DATA ANALYSIS AND CONDITIONAL INDEPENDENCE STRUCTURES

Nanny Wermuth

Psychologisches Institut
Johannes Gutenberg-Universität
Postfach 3980

D-6500 Mainz
Federal Republic of Germany

1. HISTORY AND APPLICATIONS OF MODELS FOR ANALYSING STRUCTURES

In medicine, psychology, marketing research and sociology one tries to find determinants for human behaviour and reactions. Typical questions are: which conditions enhance the development of a particular ability, what causes a change in reaction patterns, what kind of background or anamnesis is favourable to a special phenomenon? Observations from cross-sectional studies are used to evaluate empiral evidence for such reflections. The observations then concern systems of variables that are characteristics, achievements or attitudes of persons. With models for analysing structures one intends to describe and explain the relations between the observed variables. These types of models have a long tradition (compare: Wright 1923, Spearman 1926, Simon 1957, Blalock 1971, Goldberger and Duncan 1973, Goodman 1978), a longer tradition than the theory of multivariate statistical analysis (compare: Anderson 1958). Thus, the aims and first methods for analysing structures had been formulated a long time before problems and answers in multivariate statistics became well-known.

Path analysis as proposed by Wright (1921, 1934) represents one of the early attempts to meet the need for methods to analyse structures. The purposes that Wright stated at the time have remained the same, only the terminology has changed and a more realistic assessment of what can be achieved with such analyses, is possible nowadays. For a given system of variables one wants with path analysis

(1) to classify some of the relations between variable pairs as direct, others as only indirect, i.e. explainable through the remaining direct relations in the system,

(2) to decide on whether the observed relations between the variables can be well approximated by a simplifying structure, and

(3) to evaluate the relative importance of direct relations.

Translated into statistical terminology one can interpret these tasks as follows:

(1) means to specify a multivariate model, more precisely a system of dependencies and associations, in which conditional independencies are postulated for some of the variable pairs,

(2) corresponds to evaluating and testing the goodness-of-fit of such models, and

(3) amounts to defining and estimating standardized versions of measures for partial associations and partial dependencies.

In: Bulletin of the International Statistical Institute, Proceedings 45th Session, Amsterdam 1985, Vol 1. ISI: Voorburg

In the 1920's, it was still believed that finding a well-fitting model i.e. a simplyfing structure in terms of only few direct relations is to establish evidence for corresponding causal dependencies. This view is no longer tenable. If a statistical model fits some observations well, then the empiral evidence only does not (yet) contradict a causal interpretation that is compatible with the statistical model and is derived from subject matter knowledge. Some of the reasons are that

- a cross-sectional investigation lacks one of the ingredients most important to studying cause-effect relations: the possibility for observing a change in time, i.e. the effect of a hypothesized cause, and
- each relation classified as direct in one given system of variables can turn into an indirect one, once a system with some additional variables is analysed.

On the other hand one should be able to interpret the classification as an indirect relation as evidence against a causal dependence. This can be done, provided an indirect relation in the model is in fact equivalent to a conditional independence and the observable variables satisfy the assumptions of the statistical model. The development of the graphical chain models (Lauritzen and Wermuth 1984) represents one important step towards more adaequate analyses in the above sense. In these models certain missing interactions are equivalent to conditional independencies and a joint distributions is specified for discrete and continuous random variables. Thus, qualitative characteristics are modelled by discrete random variables and quantitative ones by continuous random variables. In the case one has a system with variables of only one kind the models specialise to well-known methods for analysing structures, e.g. to path analysis, and to covariance selection (Dempster 1972) in the case of only quantitative characteristics and to modified path analysis (Goodman 1973, Wermuth and Lauritzen 1983) and to graphical loglinear models (Birch 1963, Bishop, Fienberg and Holland 1975, Darroch, Lauritzen and Speed 1980) for systems containing only qualitative characteristics.

