MATHEMATIK-ARBEITSPAPIERE

SEPARATING AGREEMENT FROM ASSOCIATION IN LOG–LINEAR MODELS FOR SQUARE CONTINGENCY TABLES WITH APPLICATIONS

GERHARD OSIUS

PREPRINT NR. 43 NOVEMBER 1993

UNIVERSITÄT BREMEN

MATHEMATIK–ARBEITSPAPIERE

A: MATHEMATISCHE FORSCHUNGSPAPIERE

SEPARATING AGREEMENT FROM ASSOCIATION IN LOG–LINEAR MODELS FOR SQUARE CONTINGENCY TABLES WITH APPLICATIONS

GERHARD OSIUS

PREPRINT NR. 43 NOVEMBER 1993

FACHBEREICH MATHEMATIK/INFORMATIK
UNIVERSITÄT BREMEN

Bibliothekstraße
D–28359 Bremen
Germany

Separating Agreement from Association in Log-linear Models for Square Contingency Tables With Applications

Gerhard Osius

Contents

1. Introduction
2. The Symmetrical and Asymmetrical Components of Association
3. Sampling Models
4. Log-Linear Models for Agreement
5. Data Analysis
6. Model Extensions

Summary

In square contingency tables the agreement between row and column classification is often of primary interest. Our aim is to *model* rather than *measure* agreement and to separate agreement from the remaining association. We decompose the total association – which is summarized in the family of cross product or odds ratios – into an symmetrical component representing agreement and a remaining asymmetrical part. Modeling both components of association separately yields a variety of models for classifications on nominal and ordinal scales. All models proposed are log-linear models for the expected table and are applicable to the usual sampling schemes (Poisson or multinomial), since they do not restrict the expected margins. These models are extended to a set of square tables arizing from stratification of additional observed covariables. The methods are illustrated for a large data set on the classification of uterine cancer and for another set from the literature.

Key words: Agreement, association, cross product ratio, kappa, odds ratio, log-linear models, quasi-symmetry, square contingency tables, uterine cancer.

1. Introduction

Suppose each item under study is classified into J categories in two different ways, giving the classification factors A and B with values $1,\ldots,J$. For example the development of a particular tumor may be classified first using a low cost screening method A and later by a more expensive clinical diagnosis B. Let Y_{jk} denote the observed counts for the classification $\{A=j, B=k\}$, giving rise to a $J\times J$ square contincency table $\mathbf{Y} = (Y_{\mathrm{jk}})$.

The most popular *measure* of agreement beetwen A and B, introduced by Cohen (1960),

$$\kappa = \big(\sum_{\mathrm{j}} \pi_{\mathrm{jj}} - \sum_{\mathrm{j}} \pi_{\mathrm{j+}}\,\pi_{\mathrm{+j}}\big) \,/\, \big(1 - \sum_{\mathrm{j}} \pi_{\mathrm{j+}}\,\pi_{\mathrm{+j}}\big) ,$$

is based on the classification probabilities $\pi_{\mathrm{jk}} = \mathscr{P}\{A=j,\ B=k\}$, where "+" indicates summation over the missing subscript. An unsatisfactory feature of κ is that its value depends on the marginal distributions $(\pi_{\mathrm{j+}})$ and $(\pi_{\mathrm{+k}})$ of A and B, which are fixed in advance by certain sampling schemes and do not contain any information about the association between the two factors. Moreover, in many applications it is not sufficient to summarize the agreement by a single number. Therefore we aim to *model* the *agreement* and to separate it from the *total association* between A and B by means of log-linear models for the expected counts $\mu_{\mathrm{jk}} = \mathscr{E}(Y_{\mathrm{jk}})$. Instead of *measuring agreement globally* we then estimate and test for significance the corresponding parameters in the model.

In a 2×2 table the factors A and B agree exactly on the two main diagonal cells and disagree on the other two off-diagonal cells. Hence the cross-product ratio (or odds ratio) $\theta = \mu_{11}\,\mu_{22} / \mu_{12}\,\mu_{21}$ resp. its logarithm $\psi = \log\theta$, which characterizes the *association* independently of the margins, also describes the *agreement*. The sign of ψ gives the direction of agreement (negative values indicating disagreement) and the absolute value measures its strength, which is 0, of course, if and only if A and B are independent.

For a general $J\times J$ table, the total agreement between A and B is determined by the agreement on all subsets $\{j,k\}$ for any two categories $1\leq j<k\leq J$, which in turn are given by the cross product ratios

(1) $$\theta_{\mathrm{jk}} = \mu_{\mathrm{jj}}\,\mu_{\mathrm{kk}} / \mu_{\mathrm{jk}}\,\mu_{\mathrm{kj}}$$

of all 2×2 subtables with rows and colums in $\{j,k\}$. The family (θ_{jk}) thus summarizes the total *agreement* beetween A and B. The total *association* beetween A and B however, is determined by the family of all cross product ratios

(2) $$\theta_{\mathrm{jk\,j'\,k'}} = \mu_{\mathrm{jk}}\,\mu_{\mathrm{j'k'}}/\mu_{\mathrm{jk'}}\,\mu_{\mathrm{j'k}}$$

for all $j,\ k,\ j',\ k' \in \{1,\ldots,J\}$. Our aim is to separate agreement from association by means of decomposing the association family $(\theta_{\mathrm{jk\,j'\,k'}})$ into a symmetrical and an asymmetrical component. The symmetrical part turns out to be uniquely determined by the subfamily (θ_{jk}) and thus represents *agreement*, whereas the asymmetrical part is a family (Δ_{jk}) containing the remaining association beetween A and B. The family (θ_{jk}) of agreement has two obvious, very desirable properties:

(i) the family does not depend on the marginal distributions of the factors A and B,

(ii) the value of a single member θ_{jk} depends only on the 2×2 subtable with rows and colums in $\{j,k\}$ and hence remains unchanged if one or more of the remaining categories are deleted or modified (e.g by combining adjacent categories).

Our main goal is to present models for agreement by specifying only (θ_{jk}), thus leaving the remaining structure of the expected table $\boldsymbol{\mu}$ completely arbitrary. Allowing for additional (asymmetrical) association besides agreement distinguishes the proposed models from popular *quasi-symmetry* models (cf. Agresti 1990, chapter 10) and leads to a broader class of models. Starting with models for *nominal* classifications (which do not depend on the possibly artificial ordering of the categories $1,...,J$), we proceed to models for *ordinal* classifications, which exploit the natural ordering or additional scores assigned to the categories. Although not of primary interest here, some models for the asymmetrical association family (Δ_{jk}) are also given. Furthermore, the above models are extended to the case where the observed data is stratified according to an additional factor $C\in\{1,...,S\}$, which is typically obtained from grouping one more observed covariables attributed to each item under study.

All models given here correspond to log-linear models for the expected table $\boldsymbol{\mu}$ thus allowing an analysis of the observed table $\mathbf{Y}$ by means of standard statistical software packages. Furthermore the statistical analysis of the families (θ_{jk}), and (Δ_{jk}) under these models is invariant under the common sampling schemes: Poisson or multinomial, the latter fixing either the total Y_{++}, the row totals $Y_{\mathrm{j}+}$ or the column totals $Y_{+\mathrm{k}}$ in advance.

Example 1: Cytological and Pathological Classification of Uterine Cancer

The example, which motivated this paper, is concerned with the diagnosis of uterine cancer (carcinoma of the cervix). In Germany a preventive check-up to detect various forms of cancer is offered free of charge by the health insurances. As part of the investigations for females, a vaginal smear is analyzed histologically and categorized according to a cytological classification given in table 1-1, which goes back to Papanicolaou (1943). For the categories *Pap III* , *IV* and *V*, a suspicion of a carninoma cannot be ruled out, and a pretherapeutic cone-biopsy is suggested, in order to obtain a reliable diagnosis based on a pathological classification of the tissue using the same categories of table 1-1. Obviously, the accuracy of the cytological classification *CYT* as measured by its agreement with the final pathological classification *PAT* is of major importance.

Pap I	normal cells
Pap II	light inflamed or regenerative alterations
Pap III D L	light displasias
Pap III D M	medium displasias
Pap IV A S	severe displasias
Pap IV A Cis	carcinoma in situ
Pap IV B	carcinoma in situ with beginning infiltration
Pap V	carcinoma in epithelial tissue or adenoid carcinoma

Table 1-1: Categories for cytological and pathological classifications in example 1.

The data to be analyzed here are taken from the records of the *Cytological Lab* (headed by Prof. Dr. D. Langnickel) in the gynaecological clinic of the central hospital *Zentralkrankenhaus St. Jürgen Strasse* in Bremen, Germany. From the large body of data only those individual records from 12 years (1972-1983) are selected, for which both classifications *CYT (A)* and *PAT (B)* were given. Since a cone-biopsy was routinely performed only if *CYT* was *Pap III* or higher, our analysis is restricted to these cases. As mentioned earlier, the exclusion of the first two categories does not affect the agreement between the remaining categories. This leaves us with $J=6$ response categories from *Pap III DL* to *Pap V*, exhibiting a natural order corresponding to the severity of suspicion. The observed counts for all 12 years is given in table 1-2 .

