TAGUNGSBERICHT 26/1995

Algebraic Methods in Multivariate Statistical Analysis

2 - 8 July 1995

The conference was organized by Michael D. Perlman (Seattle, USA) and Friedrich Pukelsheim (Augsburg, Germany), and attended by 45 participants. Various topics in multivariate statistics were presented in 30 talks each followed by intensive discussions. The main interest focussed on two fields: theory and practice of graphical models, and the use of algebraic methods in multivariate statistics.Among the topics that recieved a special emphasis were: existence of the maximum likelihood estimator, factorization of the sample space and the likelihood function, causal interpretation of graphical models, counterfactuals, characterizations of important multivariate distributions, Gaussian and dual Gaussian models, inequalities and stochastic ordering, lattice conditional independence models, group symmetry models, models with a Jordan algebra structure, Bayesian analysis of graphical models, Markov equivalence classes and Markov properties in graphical models, distributive block structures in experimental layouts, invariance and incosistency.

Gefördert durch
DFG Deutsche Forschungsgemeinschaft

Abstracts

Muriel CASALIS

The Lukacs-Olkin-Rubin characterization of the Wishart distribution on symmetric cones

We characterize the Wishart distributions on a symmetric cone Ω as follows: let G denote the connected component of the structural group $G(\Omega)$ of the cone containing the identity, and let $x \in \Omega \to g(x) \in \Omega$ be a division algorithm defined by $g(x)(x) = e$ (the unity of Ω) (for example, $g(x)u = x^{-1/2}ux^{-1/2}$). Let U and V be independent non Dirac random variables valued in $\bar{\Omega}$ such that U and V are not concentrated on the same one dimensional subspace and such that $U+V$ is positive definite a.s. Define then $Z = g(U+V)U$. If the distribution of Z is invariant under orthogonal transformations of G, if Z and $U+V$ are independent, then U and V are Wishart distributed. If $\Omega = (0,\infty)$, this result is due to Lukacs in 1956. If Ω is the cone of positive definite symmetric matrices, it is due to Olkin and Rubin in 1962. We extend it to symmetric cones and shorten the proof by using three ideas:

1. we avoid to work with coordinates in differential geometry

2. we use the variance function of a natural exponential family which characterizes the family

3. we use the theory of Jordan algebras since a symmetric cone is defined naturally in a Jordan algebra framework.

(in collaboration with Gerard Letac)

David COX

General remarks on graphical models

Applications of graphs in probability and statistics range from probabilistic graphs (including explanations of Horlons laws in geomorphology), graphs in decision analysis (influence diagrams), in the combinatorics of experimental design, in probabilistic expert systems and finally in statistical methods for the analysis and interpretation of data, the objective of the present work. A key issue in statistics is the incorporation of substantive knowledge and assumptions. In many fields this is achieved by

1. partial ordering of variables as explanatory, intermediate and response,

2. hypotheses of conditional independence and dependence (the substantive research hypotheses of Wermuth and Lauritzen).

The imperative of taking account of (1) and (2) points to chain block graphs as being of central importance. Six types of graphs werde described, including graphs with two different kinds of edges. These raise issues of the interrelationships between the different graphs, for example between a chain block graph and an undirected concentration or covariance graph. Finally a number of open issues were listed.
(joint work with Nanny Wermuth)

Phil DAWID

The analysis of distributive block structures for experimental layouts

Distributive block structures form a rich class of "balanced orthogonal designs". A DBS is a distributive lattice of equivalence relations on a set of units satisfying some extra properties. In some contexts, attitudes to measurements on the units are invariant under permutations preserving the DBS. We then obtain a simple canonical decomposition of variable space. Comparing such decompositions for nested DB structures determines how hypotheses should be tested. We can also construct a "linear model" to describe the structure of the variables under a DBS. The ideas extend to cases where some factors have fixed levels giving rise to more complex "mixed models" and associated tests.