Typical examples for situations, in which one wishes to investigate the structure of a system of variables are provided e.g. by studies on achievement, stress and anxiety (compare Hodapp 1982, Lazarus 1966). There, one distinguishes between variables capturing reactions of a physical kind (like systolic and diastolic blood pressure, puls rate) and of a psychical kind (like nervousness, tiredness), furthermore between personality dependent variables (like anxiety, emotional stability and coping mechanisms) and variables charactising a given situation (like stress at work, hiercharchical status and other working conditions). Psychological theory permits to specify a complex system of dependencies and associations. For instance, the personality dependent variables are considered to be response variables of the working conditions and determinants for the reactive variables. For some of the jointly dependent variables like diastolic and systolic blood pressure one cannot assign a direction to the relationship, i.e. one can only postulate an association. A corresponding statistical model then contains (joint) conditional distributions with discrete and continuous variables, i.e. distributions like the one used for graphical chain models. In general, such models will neither be generalised linear models (McCullagh and Nelder, 1983) nor be linear structural relation models (Jöreskog and Sörbom 1978) even though there is an overlap in special cases.

The conditional Gaussian (CG-) distributions considered by Lauritzen and Wermuth for the graphical chain models have a property, which is extremely important for corresponding data analyses: they define fairly simple transformations of the observations, which have to reflect conditional independencies of variable pairs, if there are any. These transformations are the maximum-likelihood estimates of the so-called natural parameters in the saturated case, i.e. in the situation without restrictions on the natural parameters. Thus, for these distributions one gets guide-lines for how to read off from the data independencies that are compatible with the observations, i.e. the models tell how to detect simplifying structures in the data.

In the remainder of this paper we illustrate with examples how conditional independencies are reflected in data of CG-distributions. To this end we first describe the graphical representations of the independence structures and of our models (section 2), then discuss the implications of two selected structures for different constellations of quantitative and qualitatitve variables (section 3) and finally (section 4) point of possible pitfalls.

2. GRAPHS FOR CONDITIONAL INDEPENDENCE STRUCTURES AND GRAPHICAL CHAIN MODELS:

In the graphs considered here a vertex drawn as a cross represents a variable and lines and arrows between vertices describe associations and dependencies. Boxes indicate the dependence chain $(D_1, \ldots, D_T)$, i.e. a partitioning of the vertex set. A chain element D_t for $t < T$ specifies a set of jointly dependent variables, the regressands: $\alpha \in D_t$ and a set of potential regressors: $\beta \in \cup_{j>t} D_t$. For joint regressands the permitted relations are direct or indirect associations, i.e. a pair of vertices $\alpha, \beta \in D_t$ is either connected by a line or not. As relation between a regressand $\alpha \in D_i$ and a potential regressor $\beta \in D_j$, with $j > i$ a direct or indirect dependency is possible, i.e. there is either an arrow pointing from β to α or not. Each missing direct relation means that the variable pair is conditionally independent given the remaining joint regressands and potential regressors, i.e. for $\alpha \in D_i$, $\beta \in D_j$ and $j \geq i$ we have:

$$\overset{\alpha}{x} \quad \overset{\beta}{x} \; := \; \alpha \perp\!\!\!\perp \beta \mid (\cup_{t \geq i} D_t) \setminus \{\alpha,\beta\}$$

In the next section we shall further discuss the two conditional independence structures displayed in Figure 1.

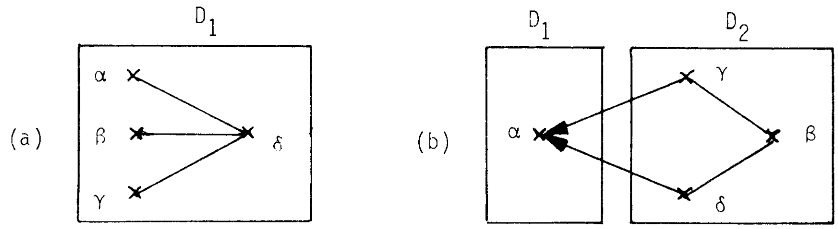


Figure 1: Two conditional independence structures

(a) an association structure with $\alpha \perp\!\!\!\perp \beta \mid (\gamma,\delta)$, $\alpha \perp\!\!\!\perp \gamma \mid (\beta,\delta)$ and $\beta \perp\!\!\!\perp \gamma \mid (\alpha,\delta)$,

(b) a dependence structure with $\alpha \perp\!\!\!\perp \beta \mid (\gamma,\delta)$ and $\delta \perp\!\!\!\perp \gamma \mid \beta$.