Cytological classification	Pathological classification III DL	III DM	IV AS	IV AC	IV B	V	Total
III DL	53	42	14	9	0	2	120
III DM	35	112	85	79	1	6	318
IV AS	4	27	108	76	3	8	226
IV AC	8	14	56	293	16	43	430
IV B	0	0	6	37	14	49	106
V	0	3	8	84	14	529	638
Total	100	198	277	578	48	637	1838

Table 1-2: Observed counts (12 years) for cytological and pathological classifications in example 1.

Example 2: Number of Pregnancies Reported by the Mother and Father

Velema et al. (1991) analyzed agreement in data from Columbia in order to compare responses from wives *(A)* and husbands *(B)* concerning the reproductive history of the women. We would like to thank the authors for generously providing us with their complete data set, which will be reanalyzed here only in one aspect to illustrate our approach. Table 1-3 gives the cross-classification of all 199 couples according to the number of pregnancies reported by each partner.

Pregnancies Reported by Mother	Pregnancies Reported by Father 1	2	3	4	5 +	Total
1	62	16	6	2	1	87
2	0	35	4	3	2	44
3	0	1	18	8	3	30
4	1	1	0	9	4	15
5 +	0	0	1	0	22	23
Total	63	53	29	22	32	199

Table 1-3: Observed counts for number of pregnancies reported by mother and father in example 2 (taken from Velema et al. 1991)

2. The Symmetrical and Asymmetrical Components of Association

We consider a $J \times J$-contingency table $\mathbf{Y} = (Y_{jk})$ with positive expectation $\mu_{jk} = \mathscr{E}(Y_{jk}) > 0$ for each cell count Y_{jk}. For any pair of cells (j,k) and (j',k') the cross product or odds ratio $\theta_{jk\,j'k'}$ for the 2×2 subtable with columns $\{j, j'\}$ and rows $\{k, k'\}$ is given by (2). It is well known (Plackett 1974, sec. 3.4) that the expected table $\boldsymbol{\mu}$ is uniquely determined by: **(a)** the row margins (μ_{j+}), **(b)** the column margins (μ_{+k}), and **(c)** the cross product ratio family $(\theta_{jkj'k'})$.

Our primary interest focuses on the subfamily (θ_{jk}) given in (1), which contains all cross product ratios $\theta_{jk} = \theta_{jjkk}$ of "quadratic" 2×2 subtables along the main diagonal. Our aim is to specify parametric models for the subfamily (θ_{jk}), leaving the remaining structure of the expected table $\boldsymbol{\mu}$ completely arbitrary. Before proposing specific models it is useful to give a convenient parametrization of the *saturated* model which imposes no structure on the expected table whatsoever. In the saturated model, the log-expectations $\eta_{jk} = \log \mu_{jk}$ may take arbitrary real values, and a common parametrization is

$$\eta_{jk} = \lambda_0 + \lambda_j^A + \lambda_k^B + \lambda_{jk} . \tag{3}$$

To identify the parameters, suitable linear constraints have to be imposed. We adopt the constraints

$$\lambda_1^A = 0 , \qquad \lambda_1^B = 0 , \qquad \lambda_{j1} = 0 \qquad \lambda_{1k} = 0 \qquad \text{for all } j \text{ and } k, \tag{4}$$

which allow the following interpretation of the parameters

$$\begin{aligned} \lambda_0 &= \eta_{11} = \log \mu_{11} \\ \lambda_j^A &= \eta_{j1} - \eta_{11} = \log (\mu_{j1}/\mu_{11}) \\ \lambda_k^B &= \eta_{1k} - \eta_{11} = \log (\mu_{1k}/\mu_{11}) \\ \lambda_{jk} &= \eta_{11} + \eta_{jk} - \eta_{j1} - \eta_{1k} = \log \theta_{11jk} . \end{aligned} \tag{5}$$

The association parameters λ_{jk} thus completely determine the family $(\theta_{jkj'k'})$ of cross product ratios resp. the family of logarithms

$$(6) \qquad \psi_{jkj'k'} := \log \theta_{jkj'k'} = \eta_{jk} + \eta_{j'k'} - \eta_{jk'} - \eta_{j'k} \,.$$

Since we are dealing with a square table, we get a decomposition

$$(7) \qquad \lambda_{jk} = \overline{\lambda}_{jk} + \Delta_{jk} \qquad \text{with}$$

$$\overline{\lambda}_{jk} := \tfrac{1}{2}(\lambda_{jk} + \lambda_{kj}) \,, \qquad \Delta_{jk} := \tfrac{1}{2}(\lambda_{jk} - \lambda_{kj}) \,.$$

In view of

$$(8) \qquad \overline{\lambda}_{jk} = \overline{\lambda}_{kj}, \qquad \Delta_{jk} = -\Delta_{kj}$$

the family $(\overline{\boldsymbol{\lambda}}_{jk})$ represents the *symmetrical* and (Δ_{jk}) the *asymmetrical* part of the association $\boldsymbol{\lambda}$. The main advantage of this decomposition is that the symmetrical part $\overline{\boldsymbol{\lambda}}$ completely characterizes the family (θ_{jk}) of agreement resp. the family of logarithms

$$(9) \qquad \psi_{jk} := \log \theta_{jk} = \eta_{jj} + \eta_{kk} - \eta_{jk} - \eta_{kj} \,.$$

In fact, inserting (3) into (9) yields

$$(10) \qquad \psi_{jk} = \lambda_{jj} + \lambda_{kk} - 2\,\overline{\lambda}_{jk} \,.$$

Using $\lambda_{jj} = \overline{\lambda}_{jj}$, gives $\boldsymbol{\psi}$ as a function of $\overline{\boldsymbol{\lambda}}$. And conversely, using $\lambda_{jj} = \psi_{1j}$ from (5), yields $\overline{\boldsymbol{\lambda}}$ as function of $\boldsymbol{\psi}$:

$$(11) \qquad \overline{\lambda}_{jk} = \tfrac{1}{2}(\psi_{1j} + \psi_{1k} - \psi_{jk}).$$

Inserting (7) and (11) into (3) now gives the desired parametrization of the saturated model

$$(\mathrm{SM}) \qquad \eta_{jk} = \alpha_0 + \alpha_j^{A} + \alpha_k^{B} + \Delta_{jk} - \tfrac{1}{2}\psi_{jk} \qquad \textit{(saturated model)},$$

with new parameters $\alpha_0 := \lambda_0$, $\alpha_j^{A} := \lambda_j^{A} + \frac{1}{2}\psi_{1j}$ and $\alpha_k^{B} := \lambda_k^{B} + \frac{1}{2}\psi_{1k}$ satisfying

$$(12) \qquad \alpha_1^{A} = 0 \,, \qquad \alpha_1^{B} = 0 \,,$$

$$\Delta_{1k} = 0 \,, \qquad \Delta_{jk} = -\Delta_{kj} \,,$$

$$\psi_{jj} = 0 \,, \qquad \psi_{jk} = \psi_{kj} \,, \qquad \text{for all } j,\ k.$$

The symmetrical part $\boldsymbol{\psi}$ is determined by the subfamily ψ_{jk} with $1 \leq j < k \leq J$ through

$$(13) \qquad \psi_{jk} = I\{j \neq k\} \cdot \psi_{(j \wedge k)\,(j \vee k)} \qquad \text{for all } j \text{ and } k.$$

Here $I\{E\}$ denotes the indicator for an event E (which equals 1 if E occurs and 0 otherwise), and $j \wedge k$ resp. $j \vee k$ is the minimum resp. maximum of (j, k). And the asymmetrical part $\boldsymbol{\Delta}$ is determined by all Δ_{jk} with $1<\mathrm{j}<\mathrm{k}\leq J$ through

(14) $$\Delta_{\mathrm{jk}} = I\{1<j\wedge k\} \cdot \mathrm{sgn(k–j)} \cdot \Delta_{(\mathrm{j}\wedge\mathrm{k})\,(\mathrm{j}\vee\mathrm{k})} \qquad \text{for all } j \text{ and } k,$$

where *sgn(m)* denotes the sign of m, taking values –1, 0, +1 for $m<0$, $=0$, >0. The dimensions of these parameter families are summarized in table 2-1.

source	parameter	dimension
constant	α_0	1
row factor A	$\boldsymbol{\alpha}^{\mathrm{A}}$	J–1
column factor B	$\boldsymbol{\alpha}^{\mathrm{B}}$	J–1
association: asymmetry	$\boldsymbol{\Delta}$	$\frac{1}{2}(J–1)(J–2)$
association: symmetry	$\boldsymbol{\psi}$	$\frac{1}{2}J(J–1)$
total		$J\times J$

Table 2-1: Dimensions of the parameters in the saturated model.

