_i^{H} \tau_j^{M} e^{\mu_i \mu_j \beta} \qquad G_{ij}^{HM} = \theta + \lambda_i^{H} + \lambda_j^{M} + \mu_i \mu_j \beta \tag{46}$$

A/18 Model log-multiplikativního mezitabulkového efektu

O více než desetiletí později Y. Xie (1992) nebo R. Erikson a J. H. Goldthorpe (1992) rozšířili log-multiplikativní princip na mezitabulkovou asociaci (trojrozměrná a vyšší interakce). Nezávisle na sobě navrhují model, v němž jsou odhadnuty parametry pro dvojrozměrnou (tabulkovou) asociaci, přitom je ale pro každou variantu třetí proměnné odhadnuta také multiplikativní odchylka od této dvojrozměrné asociace. Z hlediska interpretace tato odchylka ukazuje, jak se mění dvojrozměrná asociace podle variant třetí proměnné.

Předpokládejme, že modelujeme vztah mezi věkovou homogamií (*H*) a sňatkovým věkem muže (*M*) v jednotlivých letech (*R*). Log-lineární (aditivní) rovnice pro saturovaný model vypadá následovně:

$$G_{ijk}^{HMR} = \theta + \lambda_i^{H} + \lambda_j^{M} + \lambda_k^{R} + \lambda_{ki}^{RH} + \lambda_{kj}^{RM} + \lambda_{ij}^{HM} + \lambda_{ijk}^{HMR} \tag{47}$$

Chceme-li odhadnout model log-multiplikativního mezitabulkového efektu pro tato data, musíme součet parametrů $\lambda_{ij}^{HM} + \lambda_{ijk}^{HMR}$ v této rovnici nahradit součinem parametrů $\psi_{ij}\phi_k$. Parametr ψ_{ij} ukazuje asociaci mezi jednotlivými variantami věkové homogamie a sňatkového věku muže (bez ohledu na roky), parametr ϕ_k ukazuje násobek této asociace neboli její velikost pro jednotlivé roky. Rovnice modelu pak vypadá následovně:

$$G_{ijk}^{HMR} = \theta + \lambda_i^{H} + \lambda_j^{M} + \lambda_k^{R} + \lambda_{ki}^{RH} + \lambda_{kj}^{RM} + \psi_{ij}\phi_k \tag{48}$$

Model se nazývá log-multiplikativní, protože log-lineární rovnice obsahuje multiplikaci dvou parametrů. Jeho předpokladem je, že všechny tabulkové poměry šancí se mění stejným směrem (podle variant třetí proměnné). Z tohoto důvodu je změna v asociaci modelována pouze pomocí jednoho parametru. Díky této charakteristice je tento model v sociálněstratifikačním výzkumu nazýván jako model uniformní diference neboli *unidiff model* (Erikson, Goldthorpe 1992). Pro identifikaci vývoje či změny asociace podle variant třetí proměnné se jedná o velmi vhodný model. Problém spočívá v tom, že na jeho základě nejsme schopni popsat změnu, k níž v poměrech šancí (tabulkové asociaci) podle variant třetí proměnné dochází. Řešení tohoto problému nabízí až model navržený o šest let později Leo Goodmanem a Mikem Houtem (1998, 2001).

Oba badatelé vyšli z předpokladu, že model uniformní diference je příliš restriktivní. Z hlediska úspornosti je to nesporně výhoda, z hlediska popsání změny v tabulkové asociaci se však jedná o značnou nevýhodu. Navrhují proto model, který je dnes znám jako Goodman-Hout model nebo jako model regresního mezitabulkového efektu. V jeho rámci můžeme modelovat jak proměnu poměrů šancí (změnu vzorce tabulkové asociace), tak vývoj velikosti této asociace (trend v asociaci). Vyjdeme-li ze saturovaného modelu v rovnici 47, model regresního mezitabulkového efektu dostaneme tak, že součet parametrů $\lambda_{ij}^{HM} + \lambda_{ijk}^{HMR}$ nahradíme součtem a součinem parametrů $\lambda_{ij}^{HM} + \psi_{ij}\phi_k$. Parametr λ_{ij}^{HM} ukazuje základní vzorec tabulkové asociace, ψ_{ij} ukazuje části asociace, které se mění podle třetí proměnné – v letech – a parametr ϕ_k ukazuje velikost změny asociace pro jednotlivé roky. Log-lineární rovnice takového modelu je pak následující:

$$G_{ijk}^{HMR} = \theta + \lambda_i^{H} + \lambda_j^{M} + \lambda_k^{R} + \lambda_{ki}^{RH} + \lambda_{kj}^{RM} + \lambda_{ij}^{HM} + \psi_{ij}\phi_k \tag{49}$$

Pomocí tohoto modelu dokážeme identifikovat jak změny ve struktuře asociace (poměrech šancí), tak velikost změny asociace v jednotlivých variantách třetí proměnné. Jedná se zatím o poslední a velmi významný posun na poli log-lineárních modelů. Za jistou nevýhodu tohoto modelu lze považovat to, že zatím nebyl uspokojivě aplikován na data, která obsahují více než tři rozměry (na čtyřrozměrné a vícerozměrné tabulky).