A conditional independence structure of the described type together with distributional assumptions define a statistical model. We only consider distributions of the conditional Gaussian type (see Lauritzen and Wermuth, 1984) and speak then of graphical chain models. In graphical representations of these models the vertices have to indicate the type of variable, therefore a vertex is drawn as a circle, if the variable is continuous and as a dot, if the variable is discrete. Examples for such graphs are given in Figure 2.

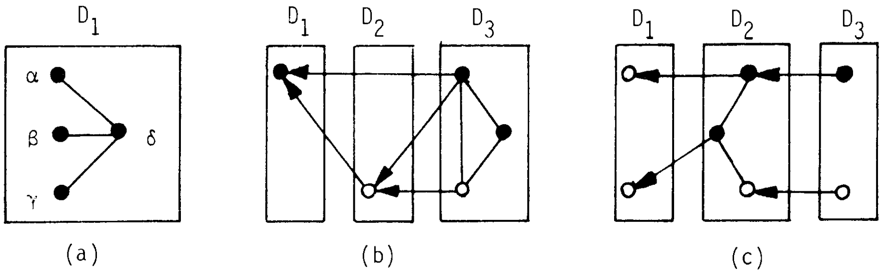


Figure 2: Examples for graphical chain models. An association structure
(a) a graphical log-linear model with $(\alpha \perp\!\!\!\perp \beta \perp\!\!\!\perp \gamma)|\delta$ and mixed dependence structures one with
(b) univariate conditional CG-distributions and one with
(c) joint conditional CG-distributions.

3. SUMMARIES OF DATA FOR GRAPHICAL CHAIN MODELS.

Independencies show up in different summaries, depending on whether the variables in the considered system are only qualitative, only quantitative or how many there are of both types. These differences can be well enough explained by using a system with only four variables. To further simplify the representation we assume here that each qualitative variable has only two possible categories.

3.1 ONLY QUALITATIVE DATA

Each graphical chain model for an association structure with only qualitative variables is a member of the well-studied class of log-linear models (Birch 1963, Bishop, Fienberg and Holland 1974, Haberman 1974, Andersen 1978, Plackett 1981). The natural parameters are here called log-linear interaction terms.

Maximum likelihood estimates for the saturated model can be computed with the help of standard statistical software like BMDP. It is well-known (see Haberman 1978, Wermuth and Lauritzen 1983) that for 2^k-tables these estimates can also be obtained by applying Yates algorithm to the observed log counts. This algorithm is illustrated here in Table 1.

Table 1: Illustration of Yates algorithm for a 2x2 table

cell i j	values a_{ij}	step (1)	step (2)	result of (2) ./. 2^2	effects*	
1 1	10	60	160	40	overall	: μ
2 1	50	100	- 80	- 20	main of i	: α_1
1 2	30	40	- 40	- 10	main of j	: β_1
2 2	70	40	0	0	interaction	: γ_{11}

* with $a_{11} = \mu + \alpha_1 + \beta_1 + \gamma_{11}$; $\Sigma\alpha_i = \Sigma\beta_j = \Sigma_i\gamma_{ij} = \Sigma_j\gamma_{ij} = 0$.

Goodman had first suggested (1971) to look at standardized estimates of the natural parameters of the saturated model for the purpose of judging the goodness-of-fit of a model. Standardisation does not change the results in 2^k-tables. The data speak for the conditional independence of a variable pair, if all the estimated parameters pertainig to this pair are close to zero. As an illustration we use the fictitions data in Table 2. It is difficult to see a simplifying structure in the counts, but the transformation into estimated λ-parameters reveals a perfect fit to an independence structure of the type shown in Figure 1 a, i.e. a perfect fit to the model represented by

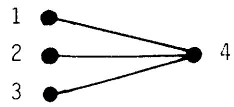
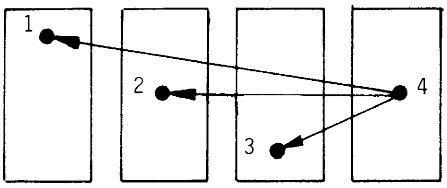


Table 2: An example for a 2^4-contingency table with perfect fit to the hypothesis $(1 \perp\!\!\!\perp 2 \perp\!\!\!\perp 3) \mid 4$.