Using (13) and (14) the saturated model may be written as

(SM)* $$\eta_{\mathrm{jk}} = \alpha_0 + \alpha_{\mathrm{j}}^{\mathrm{A}} + \alpha_{\mathrm{k}}^{\mathrm{B}} + I\{1<j\wedge k\}\cdot\mathrm{sgn}(k–j)\cdot\Delta_{(\mathrm{j}\wedge\mathrm{k})\,(\mathrm{j}\vee\mathrm{k})} - \tfrac{1}{2}I\{j\neq k\}\cdot\psi_{(\mathrm{j}\wedge\mathrm{k})\,(\mathrm{j}\vee\mathrm{k})}$$

with no further restrictions on the parameters α_0, $\{\alpha_{\mathrm{j}}^{\mathrm{A}}\,|\,1<j\leq J\}$, $\{\alpha_{\mathrm{k}}^{\mathrm{B}}\,|\,1<k\leq J\}$, $\{\Delta_{\mathrm{jk}}\,|\,1<j<k\leq J\}$ and $\{\psi_{\mathrm{jk}}\,|\,1\leq j<k\leq J\}$.

For an explorative data analysis it is instructive to compute the (corrected) observed cross product ratios

$$\tilde{\theta}_{\mathrm{jkj'k'}} = (Y_{\mathrm{jk}}+c)\,(Y_{\mathrm{j'k'}}+c)\,/\,(Y_{\mathrm{jk'}}+c)\,(Y_{\mathrm{j'k}}+c)$$

and analyse the corresponding families $(\tilde{\psi}_{\mathrm{jk}})$ and $(\tilde{\Delta}_{\mathrm{jk}})$ of agreement and asymmetry. A good choice for the correction is $c=\frac{1}{2}$ (Cox and Snell 1989, 2.1.6), but if all Y_{jk} are positive, the ML estimate under the saturated model (given by $c=0$) may be preferable.

Example 1 (ctd.): Cytological and Pathological Classification of Uterine Cancer

Table 2-2 shows the observed agreement $\tilde{\psi}_{\mathrm{jk}}$ with its standard error as computed from the saturated model for the corrected observed table $(\mathrm{Y}_{\mathrm{jk}}+\frac{1}{2})$. Not surprisingly, all values of $\tilde{\psi}_{\mathrm{jk}}$ are significantly positive and tend to increase with the "distance" between the ordered responses j and k. The observed asymmetries $\tilde{\Delta}_{\mathrm{jk}}$ (table 2-2) however do not reveal any apparent pattern and only two entries (for $j=4$) are significantly different from 0.

Cytological classification (j)		Pathological classification (k) III DL (1)	III DM (2)	IV AS (3)	IV AC (4)	IV B (5)	V (6)
III DL	(1)						
III DM	(2)	1.38 ± 0.28		0.07 ± 0.31	0.89 ± 0.30	0.64 ± 1.30	− 0.41 ± 0.85
IV AS	(3)	4.49 ± 0.56	1.65 ± 0.26		0.68 ± 0.37	0.28 ± 1.09	− 0.22 ± 0.86
IV AC	(4)	5.27 ± 0.49	3.35 ± 0.31	2.00 ± 0.21		− 0.35 ± 1.04	− 1.08 ± 0.82
IV B	(5)	8.04 ± 2.02	7.68 ± 1.66	4.24 ± 0.72	1.93 ± 0.40		− 0.19 ± 1.27
V	(6)	10.03 ± 1.56	7.87 ± 0.67	6.68 ± 0.50	3.74 ± 0.20	2.37 ± 0.40	

Table 2-2: Observed agreement $\widetilde{\psi}_{jk}$ for $j>k$ (lower triangle) and asymmetries $\widetilde{\Delta}_{jk}$ for $1<j<k$ (upper triangle) with standard errors (±S.E.) in example 1 (computed for the saturated model using corrected counts $Y_{jk}+\frac{1}{2}$).

Example 2 (ctd.): Number of Pregnancies Reported by the Mother and Father

All values $\widetilde{\psi}_{jk}$ for the observed agreement in table 2-3 are significantly positive but no apparent pattern emerges taking the standard errors into account. The observed asymmetries $\widetilde{\Delta}_{jk}$ (table 2-3) are positive for all $j<k$, but only two entries are significantly different from 0.

Pregnancies Reported by Mother (j)	Pregnancies Reported by Father (k) 1	2	3	4	5 +
1					
2	5.59 ± 1.45		1.02 ± 1.13	1.92 ± 1.01	2.00 ± 1.33
3	5.87 ± 1.49	4.58 ± 0.99		2.44 ± 1.16	1.16 ± 1.20
4	5.06 ± 1.09	4.16 ± 1.04	3.72 ± 1.51		0.80 ± 1.22
5 +	7.54 ± 1.65	6.46 ± 1.57	4.37 ± 1.03	4.55 ± 1.54	

Table 2-3: Observed agreement $\widetilde{\psi}_{jk}$ for $j>k$ (lower triangle) and asymmetries $\widetilde{\Delta}_{jk}$ for $1<j<k$ (upper triangle) with standard errors (±S.E.) in example 2 (computed for the saturated model using corrected counts $Y_{jk}+\frac{1}{2}$).

3. Sampling Models

There are basically four different methods to sample the items which are to be categorized according to A and B from the population under study. Each sampling scheme described briefly below leads to a different distribution for the observed table $\mathbf{Y}$ and hence to a specific interpretation of the expected table $\boldsymbol{\mu}$.

If the items are sampled without any restriction on the total sample size Y_{++} (e.g. by sampling over a fixed period of time) then the *Poisson model* is usually appropriate, in which all Y_{jk} are independent each having a Poisson distribution with expectation μ_{jk}. The classification probabilities are then given by $\pi_{jk} = \mu_{jk}/\mu_{++}$.

The *multinomial model* arises if n items are randomly selected from the population. Then the observed table (Y_{jk}) has a multinomial distribution of size $n = Y_{++}$ with cell probabilities $\pi_{jk} = \mathscr{P}\{A{=}j,\, B{=}k\}$. In view of $\mu_{jk} = n\,\pi_{jk}$, the cross product ratio family for the expectations coincides with the correponding family for the probabilities, i.e.

$$\theta_{jkj'k'} = \pi_{jk}\,\pi_{j'k'}/\pi_{jk'}\,\pi_{j'k},$$

and hence does not depend upon n. Furthermore the symmetrical and asymmetrical components $\boldsymbol{\psi}$ and $\boldsymbol{\Delta}$ may also be interpreted in terms of the probability table (π_{jk}).

Suppose now that for each category j of A a fixed number n_j of items are sampled conditionally upon $A{=}j$ (i.e from the corresponding subpopulation). This leads to the *row multinomial model* in which all rows of the table (Y_{jk}) are independent, each having a multinomial distribution of size $n_j = Y_{j+}$ with conditional cell probabilities $\pi^B_{jk} := \mathscr{P}\{B{=}k \mid A{=}j\}$. Again $\mu_{jk} = n_j\,\pi^B_{jk}$ implies that the cross product ratio family of $\boldsymbol{\mu}$ coincides with the corresponding family for the conditional probabilities, i.e.

$$\theta_{jkj'k'} = \pi^B_{jk}\,\pi^B_{j'k'} \,/\, \pi^B_{jk'}\,\pi^B_{j'k}\,,$$

and hence does not depend on the row totals n_j.

Finally, interchanging rows with columns (i.e A with B) in the row multinomial model yields a *column multinomial model*, based on the conditional probabilities $\pi^A_{jk} := \mathscr{P}\{A{=}j \mid B{=}k\}$.

The symmetrical and asymmetrical components of association can also be described in terms of the conditional probabilities $\boldsymbol{\pi}^B$ or $\boldsymbol{\pi}^A$, allowing for an additional interpretation of $\boldsymbol{\psi}$ and $\boldsymbol{\Delta}$. The family (π^B_{jk}) is uniquely determined by the ratios

$$\text{(15)} \qquad \gamma^B_{jk} := \mathscr{P}\{B{=}k \mid A{=}j\} \,/\, \mathscr{P}\{B{=}j \mid A{=}j\} = \pi^B_{jk} \,/\, \pi^B_{jj} = \mu^B_{jk} \,/\, \mu^B_{jj}\,,$$

representing the *odds for a disagreement* $B{=}k$ given $A{=}j$. The odds ratio family (θ_{jk}) for agreement can now be written as

$$\text{(16)} \qquad \theta_{jk} = (\gamma^B_{jk} \cdot \gamma^B_{kj})^{-1} \qquad \text{resp.} \qquad \psi_{jk} = -\log \gamma^B_{jk} - \log \gamma^B_{kj}\,.$$

Using (SM), the log-odds-ratio of disagreement is

(17) $$\varepsilon_{jk} := \log(\gamma^B_{jk} / \gamma^B_{kj}) = 2\,(\alpha^B_j - \alpha^B_k + \Delta_{jk})$$ *(log-odds-ratio of disagreement).*

The constraints (12) imply $\alpha^B_j = \frac{1}{2}\varepsilon_{j1}$, thus giving the asymmetrical part $\boldsymbol{\Delta}$ as a function of $\boldsymbol{\varepsilon}$ and hence of $\boldsymbol{\pi}^B$

(18) $$\Delta_{jk} = \tfrac{1}{2}(\varepsilon_{jk} - \varepsilon_{j1} + \varepsilon_{k1}) \,.$$

4. Log-Linear Models for Agreement

Starting with the saturated model (SM) for the expected counts $\boldsymbol{\mu}$ we can easily describe submodels which only impose restrictions on the symmetrical part $\boldsymbol{\psi}$ of the association leaving the remaining structure of $\boldsymbol{\mu}$ resp. $\boldsymbol{\eta}$ arbitrary. We will restrict ourselves here to linear models for $\boldsymbol{\psi}$, i.e. log-linear models for (θ_{jk}), which yield log-linear models for the expected counts $\boldsymbol{\mu}$ by replacing ψ_{jk} in (SM)* with corresponding model formulae for ψ_{jk}.