Morris L. EATON

Invariance and strong inconsistency

We study invariant prediction problems when the group action is transitive on the parameter space. Under certain structural assumptions, we identify a preferred inference and show that other invariant inferences are strongly inconsistent in the sense of M. Stone. The preferred inference is shown to be "not strongly inconsistent" when the group is amenable. As an application, we show that in standard MANOVA prediction models, all fully invariant inferences are strongly inconsistent.

Kai Tai FANG

On multivariate vertical densitiy representation

Motivated by the works of Troutt (1991, 1993) and Kotz and Troutt (1994), we provide a multivariate definition of the vertival density representation (vdr) and calculate its value for some basic multivariate distributions presented in Fang, Kotz and Ng (1990) and Johnson (1987). Utilizing the multivariate vdr we can generalize the well-known Box-Muller method to generate the multivariate distributions whose vdrs are derived.
(in co-authors with Samuel Kotz and Jia-Juan Liang)

Dan GEIGER

A characterization of the Dirichlet, Normal and Wishart distribution, with applications to learning Bayesian networks

We provide a new characterization of the Dirichlet distribution. This characterization implies that under assumptions made by several previous authors for learning Bayesian networks, a Dirichlet prior on the parameters is inevitable.
This article will appear in the Proceedings of the 11th conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann, Canada, August 1995.
(Joint work with David Heckerman)

Søren Tolver JENSEN

Classification of full exponential transformation models

The full natural exponential transformation models with a transitive action on the parameter set D are classified. The action on the sample space $V = \mathbb{R}^k$ must be linear $gy = A(g)y + a(g)$ and the exponential family $\frac{dP_\theta}{d\nu}(y) = \frac{1}{\varphi(\theta)} \exp\{\theta' y\}$, $\theta \in D$, is an exponential transformation model if and only if the structural measure ν is quasiinvariant with a multiplicator of the form

$$\chi(g)(y) = \exp\{b(g^{-1})'y + d(g)\},$$

and then the action on the parameter space must be

$$g\theta = A(g^{-1})'\theta + b(g).$$

The problems are to find the possible group actions and the possible multiplicators. If the action on the parameter space is transitive, there is a complete solution to the problem.

Göran Kauermann

On a dualization of graphical Gaussian models

Graphical Gaussian models as defined by Speed and Kiiveri (Ann.Stat., 1986) present the conditional independence structure of normaly distributed variables by a graph. A similar approach was motivated by Cox and Wermuth (Stat. Science, 1993) by introducing graphs showing the marginal independence structure. We define different Markov properties playing the role of a link between the graphical approach and the induced independence relations, either conditional or marginal. We derive similar results for marginal independencies in the normal distributions as have been proven by Frydenberg (Ann. Stat., 1990) for conditional independencies.

Finally, the focus is on parameter estimation. We show how dual maximum likelihood estimation allows to adapt familiar results on the estimation theory for conditional independence restrictions to models for marginal independencies.

Jan KOSTER

Markov properties of non-recursive causal models

A new class of conditional independence (CI) probability models is introduced, determined by a so-called reciprocal graph. This class strictly includes the well-known class of graphical chain models studied by Frydenberg, Lauritzen and Wermuth etc., and the class of probability models determined by a directed graph (DCG), studied recently by Spirtes and Glymour. The class of "lattice CI models", introduced by Andersson and Perlman, is also strictly included. It is shown that the Markov property determined by a reciprocal graph is equivalent to the existence of a Gibbs factorization of the density (assumed positive). The class of lattice CI models plays a key role in understanding the Markov structure of reciprocal graphical models.