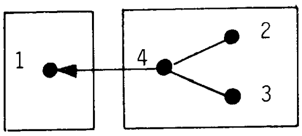
cell i j k ℓ	counts $n_{ijk\ell}$	estimated parameters type	value
1 1 1 1	216	λ	3.94
2 1 1 1	504	$\lambda^{(1)}$	.34
1 2 1 1	24	$\lambda^{(2)}$	.20
2 2 1 1	56	$\lambda^{(1,2)}$	0
1 1 2 1	54	$\lambda^{(3)}$	.00
2 1 2 1	126	$\lambda^{(1,3)}$	0
1 2 2 1	6	$\lambda^{(2,3)}$	0
2 2 2 1	14	$\lambda^{(1,2,3)}$	0
1 1 1 2	36	$\lambda^{(4)}$	.07
2 1 1 2	4	$\lambda^{(1,4)}$	-.76
1 2 1 2	144	$\lambda^{(2,4)}$	.90
2 2 1 2	16	$\lambda^{(1,2,4)}$	0
1 1 2 2	144	$\lambda^{(3,4)}$	.06
2 1 2 2	16	$\lambda^{(1,3,4)}$	0
1 2 2 2	576	$\lambda^{(2,3,4)}$	0
2 2 2 2	64	$\lambda^{(1,2,3,4)}$	0

It cannot be emphasized enough how important it is for good data analyses to look at such transformations and not only at corresponding χ^2-statistics for the goodness-of-fit: all kinds of dependencies, i.e. nonzero interaction parameters can hide behind high-degree of freedom test-statistics that appear to indicate a good fit of the model.

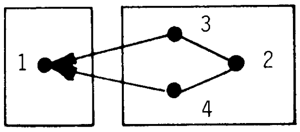
Results on the equivalence of graphical and recursive models (Wermuth and Lauritzen 1983) explain why the above table of counts also gives a perfect fit to a recursive model with three response variables to



Results by Kiiveri, Speed and Carlin (1984) or Lauritzen and Wermuth (1984) tell that it also fits perfectly a model in one response and a conditional independence condition on the regressors:



On the other hand it is, in general, not enough to check the fit of a dependence structure like:



in the described transformation of the 2^4-table: for $3 \perp\!\!\!\perp 4 \mid 2$ one has to look at a similar transformation of a second table, the one obtained from the 2^4-table by marginalising over variable 1.

3.2 ONLY QUANTITATIVE DATA

Each graphical chain model for an association structure with only continuous variables is a covariance selection model (Dempster 1972), applications of which have for instance been described by Wermuth (1978), Hodapp and Wermuth (1983). The natural parameters are here the concentrations, i.e. the elements in the inverse covariance matrix. Data speak for the conditional independence of a variable pair (i,j) given the remaining variables, if an observed concentration s^{ij}, or the partial correlation coefficient given all other variables K is zero:

$$r_{ij.K} = -s^{ij}/(s^{ii}s^{jj})^{1/2}.$$

As an illustration we use the fictitions correlation matrix in Table 3. Again, it is not easy to recognize a simplifying structure in the correlation matrix. The transformation of the data into estimated concentrations of the standardized variables, i.e. the elements in the inverse correlation matrix however, exhibit a perfect fit to the model represented by

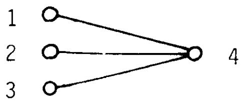
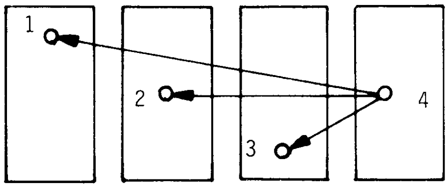
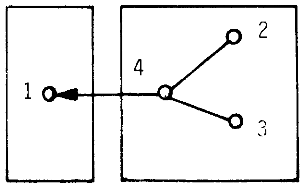


Table 3: An example for a 4x4 correlation matrix with perfect fit to the hypothesis $(1 \perp\!\!\!\perp 2 \perp\!\!\!\perp 3) \mid 4$.