We first look at models which do not refer to any ordering of the response categories 1,...,*J* and hence apply to nominal reponses (e.g. blood group). The most elementary such models are

(ZS) $$\psi_{jk} = 0 \qquad \text{for all } j \neq k$$ *(zero symmetry),*

(CS) $$\psi_{jk} = \psi_0 \qquad \text{for all } j \neq k$$ *(constant symmetry),*

with a real parameter ψ_0 . In view of (16) the last model states that the odds ratio of disagreement γ_{jk} is proportional to the inverse of its counterpart γ_{kj}

(CS)′ $$\gamma_{jk} = \exp(-\psi_0) / \gamma_{kj} \qquad \text{for all } j \neq k\,.$$

A natural extension allowing for additive effects of the row and column categories is

$$\psi_{jk} = \psi_0 + \tau_j + \nu_k \qquad \text{for all } j \neq k$$

with parameter vectors $\boldsymbol{\tau}, \boldsymbol{\nu} \in \mathbb{R}^J$ satisfying the constraints $\tau_1 = 0$, $\nu_1 = 0$. However the constraints upon ψ require $\boldsymbol{\tau} = \boldsymbol{\nu}$ thus reducing the above model to

(AS) $$\psi_{jk} = \psi_0 + \tau_j + \tau_k \qquad \text{for all } j \neq k$$ *(additive symmetry).*

Suppose now, we are dealing with an *ordinal* response variable with category labels 1,...,*J* reflecting the order of the response. The additive symmetry model (AS) may then be extended to allow for different contribution according to the order of the categories *j* and *k*

(OAS) $$\psi_{jk} = \psi_0 + \tau_{j\wedge k} + \nu_{j\vee k} \qquad \text{for all } j \neq k$$ *(order additive symmetry)*

with parameter vectors $\boldsymbol{\tau}, \boldsymbol{\nu} \in \mathbb{R}^J$ satisfying suitable constraints, e.g. $\tau_1 = \tau_J = 0$, $\nu_1 = \nu_J = 0$. A reparametrization $\tau'_j = \tau_j - \tau_{j-1}$ and $\nu'_j = \nu_j - \nu_{j-1}$ for $j > 1$ with $\tau'_1 = 0$ and $\nu'_1 = 0$ leads to

(OAS)′ $\quad \psi_{jk} = \psi_0 + \sum_{j' \leq j} \tau'_j + \sum_{k' \leq k} \nu'_k \qquad$ for all $j < k$

with the following interpretation of the parameters. $\psi_0 = \psi_{1J}$ represents the agreement for the extreme responses 1 and J, and the parameter $\tau'_j = \psi_{jk} - \psi_{(j-1)\,k}$ resp. $\nu'_k = \psi_{jk} - \psi_{j\,(k\text{-}1)}$ measures the change in agreement due to increasing the lower resp. decreasing the upper category. For a satisfactory overall agreement we expect ψ_0 to be positive (and large), but all τ'_j and ν'_k to be negative. Note that additive symmetry is obtained as a submodel for $\boldsymbol{\tau}' = \boldsymbol{\nu}'$.

If meaningful real-valued scores $z_1 \leq z_2 \ldots \leq z_J$ can be assigned to the response categories, which reflect their natural ordering, then further models based on the the *score distance* $d_{jk} = |z_k - z_j|$ for two categories j and k are available. The *linear distance symmetry model*

(LDS) $\quad \psi_{jk} = \psi_0 + \varrho\, d_{jk} \qquad$ for all $j \neq k \qquad$ *(linear distance symmetry)*

is an extension of (CS) with an additional parameter ϱ, and may be extended further to the *quadratic distance symmetry* model

(QDS) $\quad \psi_{jk} = \psi_0 + \varrho\, d_{jk} + \varrho_2\, d_{jk}^2 \qquad$ for all $j \neq k \qquad$ *(quadratic distance symmetry)*

or to a general polynomial distance symmetry model by adding further powers of d_{jk}. If in particular equally spaced scores are appropriate, i.e. $d_{jk} = d \cdot |j - k|$, then the *unit scores* $z_j := j$ (i.e. $d = 1$) may be used with no loss in generality. Now, viewing the score distance as a factor with J–1 levels, gives the *general distance symmetry model*

(GDS) $\quad \psi_{jk} = \psi_0 + \varrho_{\,|k-j|} \qquad$ for all $j \neq k \qquad$ *(general distance symmetry)*

with parameters $\varrho_1, \ldots, \varrho_{J-1}$ and the constraint $\varrho_1 = 0$. For unit scores, this is the most general model in which ψ_{jk} depends on j and k only through $d_{jk} = |j - k|$.

Further models may be designed along these ideas, in particular by passing to submodels or by extending two models to a common model containing both, e.g. (AS) and (QDS) may be extended to

(AS+QDS) $\quad \psi_{jk} = \psi_0 + \tau_j + \tau_k + \varrho\, d_{jk} + \varrho_2\, d_{jk}^2 \qquad$ for all $j \neq k$.

Although our prime interest focuses on models for agreement and hence for $\boldsymbol{\psi}$, it may also be desirable to specifiy the asymmetrical association, in particular to reduce the dimension $(J\text{–}1)(J\text{–}2)/2$ of the nuisance parameter $\boldsymbol{\Delta}$. Models for $\boldsymbol{\Delta}$ may be constructed along the same line as for $\boldsymbol{\psi}$, leading to the sequence of nested models

(ZA) $\quad \Delta_{jk} = 0 \qquad$ for all $j \neq k \qquad$ *(zero asymmetry),*

(CA) $\quad \Delta_{jk} = \Delta_0 \qquad$ for all $1 < j < k \qquad$ *(constant asymmetry),* (AA) $\Delta_{jk} = \Delta_0 + \tau_j^* + \tau_k^*$ for all $1<j<k$ *(additive asymmetry),*

(OAA) $\Delta_{jk} = \Delta_0 + \tau_{j\wedge k}^* + \nu_{j\vee k}^*$ for all $1<j<k$ *(ord. add. asymmetry).*

And in the presence of scores, further models for asymmetry are available

(LDA) $\Delta_{jk} = \Delta_0 + \varrho^* d_{jk}$ for all $1<j<k$ *(linear dist. asymmetry)*

(QDA) $\Delta_{jk} = \Delta_0 + \varrho^* d_{jk} + \varrho_2^* d_{jk}^2$ for all $1<j<k$ *(quadratic dist. asymmetry)*

(GDA) $\Delta_{jk} = \Delta_0 + \varrho^*_{|k-j|}$ for all $1<j<k$ *(general dist. asymmetry)*

Combining any of the above models for the symmetrical part $\boldsymbol{\psi}$ (including the saturated model) with one for the asymmetrical part $\boldsymbol{\Delta}$ yields a variety of log-linear models for the expected counts, some of which have been introduced before by other authors under different names (cf. Agresti 1990, Chapter 10). In particular, the model with zero asymmetry and saturated symmetry, i.e. omitting $\boldsymbol{\Delta}$ in $(SM)^*$, was introduced by Caussinus (1965) as the *quasi-symmetry model.* And submodels thereof, given by (AS) resp. (AS+QDS) with ϱ=0 are known as *quasi-independence* resp. *quasi-(uniform) association models.* Tanner and Young (1985b) already used the simple models (CS)+(ZA) and (CS)+(CA) as well as the combination of (GDS) with (ZA) resp. (GDA) which they called the model of *symmetric* resp. *asymmetric band disagreement.* However, models for ordinal responses containing order additive symmetry (OAS) or asymmetry (OAA) seem to be new.

Assuming that A and B are classifications of the same item by two observers, Darroch and McCloud (1986) derived the quasi-symmetry model from an additonal assumption and studied submodels given by the *indistinguishability* of two categories j and k, defined through $\psi_{jm} = \psi_{km}$ for all $m=1, .. J$ (which implies $\psi_{jk}=0$). They also gave an overall measure of agreement based only on the symmetric family $\boldsymbol{\psi}$, which is preferable to κ.