A/19 Podoba dat pro log-lineární analýzu

Data pro statistickou analýzu mají buď individuální, nebo agregovanou podobu. V případě, že pracujeme s individuálními daty, analyzujeme matici, v níž je (v jednotlivých polích) zapsaná pozorovaná (měřená) varianta proměnné (bývá obvykle ve sloupcích matice) pro jednotlivé případy (obvykle bývají v řádcích matice). V log-lineární analýze s tímto typem dat nepracujeme. Pokud bychom měli individuální data a chtěli bychom je analyzovat pomocí log-lineárních modelů, bylo by nezbytné je převést na data agregovaná.[88]

88 Jiným řešením je použít logistickou regresi, kterou lze aplikovat jak na individuální, tak agregovaná data, přičemž hodnoty koeficientů logitových a log-lineárních modelů, které se neliší svou strukturou a jsou aplikovány na stejná data, jsou totožná.

Tabulka A.9 Data z tabulky A.1 ve formě četnostních záznamů pro kombinace variant analyzovaných proměnných

Roky	Typ věkového sňatku	Sňatkový věk muže	Věková homogamie a heterogamie	Četnost
1	1	1	1	18 554
1	1	1	2	11 728
1	1	1	3	4 655
1	1	2	1	1 109
1	1	2	2	1 580
1	1	2	3	4 469
1	2	1	1	4 294
1	2	1	2	1 666
1	2	1	3	846
1	2	2	1	361
1	2	2	2	276
1	2	2	3	115
2	1	1	1	11 408
2	1	1	2	6 347
2	1	1	3	1 819
2	1	2	1	3 191
2	1	2	2	4 574
2	1	2	3	6 079
2	2	1	1	4 066
2	2	1	2	2 106
2	2	1	3	1 018
2	2	2	1	771
2	2	2	2	516
2	2	2	3	147

Agregovaná data, prezentovaná obvykle ve formě kontingenčních tabulek, ukazují počet opakujících se pozorování pro jednotlivé kombinace variant proměnných. V tomto případě se nejedná o nic jiného než o přepis (jakkoliv mnohorozměrné) kontingenční tabulky podle variant jednotlivých proměnných do řádků a sloupců matice.

V tabulce A.1 máme agregovaná data, která ukazují počet věkově homogamních a heterogamních sňatků podle sňatkového věku muže a typu věkového sňatku v letech 1994 a 2004 v České republice. Tato data můžeme zapsat také v podobě četností pro jednotlivé kombinace tabulkových proměnných. Tabulka A.9 ukazuje tento zápis (názvy variant jednotlivých proměnných jsou nahrazeny čísly). Jedná se o vymezení všech možných případů z hlediska variant jednotlivých proměnných. Každá četnost ukazuje, kolikrát se daná kombinace variant v datech vyskytuje. Tato data analyzujeme naprosto stejným způsobem jako data individuální, pouze kombinacím jednotlivých proměnných přiřadíme (jim odpovídající) četnosti jako váhy. V log-lineárním modelování pracujeme buď s tímto zápisem dat, nebo s daty v podobě kontingenční tabulky (věcně se jedná o jedno a totéž).

Agregovaná data lze jednoduše převést na individuální tak, že do každého řádku matice (v němž předpokládáme případy) vepíšeme odpovídající počty kombinací jednotlivých variant proměnných. V našem případě víme, že kombinace variant 1 1 1 1 se vyskytuje 18 554 (tabulka A.9). Je nezbytné tedy vepsat 18 554 řádků s hodnotou 1 u každé proměnné. Podobně pak zapíšeme počet řádků daných četnostmi pro všechny zbylé kombinace variant proměnných. Celkový počet řádků v matici pak odpovídá celkovému počtu případů v kontingenční tabulce. V případě tabulky A.9 by to bylo 91 695.

Tabulka A.1 Věkově homogamní a heterogamní sňatky podle sňatkového věku muže a typu věkového sňatku v letech 1994–2004 v ČR

Roky	Typ věkového sňatku	Sňatkový věk muže	Věková homogamie 0–2 roky	Věková heterogamie 3–5 let	Věková heterogamie 6+ let	Celkem
1994	tradiční	18–29	18 554	11 728	4 655	34 937
			20,23 %	12,79 %	5,08 %	38,10 %
		30+	1 109	1 580	4 469	7 158
			1,21 %	1,72 %	4,87 %	7,81 %
	netradiční	18–29	4 294	1 666	846	6 806
			4,68 %	1,82 %	0,92 %	7,42 %
		30+	361	276	115	752
			0,39 %	0,30 %	0,13 %	0,82 %
2004	tradiční	18–29	11 408	6 347	1 819	19 574
			12,44 %	6,92 %	1,98 %	21,35 %
		30+	3 191	4 574	6 079	13 844
			3,48 %	4,99 %	6,63 %	15,10 %
	netradiční	18–29	4 066	2 106	1 018	7 190
			4,43 %	2,30 %	1,11 %	7,84 %
		30+	771	516	147	1 434
			0,84 %	0,56 %	0,16 %	1,56 %
Celkem			43 754	28 793	19 148	91 695
			47,72 %	31,40 %	20,88 %	100 %

Poznámka: Procenta jsou sdružené (celkové) relativní četnosti.