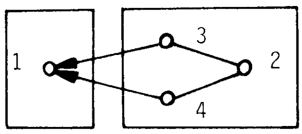
R =	1	.431	.367	.612	R^{-1} =	1.60	0	0	-.98
		1	.422	.704			1.98	0	- 1.40
			1	.600				1.56	-.94
				1					3.14

As for discrete variables, this correlation matrix has at the same time a perfect fit to other types of graphical chain models. This is explained by results on the equivalence of covariance selection models to systems of linear recursive equations with independent errors (Wermuth 1980) and to other recursive systems (Kiiveri, Speed and Carlin 1983). Thus the above association structure is e.g. equivalent to:

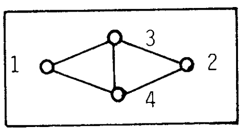


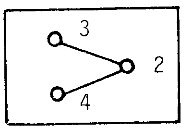


In general, it is not enough to look at the inverse 4x4 matrix R to check the fit of a dependence structure such as



Instead this task can be accomplished by judging the fit of two seperate association structures, of





i.e. $r_{12.34}$ and $r_{34.2}$ have to be near zero.

3.3 TWO QUANTITATIVE AND TWO QUALITATIVE VARIABLES

A graphical chain model for an association structure with both continuous and discrete variables has been called a mixed interaction model (Lauritzen and Wermuth, 1984). If we use a notation for the interaction parameters analogous to the one common for log-linear models, denote the realisations of the continuous variables by x and y and of the discrete variables by i and j, then we can write the joint log-density in terms of the <u>natural interaction parameters</u> as follows:

$$
\begin{aligned}
& \log f(x,y,i,j) \\
= \; & [\eta^{1} + \eta_{i}^{1(3)} + \eta_{j}^{1(4)} + \eta_{i,j}^{1(3,4)}]\, x \\
& + [\eta^{2} + \eta_{i}^{2(3)} + \eta_{j}^{2(4)} + \eta_{i,j}^{2(3,4)}]\, y \\
& - \tfrac{1}{2} [\psi^{1} + \psi_{i}^{1(3)} + \psi_{j}^{1(4)} + \psi_{i,j}^{1(3,4)}]\, x^{2} \\
& - \tfrac{1}{2} [\psi^{2} + \psi_{i}^{2(3)} + \psi_{j}^{2(4)} \quad \psi_{i,j}^{2(3,4)}]\, y^{2} \\
& - [\psi^{12} + \psi_{i}^{1,2(3)} + \psi_{j}^{1,2(4)} + \psi_{i,j}^{1,2(3,4)}]\, xy \\
& + [\lambda + \lambda_{i}^{(3)} + \lambda_{j}^{(4)} + \lambda_{i,j}^{(3,4)}] \, .
\end{aligned}
\qquad (1)
$$

Here, we have e.g. as two-factor interactions terms a pure discrete term: $\lambda_{ij}^{(3,4)}$, a pure continuous term: ψ^{12}, mixed linear terms: $\eta^{1(3)}$, $\eta^{2(3)}$. $\eta^{1(4)}$, $\eta^{2(4)}$ and mixed quadratic terms, like $\psi^{1(3)}$, $\psi^{2(3)}$. Similarly, there are main effect terms and higher-order mixed terms. To get the maximum-likelihood estimates of the interactions in the saturated model, one needs to compute for each combination i,j the

$$
\begin{aligned}
n_{ij} &:= \text{count}, \\
\bar{x}_{ij} &:= \text{mean of } x, \\
\bar{y}_{ij} &:= \text{mean of } y, \\
S_{ij} &:= \text{covariance matrix for } x \text{ and } y.
\end{aligned}
\qquad (2)
$$