There is a convenient description for models by means of model formulae (McCullagh and Nelder 1989, 3.4), which is used by computer programs like GLIM (Francis et al. 1993). Let SYM and $ASYM$ denote the model formulae for the symmetrical and unsymmetrical components $\boldsymbol{\psi}$ and $\boldsymbol{\Delta}$ of association. The corresponding model formulae for the log-expectation (linear predictor) is then given by

$$\log \boldsymbol{\mu} = A + B + S.SYM + AS.ASYM$$

where S and AS represent continuous covariates given by

$$S(j,k) := -\tfrac{1}{2}\, I\{j \neq k\}\,, \quad AS(j,k) := I\{1<j\wedge k\} \cdot \mathrm{sgn}(k-j)\,\cdot$$

In the *saturated* model SYM is given by a factor $SatSym$, taking $\binom{J}{2}$ different values for all subsets $\{j,k\}$ of $\{1,\dots,J\}$ with $j\neq k$, and an additional value on the diagonal $j=k$. A convenient version is

$$SatSym(j,k) := 1 + I\{j \neq k\} \cdot \left[\, (\mathrm{j} \wedge \mathrm{k}) + \tfrac{1}{2}\,(\mathrm{j} \vee \mathrm{k} - 1)\,(\mathrm{j} \vee \mathrm{k} - 2)\,\right],$$

with levels ranging from 1 to $1+\binom{J}{2}$. And *ASYM* is given by a factor *SatAsym* taking $\binom{J-1}{2}$ different values for all subsets $\{j,k\}$ of $\{2,...,J\}$ with $j\neq k$, and an additional value in the remaining cases $j=k$, $j=1$, $k=1$, e.g.

$$SatAsym(j,k) := 1 + I\{j \neq k\,,\, 1 < j \wedge k\} \cdot \left[\, (\mathrm{j} \wedge \mathrm{k} - 1) + \tfrac{1}{2}\,(\mathrm{j} \vee \mathrm{k} - 2)\,(\mathrm{j} \vee \mathrm{k} - 3)\,\right],$$

with levels ranging from 1 to $1+\binom{J-1}{2}$.

The model of order additive symmetry, for example, can briefly be described by

(OAS)′ $\qquad SYM = MIN + MAX,$

with factors *MIN* and *MAX* representing the minimum and maximum of *A* and *B*. Leaving the remaining structure of $\boldsymbol{\mu}$ unspecified produces the model

$$\log \boldsymbol{\mu} = A + B + AS.\,SatAsym + S.(MIN + MAX)\ .$$

But restricting asymmetry through (OAA) leads to

(OA) $\qquad \log \boldsymbol{\mu} = A + B + (AS + S).(MIN + MAX) \qquad$ *(order additive association).*

5. Data Analysis

The analysis of an observed table $\mathbf{Y} = (Y_{\mathrm{jk}})$ using the above models for $\boldsymbol{\psi}$ and $\boldsymbol{\Delta}$ is straight forward, because these models belong to the family of log-linear models for the expected counts $\boldsymbol{\mu}$. Since log-linear models are treated in detail by many textbooks, e.g. Haberman (1974), Bishop, Fienberg and Holland (1975), McCullagh and Nelder (1989), Agresti (1990), only a few remarks are required here.

Because our models for $\boldsymbol{\psi}$ and $\boldsymbol{\Delta}$ restrict only the family $(\psi_{\mathrm{jkj'k'}})$ of log-cross product ratios, but not the remaining structure of $\boldsymbol{\mu}$ resp. $\boldsymbol{\eta}$, the following invariance property with respect to sampling models can easily be derived from the results of Haberman (1974, Theorems 2.4, 4.1).

Invariance property: *For any log-linear model specifying only the family $(\psi_{jkj'k'})$, the maximum likelihood estimate of $(\psi_{jkj'k'})$ and its asymptotic covariance matrix is the same in the Poisson, multinomial and row or column multinomial sampling model.*

This invariance property is inherited by the estimates of (ψ_{jk}) and (Δ_{jk}) of the symmetrical and asymmetrical component and the parameters specifying them in a particular model, e.g. ψ_0, τ, ν.... Hence asymptotic inference about (ψ_{jk}) and (Δ_{jk}) can be based on *any* of the four sampling models above, neglecting the actual distribution for the observed table (provided it belongs to one of sampling models above). This allows for a flexible analysis of the data according to the availabily of suitable software packages. For example GLIM may be applied on the basis of the Poisson model.

Example 1 (ctd.): Cytological and Pathological Classification of Uterine Cancer

Having ordered categories here, we first fitted the order additive symmetry model (OAS) leaving the remaining association arbitrary, i.e. assuming saturated asymmetry. The observed and expected counts together with their (scaled deviance) residuals given in table 5-1 indicate a satisfactory fit. The deviance $D = 7.25$ with 6 d.f. has a P-value of 29.9% (based on its asymptotic chi-squared distribution). The relevant estimates for the parameters of agreement are given in tables 5-2 and 5-3 together with their standard errors. Note that all entries are significantly non-zero.

Looking at submodels for the *symmetric* part, we find for the additive symmetry model (AS) (which does not exploit the ordering) a dramatic increase of the deviance by 330.1 with 3 d.f. which clearly rejects this submodel. On the other hand, restricting the *asymmetrical* part by a model does not change the picture for agreement. Even for the extreme submodel of zero asymmetry (ZA) (which increases the deviance only by 14.0 with 10 d.f.), the estimates for agreement and their standard errors are only slightly different from those in tables 5-2 and 5-3. However, the residuals have higher absolute values (an extreme value of 2.4 occouring in cell $j = 4$, $k = 1$). The same picture emerges for the submodels (CA), (AA) and (OAA), the residuals of the latter model are comparable to those in table 5-1. We conclude, that for this data set the estimates of agreement parameters are fairly robust under model changes for the remaining asymmetric association.

Cytological classification (j)		Pathological classification (k) III DL (1)	III DM (2)	IV AS (3)	IV AC (4)	IV B (5)	V (6)
III DL	(1)	53.00	42.00	15.32	7.12	0.30	2.26
III DM	(2)	35.00	112.00	83.68	79.48	2.13	5.70
IV AS	(3)	5.32	25.68	108.00	77.39	1.90	7.70
IV AC	(4)	6.12	14.48	57.39	293.00	15.66	43.34
IV B	(5)	0.30	1.13	4.90	36.66	14.00	49.00
V	(6)	0.26	2.70	7.70	84.34	14.00	529.00
III DL	(1)	0.00	0.00	− 0.34	0.68	− 0.78	− 0.18
III DM	(2)	0.00	0.00	0.14	− 0.05	− 0.87	0.12
IV AS	(3)	− 0.60	0.26	0.00	− 0.16	0.73	0.11
IV AC	(4)	0.72	− 0.13	− 0.18	0.00	0.09	− 0.05
IV B	(5)	− 0.78	− 1.51	0.48	0.06	0.00	0.00
V	(6)	− 0.72	0.18	0.11	− 0.04	0.00	0.00

Table 5-1: Fitted counts (top) and scaled deviance residuals (bottom) for the order additive symmetry model (with saturated asymmetry) in example 1.

Parameter	ψ_0	τ'_2	τ'_3	τ'_4	τ'_5	ν'_2	ν'_3	ν'_4	ν'_5
Estimate ± S. E.	10.78 ± 0.54	− 2.53 ± 0.43	− 1.39 ± 0.32	− 3.12 ± 0.41	− 1.37 ± 0.45	− 2.86 ± 0.47	− 1.62 ± 0.34	− 3.12 ± 0.52	− 1.78 ± 0.42

Table 5-2: Estimates with standard errors in example 1 for parameters of agreement in the order additive symmetry model (OAS)´ with saturated asymmetry.

Cytological classification (j)		Pathological classification (k) III DL (1)	III DM (2)	IV AS (3)	IV AC (4)	IV B (5)	V (6)
III DL	(1)						
III DM	(2)	1.40 ± 0.28					
IV AS	(3)	4.25 ± 0.43	1.73 ± 0.26				
IV AC	(4)	5.88 ± 0.43	3.35 ± 0.28	1.96 ± 0.20			
IV B	(5)	9.00 ± 0.64	6.48 ± 0.54	5.09 ± 0.52	1.97 ± 0.39		
V	(6)	10.78 ± 0.54	8.26 ± 0.42	6.87 ± 0.39	3.75 ± 0.20	2.38 ± 0.41	

Table 5-3: Fitted agreement $\hat{\psi}_{jk}$ for $j > k$ with standard errors (±S.E.) in example 1 for the order additive symmetry model (OAS)´ with saturated asymmetry.

Example 2 (ctd.): Number of Pregnancies Reported by the Mother and Father

Velema et al. (1991) fitted a model proposed by Agresti (1988), which can be described in our notation by zero asymmetry (ZA) and the submodel of quadratic distance symmetry (QDS) (with unit scores) given by ϱ=0 , i.e.

(19) $$\psi_{jk} = \psi_0 + \varrho_2 (j-k)^2 , \qquad \Delta_{jk} = 0 \qquad \text{for all } j \neq k$$

Although the deviance $D = 20.5$ with 14 d.f. seems to indicate an acceptable fit, there are at least two larger residuals. The fit can not be improved significantly as long as we assume zero asymmetry: even the quasi-symmetric model (with saturated symmetry) decreases the deviance only by 4.58 with 8 d.f. However, allowing for constant asymmetry (CA) reduces the deviance by 10.26 (with 1 d.f.) to 10.22 (with 13 d.f.), thus greatly improving the fit. The fitted counts and their residuals for this model are given in table 5-4 and confirm a satisfactory fit. The estimates for the parameters of interest are given below together with their standard errors

$$\hat{\psi}_0 \pm \text{S.E.} = 5.12 \pm 0.79 , \qquad \hat{\varrho}_2 \pm \text{S.E.} = 0.20 \pm 0.12 \qquad \hat{\Delta}_0 \pm \text{S.E.} = 1.74 \pm 0.57 ,$$

and the resulting estimates $\hat{\psi}_{jk}$ for agreement are contained in table 5-5. Note, that ϱ_2 is not significantly different from 0 (on the 5% level), but the remaining parameters (and all ψ_{jk}) are.

In fact, the submodel with constant symmetry (CS) also provides an acceptable fit as judged by its residuals and deviance $D = 13.18$ with 14 d.f.

Pregnancies Reported by Mother	Pregnancies Reported by Father 1	2	3	4	5 +
1	61.17	17.64	4.81	2.33	1.05
2	0.58	33.93	4.96	2.93	1.60
3	0.65	0.63	18.79	5.96	3.97
4	0.28	0.33	0.16	10.48	3.75
5 +	0.33	0.47	0.28	0.30	21.63
1	0.11	− 0.40	0.52	− 0.22	− 0.05
2	− 1.07	0.18	− 0.45	0.04	0.30
3	− 1.14	0.43	− 0.18	0.79	− 0.51
4	1.05	0.94	− 0.57	− 0.47	0.13
5 +	− 0.81	− 0.97	1.05	− 0.77	0.08

Table 5-4: Fitted counts (top) and scaled deviance residuals (bottom) in example 2 for the model given by (19) and constant asymmetry (CA).

Pregnancies Reported by Mother (j)	Pregnancies Reported by Father (k) 1	2	3	4	5 +
1					
2	5.32 ± 0.71				
3	5.91 ± 0.56	5.32 ± 0.71			
4	6.89 ± 0.80	5.91 ± 0.56	5.32 ± 0.71		
5 +	8.26 ± 1.55	6.89 ± 0.80	5.91 ± 0.56	5.32 ± 0.71	

Table 5-5: Fitted agreement $\hat{\psi}_{jk}$ for $j > k$ with standard errors (±S.E.) in example 2 for the model given by (19) and constant asymmetry (CA).

6. Model Extensions

The models considered so far can easily be extended when in addition to the responses A and B, further covariables are observed, provided they are *discrete* (which can always be achieved by grouping continuous variables). Since all additional variables can in principle be combined into a *single* discrete multivariate variable, it is sufficient to illustrate the model extensions for just one further variable C. Suppose that C takes S different values, which are for simplicity identified with the numbers 1,...,S giving rise to S strata $\{C=1\}$,..., $\{C=S\}$ on which C is constant. The previous notation will be extended by a subscript s=1,...,S referring to stratum s, e.g. Y_{jks} is the oberved count for the event $\{A=j,\ B=k,\ C=s\}$, $\mu_{jks} = \mathscr{E}(Y_{jks})$ the expected count etc.

The saturated model for the log-expectation $\boldsymbol{\eta}$ may be parametrized using (SM) for each stratum $s=1,\dots,S$

(SSM) $\quad \eta_{\mathrm{jks}} = \alpha_{0\mathrm{s}} + \alpha^{\mathrm{A}}_{\mathrm{js}} + \alpha^{\mathrm{B}}_{\mathrm{ks}} + \Delta_{\mathrm{jks}} - \frac{1}{2}\psi_{\mathrm{jks}}$ *(saturated stratum model),*

with the obvious constraints corresponding to (12) for each stratum s. As before our main interest focuses on models for the symmetrical part (ψ_{jks}) and possibly on (Δ_{jks}) leaving $(\alpha_{0\mathrm{s}})$, $(\alpha^{\mathrm{A}}_{\mathrm{js}})$, and $(\alpha^{\mathrm{B}}_{\mathrm{ks}})$ completely arbitrary to maintain the invariance property with respect to the sampling models. Concerning the dependence of ψ_{jks} on stratum s, the most simple model assumes independence from s

(CSS) $\quad \psi_{\mathrm{jks}} = \beta_{0\mathrm{jk}}$ for all s and $j \neq k$ *(constant stratum symmetry).*

with a parameter matrix $\boldsymbol{\beta_0} \in \mathbb{R}^{\mathrm{J}\times\mathrm{J}}$ satisfying the constraints $\beta_{0\mathrm{jk}} = \beta_{0\mathrm{kj}}$.

Strata are often defined by values of an additonal observed covariable x (like calendar year, age etc.) which is either discrete itself or discretized by appropriate rounding or classification. If x_{s} denotes the observed value of x in stratum s, then a *linear stratum symmetry* model for $\boldsymbol{\psi}$ is available

(LSS) $\quad \psi_{\mathrm{jks}} = \beta_{0\mathrm{jk}} + \beta_{\mathrm{jk}}\,\mathrm{x}_{\mathrm{s}}$ for all s and $j \neq k$ *(linear stratum symmetry).*

The parameter matrix $\boldsymbol{\beta_0} \in \mathbb{R}^{\mathrm{J}\times\mathrm{J}}$ gives the *baseline* symmetry corresponding to $x=0$ and the *slope* $\boldsymbol{\beta} \in \mathbb{R}^{\mathrm{J}\times\mathrm{J}}$ measures the increase of symmetry per unit increase of x, (both with the constraints $\beta_{0\mathrm{jk}} = \beta_{0\mathrm{kj}}$ and $\beta_{\mathrm{jk}} = \beta_{\mathrm{kj}}$). A generalization of this model to a *vector* $\mathbf{x}$ of covariates is obvious.

The linear stratum model which contains (CSS), represents a log-linear model for the expected counts. A number of interesting submodels are obtained by imposing additional structure on the baseline $\boldsymbol{\beta_0}$ and the slope $\boldsymbol{\beta}$ along the lines of section 4, e.g. order additive baseline and slope

(OABS) $\quad \beta_{0\mathrm{jk}} = \psi_{00} + \tau_{0\,(\mathrm{j}\wedge\mathrm{k})} + \nu_{0\,(\mathrm{j}\vee\mathrm{k})}$ for all s and $j \neq k$ *(ord. add. baseline sym.)*

(OASS) $\quad \beta_{\mathrm{jk}} = \psi_0 + \tau_{\mathrm{j}\wedge\mathrm{k}} + \nu_{\mathrm{j}\vee\mathrm{k}}$ for all s and $j \neq k$ *(ord. add. slope sym.)*

It may also be desirable (in particular in the presence of many strata), to restrict the dimension of asymmetrical part $\boldsymbol{\Delta}$ by additional model assumptions, like *constant strata asymmetry*

(CSA) $\quad \Delta_{\mathrm{jks}} = \delta_{0\mathrm{jk}}$ for all s and $j \neq k$ *(constant strata asymmetry),*

or more generally *linear stratum asymmetry*

(LSA) $\Delta_{\mathrm{jks}} = \delta_{0\mathrm{jk}} + \delta_{\mathrm{jk}}\,\mathrm{x}_{\mathrm{s}}$ all s and $j{\neq}k$ *(linear stratum asymmetry).*

The *baseline asymmetry* $\boldsymbol{\delta}_{\mathbf{0}} \in \mathrm{I\!R}^{\mathrm{J}\times\mathrm{J}}$ and *slope of asymmtery* $\boldsymbol{\delta} \in \mathrm{I\!R}^{\mathrm{J}\times\mathrm{J}}$ (which must satisfy the constraints $\delta_{0\mathrm{jk}} = -\delta_{0\mathrm{kj}}$ and $\delta_{\mathrm{jk}} = -\delta_{\mathrm{kj}}$) may again be restricted to one of the models for asymmetry in section 4, e.g. order additive baseline and slope.