We write the elements in the inverse covariance matrices as

$$
S_{ij}^{-1} = \begin{bmatrix} s_{ij}^{xx} & s_{ij}^{xy} \\ & s_{ij}^{yy} \end{bmatrix} \qquad (3)
$$

Then, the following six 2x2 tables are the basis for obtaining the desired estimates in each of the consecutive rows listed in (1) :

$$
\begin{array}{l}
\bar{x}_{ij}\, s^{xx}_{ij} + \bar{y}_{ij}\, s^{xy}_{ij} \,, \\[2ex]
\bar{y}_{ij}\, s^{yy}_{ij} + \bar{x}_{ij}\, s^{xy}_{ij} \,, \\[2ex]
s^{xx}_{ij} \,, \\[2ex]
s^{yy}_{ij} \,, \\[2ex]
s^{xy}_{ij} \,, \\[2ex]
\log n_{ij} - \frac{n}{2} [\log \det S_{ij} + (\bar{x}_{ij}, \bar{y}_{ij})\, S_{ij}^{-1} (\bar{x}_{ij}, \bar{y}_{ij})^T].
\end{array}
\qquad (4)
$$

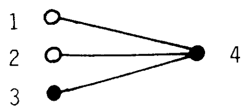
Again, one may use Yates algorithm to determine from each 2x2 table an overall effect, the main effects and an interaction effect and again, the data will support the conditional independence of a variable pair given all remaining variables, if the estimates of <u>all</u> interaction terms pertaining to this variable pair are close to zero.

To have a good fit to $1 \perp\!\!\!\perp 2 \mid (3,4)$, i.e. $x \perp\!\!\!\perp y \mid (i,j)$ the concentrations s^{xy}_{ij} have to be near zero for all i,j. For $1 \perp\!\!\!\perp 3 \mid (2,4)$, i.e. $x \perp\!\!\!\perp i \mid (y,j)$ no interaction effect and no main effect for i is permitted in each of the three tables with $\bar{x}_{ij}\, s^{xx}_{ij} + \bar{y}_{ij}\, s^{xy}_{ij}$, s^{xx}_{ij} and s^{xy}_{ij} and finally, for $3 \perp\!\!\!\perp 4 \mid (1,2)$, i.e. $i \perp\!\!\!\perp j \mid (x,y)$ one has to look at all six 2x2 tables of (4) and only main effects are allowed in each of them.

In conditional independence structures there are usually several pairwise independencies, e.g. for $(1 \perp\!\!\!\perp 2 \perp\!\!\!\perp 3) \mid 4$ the terms for pairs interaction (1,2), (1,3) and (2,3) vanish, i.e. (1) reduces to

$$
\begin{aligned}
&\log f(x,y,i,j) \\
&= [\eta^1 + \eta_j^{1(4)}]\, x + [\eta^2 + \eta_j^{2(4)}]\, y \\
&- \frac{1}{2} [\psi^1 + \psi_j^{1(4)}]\, x^2 - \frac{1}{2} [\psi^2 + \psi_j^{2(4)}]\, y^2 \\
&+ [\lambda + \lambda_i^{(3)} + \lambda_j^{(4)} + \lambda_{i,j}^{(3,4)}]
\end{aligned}
$$

Consequently, data will fit this model :

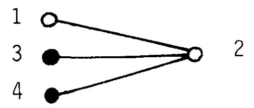


well, if the first four tables of (4) show neither a main effect for i nor an interaction effect and if the concentrations s^{xy}_{ij} are close to zero for all i,j. This turns out to be equivalent to asking for uncorrelated x and y in all classifications and for equal means, like $\bar{x}_{ij} = \bar{x}_j$ and equal variances, like $s_{xx,ij} = s_{xx,j}$.