Example 1 (ctd.): Cytological and Pathological Classification of Uterine Cancer

In order to investigate whether agreement has changed over time, the data was divided into $S=6$ strata according to two-year time intervalls 1972-73, ..., 1982-83. The observed counts are given in table 6-1. The covariate *time* in stratum s was taken as $x_{\mathrm{s}} = 2(s-1)$ to represent the number of years elapsed since the beginning of the study. We first assume a linear trend for the total association, i.e. linear stratum symmetry (LSS) in combination with linear stratum asymmetry (LSA), which leads to the following log-linear model

(20) $$\eta_{\mathrm{jks}} = \alpha_{0\mathrm{s}} + \alpha^{\mathrm{A}}_{\mathrm{js}} + \alpha^{\mathrm{B}}_{\mathrm{ks}} + \delta_{0\mathrm{jk}} - \tfrac{1}{2}\beta_{0\mathrm{jk}} + \left(\delta_{\mathrm{jk}} - \tfrac{1}{2}\beta_{\mathrm{jk}}\right)\mathrm{x}_{\mathrm{s}}\,,$$

with obviuos constraints for the parameters. Since $x_1 = 0$, the baseline symmetry $\boldsymbol{\beta}_0$ in this model gives the agreement in the first stratum, i.e. $\psi_{\mathrm{jk1}} = \beta_{\mathrm{jk0}}$, and the slope $\boldsymbol{\beta}$ of symmetry represents the increase of agreement per year, i.e. $\psi_{\mathrm{jk(s+1)}} - \psi_{\mathrm{jks}} = 2\beta_{\mathrm{jk}}$. Similiarly, the baseline asymmetry $\boldsymbol{\delta}_0$ is the asymmetry in first stratum and the slope $\boldsymbol{\delta}$ of asymmetry measures the increase per year. Restricting in (20) the parameters $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}$ for symmetry by (OABS) and (OASS) respectively, but leaving the asymmetry parameters $\boldsymbol{\delta}_0$ and $\boldsymbol{\delta}$ arbitrary, gives what will be called the *basic model.*

The basic model provides a satisfactory fit (based on an eximination of the residuals) and has a deviance D=107.51 with 112 d.f. The estimates $\hat{\beta}_{0\,\mathrm{jk}}$ resp. $\hat{\beta}_{\mathrm{jk}}$ for the baseline resp. slope of agreement and their standard errors are given in table 6-2. Not surprisingly, all estimates are positive, and the majority of them (marked by a star * in the table) are *significantly* positive, using Wald's one sided test on the 5% level. In particular, the baseline $\hat{\beta}_{0\mathrm{jk}}$ was not significant only for two adjacent pairs. And the slope was not signifikant for all pairs containing category Pap IV B, as well as for two "adjacent" pairs (i.e. with adjacent categories) indicating a non significant increase over time for these pairs. However, the estimates $\hat{\psi}_{\mathrm{jks}}$ (not tabled here) obtained from (LSS) show a steady increase with s, and are significantly positive for s>2 and all j, k. In summary, the basic model reveals positive agreement increasing over time for all pairs, being significantly positive for most pairs.

Cytological classification	Pathological classification III DL	III DM	IV AS	IV AC	IV B	V	Total
1972 – 1973							
III DL	2	0	1	1	0	0	4
III DM	0	0	0	1	1	0	2
IV AS	1	2	2	5	1	3	14
IV AC	2	4	6	51	9	13	85
IV B	0	0	1	5	7	31	44
V	0	1	1	28	10	139	179
Total	5	7	11	91	28	186	328
1974 – 1975							
III DL	4	4	0	0	0	0	8
III DM	2	18	7	16	0	2	45
IV AS	0	3	4	7	1	1	16
IV AC	2	1	7	55	1	9	75
IV B	0	0	1	5	1	9	16
V	0	0	2	23	0	103	128
Total	8	26	21	106	3	124	288
1976 – 1977							
III DL	8	15	6	3	0	0	32
III DM	7	17	17	26	0	1	68
IV AS	1	7	16	10	0	0	34
IV AC	1	2	6	55	3	4	71
IV B	0	0	1	10	2	7	20
V	0	0	0	14	1	90	105
Total	17	41	46	118	6	102	330
1978 – 1979							
III DL	18	14	6	2	0	0	40
III DM	11	19	32	18	0	1	81
IV AS	1	2	20	19	0	2	44
IV AC	2	4	18	57	1	12	94
IV B	0	0	2	8	2	1	13
V	0	1	2	8	1	60	72
Total	32	40	80	112	4	76	344
1980 – 1981							
III DL	9	4	1	2	0	1	17
III DM	9	26	16	11	0	0	62
IV AS	1	5	23	17	0	1	47
IV AC	1	3	10	39	0	4	57
IV B	0	0	0	3	2	0	5
V	0	0	2	4	1	73	80
Total	20	38	52	76	3	79	268
1982 – 1983							
III DL	12	5	0	1	0	1	19
III DM	6	32	13	7	0	2	60
IV AS	0	8	43	18	1	1	71
IV AC	0	0	9	36	2	1	48
IV B	0	0	1	6	0	1	8
V	0	1	1	7	1	64	74
Total	18	46	67	75	4	70	280

Table 6-1: Observed counts of cytological and pathological classifications for each of 6 two-year time periods in example 1.

Cytological classification (j)		III DL (1)	III DM (2)	IV AS (3)	IV AC (4)	IV B (5)	V (6)
		Pathological classification (k)					
III DL	(1)						
III DM	(2)	0.46 ± 0.78					
IV AS	(3)	2.14 * ± 1.03	0.04 ± 0.65				
IV AC	(4)	4.35 * ± 0.96	2.26 * ± 0.60	1.29 * ± 0.50			
IV B	(5)	7.50 * ± 1.21	5.40 * ± 0.97	4.44 * ± 0.91	2.00 * ± 0.57		
V	(6)	8.52 * ± 1.09	6.42 * ± 0.82	5.46 * ± 0.75	3.02 * ± 0.30	1.31 * ± 0.50	
III DL	(1)						
III DM	(2)	0.30 ± 0.23					
IV AS	(3)	0.77 * ± 0.33	0.51 * ± 0.18				
IV AC	(4)	0.65 * ± 0.32	0.38 * ± 0.20	0.16 ± 0.14			
IV B	(5)	0.59 ± 0.43	0.32 ± 0.35	0.09 ± 0.32	0.04 ± 0.24		
V	(6)	0.94 * ± 0.38	0.67 * ± 0.28	0.44 * ± 0.25	0.39 * ± 0.13	0.93 * ± 0.28	

Table 6-2: Estimated baselines $\hat{\beta}_{0\,jk}$ (top) resp. slopes $\hat{\beta}_{jk}$ (bottom) of agreement with standard errors (±S.E.) for the basic model (20) in example 1. A *significantly positive* entry (based on Wald's one sided test on the 5% level) is marked by a star *.

Although the basic model gives a satisfactory description of the data, it it instructive to investigate whether the above conclusions for agreement depend on the chosen model for the remaining association Δ_{jks}. Passing to the enlarged model which leaves Δ_{jks} completely arbitrary (by abandoning the condition (LSA) of linear stratum asymmetry) yields a decrease of the deviance by 27.50 with 40 d.f. which is not significant. Hence there is no need for a model extension concerning asymmetry. On the other hand, the submodel of the basic model given by $\boldsymbol{\delta} = 0$, i.e. having constant stratum asymmetry (CSA), increases the deviance by 19.13 with 10 d.f., leading to a rejection of the submodel ($p = 3.9\%$).

Another way to reduce the the dimension of asymmetry in our basic model is to restrict the baseline $\boldsymbol{\delta}_0$ resp. slope $\boldsymbol{\delta}$ by assuming an order additive, additive or even constant structure for the baseline resp. slope in the same way we restricted Δ in section 4 by (OAA), (AA) or (CA). Passing from the basic model to the submodel with an *order additive* baseline $\boldsymbol{\delta}_0$ and slope $\boldsymbol{\delta}$, the deviance increases by 2.72 with 6 d.f., which is not significant. The estimates $\hat{\beta}_{0jk}$, $\hat{\beta}_{0jk}$ and their standard errors under this submodel differ only slightly from those under the basic model, thus leading to the same conclusions concerning agreement. The same picture emerges if we pass to further submodels with *additive* or even *constant* baseline $\boldsymbol{\delta}_0$ and slope $\boldsymbol{\delta}$, although the fit for the latter (but not the former) submodel deteriorates singificantly. In summary, the chosen model for asymmetry is not crucial concerning the conclusions for agreement, as long as we allow for a (linear) change over time.

References

Agresti, A. (1988). *A model for agreement between ratings on an ordinal scale*. Biometrics 44, 539-548.

Agresti, A. (1990). *Categorical data analysis*. New York, Wiley.

Bishop, Y.M.M., Fienberg, S.E. & Holland, P.W. (1975). *Discrete mutlivariate analysis : Theory and practice*. The MIT Press, Cambridge, Massachusettes, and London, England.

Cohen, J. (1960). *A coefficient of agreement for nominal scales*. Educational and Psychological Measurement 20, 37-46.

Darroch, J.N. and McCloud, P.I. (1986). *Category distinguishability and observer agreement*. Austr. J. Statist. 28, 371-388.

Francis, B., Green, M., Payne, C. (Eds.) (1993). *The GLIM system. Release 4 manual.* Clarendon Press, Oxford.

Goodman, L.A. (1979). *Simple models for the analysis of association in cross-classifications having ordered categories*. J. Amer. Statist. Ass. 74, 537-552.

Goodman, L.A. (1985). *The analysis of cross-classified data having ordered and/or unordered categories : Association models, correlation models, and asymmetrie models for contingency tables*. Ann. Statist. 13, 10-69.

Haberman, S.J. (1974). *The analysis of frequency data*. The University of Chicago Press, Chicago and London.

McCullagh, P. & Nelder, J.A. (1989). *Generalized linear models (second edition)*. Chapman and Hall, London, New York.

Papanicolaou, G.N and Traut, H.F. (1943). *Diagnosis of uterine cancer by the vaginal smear*. The Commonwealth Fund, New York..

Plackett, R.L. (1974). *The analysis of categorical data*. Griffin, London.

Velema, J.P., Blettner, M., Restrepo, M. and Muños, N. (1991). *The evaluation of agreement by means of log-linear models: proxy interviews on reproductive history among floriculture workers in Columbia*. Epidemiology 2, 107-115.

Tanner, M.A. and Young, M.A. (1985a). *Modeling agreement among raters*. J. Amer. Statist. Ass. 80, 175-180.

Tanner, M.A. and Young, M.A. (1985b). *Modeling ordinal scale disagreement*. Psych. Bull. 98, 408-415.

Adress for correspondance: Prof. Dr. Gerhard Osius
Institut für Statistik, FB3, Universität Bremen
Postfach 33 04 40
D–28334 Bremen, Germany
Fax: +49+421–2184020

Date: 15-Nov-1993 (printed edition)
23-Jan-2001 (PDF-File, with minor corrections)

Universität Bremen Mathematik-Arbeitspapiere

Stand November 2000 ISSN 0173 - 685 X

Vertrieb der Hefte 4, 14, 23, 26 durch Universitätsbuchhandlung, Bibliothekstr. 3, D-28359 Bremen. Vertrieb der übrigen Hefte (soweit nicht vergriffen) durch die Autoren oder FB 3 Mathematik/Informatik Universität Bremen, Postfach 330440, D-28334 Bremen.

1. Ulrich Krause (1976): Strukturen in unendlichdimensionalen konvexen Mengen, 74 S.
2. Fritz Colonius, Diederich Hinrichsen (1976): Optimal control of heriditary differential systems. Part I, 66 S.
3. Günter Matthiessen (1976): Theorie der heterogenen Algebren, 88 S.
4. H. Wolfgang Fischer, Jens Gamst, Klaus Horneffer (1976): Skript zur Analysis, Band 1 (11. Auflage 2000), 286 S.
5. Wolfgang Schröder (1977): Operator-algebraische Ergodentheorie für Quantensysteme, 59 S.
6. Rolf Röhrig, Michael Unterstein (1977): Analyse multivariabler Systeme mit Hilfe komplexer Matrixfunktionen, 216 S.
7. Horst Herrlich, Hans-Eberhard Porst, Rudolf-Eberhard Hoffmann, Manfred Bernd Wischnewsky (1976): Nordwestdeutsches Kategorienseminar, 193 S.
8. Fritz Colonius, Diederich Hinrichsen (1977): Optimal Control of Hereditary Differential Systems. Part II: Differential State Space Description, 36 S.
9. Ludwig Arnold (1977): Differentialgleichungen und Regelungstheorie, 185 S.
10. Rudolf Lorenz (1977): Iterative Verfahren zur Lösung großer, dünnbesetzter symmetrischer Eigenwertprobleme, 104 S.
11. Konrad Behnen, Hans-Peter Kinder, Gerhard Osius, Rüdiger Schäfer, Jürgen Timm (1977): Dose-Response-Analysis, 206 S.
12. Hans-Friedrich Münzner, Dieter Prätzel-Wolters (1978): Minimalbasen polynomialer Moduln, Strukturindizes und BRUNOVSKY-Transformationen, 53 S.
13. Konrad Behnen (1978): Vorzeichen-Rangtests mit Nullen und Bindungen, 53 S.
14. H. Wolfgang Fischer, Jens Gamst, Klaus Horneffer, Eberhard Oeljeklaus (1978): Skript zur Linearen Algebra, Band 1 (13. Auflage 2000), 249 S.
15. Günter Ludyk (1978): Abtastregelung zeitvarianter Einfach- und Mehrfachsysteme, 54 S.
16. Momme Johs Thomsen (1977): Zur Theorie der Fastalgebren, 146 S.
17. Klaus Horneffer, Horst Diehl (1978): Modellrechnungen zur anaeroben Reduktionskinetik des Cytochroms P-450, 34 S.
18. Horst Herrlich, Rudolf-Eberhard Hoffmann, Hans-Eberhard Porst, Manfred Bernd Wischnewsky (1979): Structure of Topological Categories, 252 S.
19. Hans-Friedrich Münzner, Dieter Prätzel-Wolters (1979): Geometric and moduletheoretic approach to linear systems. Part I: Basic categories and functors, 28 S.
20. Hans-Friedrich Münzner, Dieter Prätzel-Wolters (1979): Geometric and moduletheoretic approach to linear systems. Part II: Moduletheoretic characterization and reachability, 28 S.
21. Eckart Beutler, Hans Kaiser, Günter Matthiessen, Jürgen Timm (1979): Biduale Algebren, 165 S.
22. Horst Diehl, Detlef Harbach, Jürgen Timm (1980): Planung und Auswertung von Atomabsorptions-Spektrometrie-Untersuchungen mit der Additionsmethode, 44 S.
23. H. Wolfgang Fischer, Jens Gamst, Klaus Horneffer (1981): Skript zur Analysis, Band 2 (6. Auflage 1996), 299 S.
24. Horst Herrlich (1981): Categorical Topology 1971-1981, 105 S.
25. Horst Herrlich, Rudolf-Eberhard Hoffmann, Hans-Eberhard Porst, Manfred Bernd Wischnewsky (1981): Special Topics in Topology and Category Theory, 108 S.

26. H. Wolfgang Fischer, Jens Gamst, Klaus Horneffer (1984): Skript zur Linearen Algebra, Band 2 (7. Auflage 1999), 257 S.
27. Rudolf-Eberhard Hoffmann (1982): Continuous Lattices and Related Topics, 314 S.
28. Horst Herrlich, Rudolf-Eberhard Hoffmann, Hans-Eberhard Porst (1987): Workshop on Category Theory, 169 S.
29. Harald Boehme (1987): Zur Berufspraxis des Diplommathematikers, 16 S.
30. Jürgen Timm (1986): Mathematische Modelle der Dosis-Wirkungsanalyse bei den experimentellen Untersuchungen der Arbeitsgruppe zur karzinogenen Belastung des Menschen durch Luftverunreinigung, 65 S.
31. Dieter Denneberg (1988): Mathematik für Wirtschaftswissenschaftler. I. Lineare Algebra, 97 S.
32. Peter E. Crouch, Diederich Hinrichsen, Anthony J. Pritchard, Dietmar Salamon (1988, previous edition University of Warwick 1981): Introduction to Mathematical Systems Theory, 244 S.
33. Gerhard Osius (1989): Some Results on Convergence of Moments and Convergence in Distribution with Applications in Statistics, 27 S.
34. Dieter Denneberg (1989): Verzerrte Wahrscheinlichkeiten in der Versicherungsmathematik, Quantilsabhängige Prämienprinzipien, 24 S.
35. Eberhard Oeljeklaus (1989): Birational splitting of homogeneous Albanese bundles, 30 S.
36. Gerhard Osius, Dieter Rojek (1989): Normal Goodness-of-Fit Tests for Parametric Multinomial Models with Large Degrees of Freedom, 38 S.
37. Dieter Denneberg (1990): Mathematik zur Wirtschaftswissenschaft. II. Analysis, 59 S.
38. Ulrich Krause, Cornelia Zahlten (1990): Arithmetik in Krull monoids and the cross number of divisor class groups, 29 S.
39. Dieter Denneberg (1990): Subadditive Measure and Integral, 39 S.
40. Ulrich Krause, Peter Ranft (1991): A limit set trichotomy for monotone nonlinear dynamical systems, 31 S.
41. Angelika van der Linde (1992): Statistical analyses with splines: are they well defined? 22 S.
42. Dieter Denneberg (1992): Lectures on non-additive measure and integral (new edition: Non-additive measure and integral. TDLB 27, Kluwer Academic, Dordrecht (1994)), 114 S.
43. Gerhard Osius (1993): Separating Agreement from Association in Log-linear Models for Square Contingency Tables With Applications, 23 S.
44. Hans-Peter Kinder, Friedrich Liese (1995): Bremen-Rostock Statistik Seminar, 5. - 7. März 1992, 110 S.
45. Dieter Denneberg (1995): Extension of a measurable space and linear representation of the Choquet Integral, 30 S.
46. Dieter Denneberg, Michael Grabisch (1996): Shapley value and interaction index, 20 S.
47. Angelika Bunse-Gerstner, Heike Faßbender (1996): A Jacobi-like method for solving algebraic Riccati equations on parallel computers, 24 S.
48. Hans-Eberhard Porst editor (1997): Categorical methods in algebra and topology - a collection of papers in honour of Horst Herrlich, 498 S.
49. Angelika van der Linde, Gerhard Osius (1997): Estimation of nonparametric risk functions In matched case-control studies, 28 S.
50. Angelika van der Linde (1997): Estimating the smoothing parameter in generalized spline-based regression, 46 S.
51. Ursula Müller, Gerhard Osius (1998): Asymptotic normality of goodness-of-fit statistics for sparse Poisson data, 15 S.
52. Ursula Müller (1999): Nonparametric regression for threshold data, 18 S.
53. Gerhard Osius (2000): The association between two random elements – A complete characterization in terms of odds ratios, 32 S.