In the model with e.g. $(1 \perp\!\!\!\perp 3 \perp\!\!\!\perp 4) \mid 2$ all interaction terms with (1,3), (1,4) and (3,4) are zero, i.e. (1) reduces to

$$\begin{aligned}
&\log f(x, y, i, j) \\
&= [\eta^1] x + [\eta^2 + \eta_i^{2(3)} + \eta_j^{2(4)}] y \\
&- \frac{1}{2} [\psi^1] x^2 - \frac{1}{2} [\psi^2 + \psi_i^{2(3)} + \psi_j^{2(4)}] y^2 \\
&- [\psi^{12}] xy + [\lambda + \lambda_i^{(3)} + \lambda_j^{(4)}].
\end{aligned}$$

Data will therefore be compatible with this model



if there is only an overall effect in the tables with $\bar{x}_{ij}\, s_{ij}^{xx} + \bar{y}_{ij}\, s_{ij}^{xy}$, s_{ij}^{xx} and s_{ij}^{xy} and no interaction effect in the remaining tables listes in (4). Thus, the same type of independence structure can lead to rather simple and fairly complicated implications for the data, depending mainly on which of the variables under investigation are qualitative and which are quantitative.

4. SOME WARNINGS

To deduce conditional independence just from looking at data can of course be misleading. One should for example consider the possible effects of

- mere sampling fluctuations
- violated distributional assumptions
- reclassifications of categories.

If there is for instance a small but definitely nonzero conditional association in the population, then it still will appear as a zero association in many samples. Thus, if a conditional independence is judged to be implausible from subject matter considerations, then an apparent zero interaction in some data need not be taken as evidence for the conditional independence, but is more likely to have occurred by mere chance, i.e. because of sampling fluctuations.

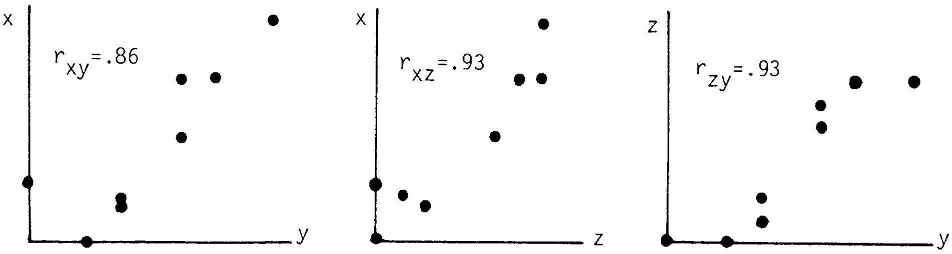
If we have the following correlation matrix

$$R = \begin{bmatrix} 1 & .62 & .82 \\ & 1 & .76 \\ & & 1 \end{bmatrix}$$

then the partial correlation $r_{12.3}$ and the concentration r^{12} are zero, since $r_{12} = r_{13}\, r_{23}$. This matrix may have been computed in a sample from a joint normal distribution, then the data point at a conditional independence: $1 \perp\!\!\!\perp 2 \mid 3$. If it results however from the following "observations"

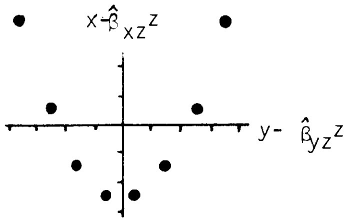
x_i :	10	28	28	18	6	0	8	38
y_i :	0	26	32	26	16	10	16	42
z_i :	0	24	28	20	8	0	4	28

with even apparent pairwise linear relations:



then a simple plot corresponding to the partial correlation coefficient, i.e. the plot of residuals indicates instead a strong nonlinear conditional dependence for the variables 1 and 2 given 3:

$r_{xy.z} = 0$



Similarly, if we have the following "observed" table of counts

	j	
i	1	2
1	50	50
2	50	50

it can point at the independence of the two variables, but it may also have resulted from a particular sampling plan, i.e. from taking equal numbers of observations in all groups. Then, the numbers are of course not the summary of a random sample for the two qualitative variables.

Another possibility is that the table does not correspond to the original categories of one of the variables. If, for instance j = 1 is made up from categories j* = 1 and 2 and j = 2 from the original levels j* = 3 and 4, then this reclassification can cover up an actual dependence, as is illustrated with the next table:

	j*			
i	1	2	3	4
1	40	10	10	40
2	10	40	40	10

Undoubtedly, one will be able to produce similar examples for possible pitfalls in interpreting associations and between quantitative and qualitative variables, once they have been better studied. This can be done with graphical chain models: they combine original aims in analysing structures with probabilistic concepts and distributional assumptions and tie them to graphtheoretic results.

BIBLIOGRAPHY

Andersen, E.B. (1980). Discrete statistical models with social science applications. Amsterdam: North-Holland.

Anderson, T.W. (1958). An introduction to multivariate statistical analysis. New York: Wiley.

Birch, M.W. (1963). Maximum-likelihood in three-way contingency tables. J.Roy.Statist.Soc. B, 25, 220-223.

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975). Discrete multivariate analysis theory and practi e. Cambridge: M.I.T.

Blalock, H.M.Jr. (ed.) (1971). Causal models in the social sciences. Chicago, Aldine-Altherton.

Darroch, J.N., Lauritzen, S.L. and Speed, T.P. (1980). Markov fields and log-linear interaction models for contingency tables. Ann. Statist., 8, 522-539.

Dempster, A.P. (1972). Covariance selection. Biometrics, 28, 157-175

Goldberger, A.S. and Duncan, O.D. (eds.) (1973). Structural equation models in the social sciences. New York: Seminar Press.

Goodman, L.A. (1970). The multivariate analysis of qualitative data: interactions among multiple classifications. J.Amer.Statist.Assoc. 65, 226-256.

- (1971). The analysis of contingency tables when some variables are posterior to others: a modified path analysis approach. Biometrika, 60, 179-192.

- (1978). Analyzing qualitative / categorical data. London: Addison Wesley.

Haberman, S.J. (1974). The analysis of frequency data. Chicago: Univ. of Chicago Press.

Hodapp, V. (1982). Causal inference from nonexperimental research on anxiety and educational achievement. In: H.W.Krohme and L.Laux (eds.) Achievement, stress and anxiety. Washington, D.C.: Hemisphere.

- and Wermuth, N. (1983). Decomposable models: a new look at interdependence and dependence structures in psychological research. Mult.Behav.Res. 13. 361-390.

Jöreskog, K.G. and Sörbom, D. (1978). LISREL IV: Analysis of linear structural relationships by the method of maximum-likelihood. Chicago: International Educational Services.

Kiiveri, H., Speed, T.P. and Carlin, J.B. (1984). Recursive causal models. J.Austral.Math. Soc., A, 36, 30-52.

Lauritzen, S.L. and Wermuth, N. (1984). Mixed interaction models. Research Report. Inst. for Electronic Systemer, Aalborg Universitetscenter (ISSN-0106-0791).

Lazarus, R.S. (1966). Psychological stress and the coping process. New York: McGraw Hill.

McCullagh, P. and Nelder, J.A. (1983). Generalised linear models. London: Chapman and Hall.

Plackett, R.L. (1981). Analysis of categorical data. New York: Mc Millan.

Simon, H.A. (1957). Models of man. New York: Wiley.

Spearman, C. (1926). The abilities of man. London: Mac Millan.

Wermuth, N. (1978). Zusammenhangsanalysen medizinischer Daten. Berlin: Springer.

- (1980). Linear recursive equations, covariance selection, and path analysis. J.Amer.Statist. Assoc., 75, 963-972.

- and Lauritzen, S.L. (1983). Graphical and recursive models for contingency tables. Biometrika, 70, 537-552.

Wright, S. (1923). The theory of path coefficients: a reply to Niles' criticism. Genetics, 8, 239-255.

- (1934). The method of path coefficients. Ann.Math.Statist. 5, 161-215.

SUMMARY

We give some historical background on models for analysing structures, describe the graphical representation of conditional independence structures and illustrate how differently these structures are reflected in data from graphical chain models depending on how many and which of the variables in the investigated system are qualitative, which are quantitative.