STEFFEN L. LAURITZEN
On the statistical theory of graphical models
The lecture surveys the state-of-art for the sampling statistical theory of graphical models in the Gaussian and mixed discrete-Gaussian case. After introducing basic concepts for undirected graphical models, the decomposable and non-decomposable cases were contrasted in the Gaussian case: when graphs are decomposable, i.e., all cycles of length ≥ 4 have chords, explicit computational and distributional results are available. In the non-decomposable case, methods are essentially limited to iterations and approximate distributions. Even the problem of existence of the MLE (i.e. boundedness of the likelihood function) is not cleared completely. In the mixed discrete-Gaussian case, explicit results are also available for decomposable graphs. But some non-graphical, hierarchical interaction models do also have explicit solutions and simple interpretations.

David MADIGAN
Bayesian analysis of Markov equivalence classes of acyclic digraphs
Acyclic digraphs (ADG) are widely used to describe dependencies among variables in multivariate distributions. In particular, the likelihood functions of ADG models admit convenient recursive factorizations that often allow explicit maximum likelihood estimates and straightforward Bayesian analyses. There may, however, be many ADGs that determine the same dependence (= Markov) model. Thus, the family of all ADGs with a given set of vertices decomposes into Markov equivalence classes, each class being associated with a unique statistical model. Statistical procedures such as Bayesian model averaging that fail to take into account these equivalence classes may incur substantial computational or other inefficiencies. In recent work, we have shown that each Markov equivalence class is uniquely determined by a single chain graph, the essential graph that is itself Markov equivalent to all the ADGs in the equivalence class. Here we propose two stochastic Bayesian model averaging and selection algorithms and describe some simple applications.
(joint work with Michael Perlman, Steen Andersson and Chris Volinsky)

Jesper MADSEN
Models combining group symmetry and conditional independence in a multivariate normal distribution
We define and study a class of models (GS-LCI-models) for the multivariate normal distribution by means of intersections of the class of group symmetry (GS) models and the class of lattice conditional independence (LCI) models. If G denotes a finite group and $\rho : G \to O(I)$ an orthogonal group representation of G on $\mathbb{R}^I$ then the GS-model given by G is defined to be the family of all normal distributions $N(0, \Sigma)$ on $\mathbb{R}^I$ that remains invariant under all transformations $\rho(g) : \mathbb{R}^I \to \mathbb{R}^I$, $g \in G$. When $\mathcal{K}$ denotes a lattice of subsets of I then the LCI model given by $\mathcal{K}$ is defined to be the class of all normal distributions $N(0, \Sigma)$ on $\mathbb{R}^I$ for which $X_L \perp\!\!\!\perp X_M \mid X_{L\cap M}$ for all L, M in $\mathcal{K}$ whenever $X \sim N(0, \Sigma)$.

Under the assumption that G and $\mathcal{K}$ "work together", in the sense, that $(\rho(g)x)_K = \rho(g)_K x_K$ for all $g \in G$, $x \in I\!R^I$ and $K \in \mathcal{K}$, we then define the GS-LCI model given by G and $\mathcal{K}$. For these models explicit likelihood inference (MLEs, LR-tests etc.) can be done.

James MALLEY

Statistical applications of Jordan algebras

After stating the abstract definition for a Jordan algebra we consider alternatives useful in practice, and list matrix space examples. Our list of applications includes:

- optimal variance component estimation, both unbiased and nonnegative
- frameworks for linear hypothesis testing
- analytic, closed-form solutions for the M-step in the EM algorithm, as applied to patterned covariance problems, in the presence of missing data
- Bayes optimal decision rule for quantum mechanical events.

John MARDEN

Multivariate ranks and their application to invariant normal models

There is very little on using rank methods for developing robust tests on covariance matrices. Given multivariate observations, one can replace the data by the ranked data, where each variable is ranked separately. However, doing so destroys the possibility of testing many popular structural hypotheses on covariance matrices. Instead, by using Smalls multivariate ranks, which are defined using all variables at once, models defined on the original variables using invariance under a compact group hold also for the rank-transformed data.

Hélène MASSAM

Lattice conditional independence models in a Jordan algebraic framework

Lattice conditional independence models have been defined for $y \sim N(0,\Sigma)$, $y \in I\!R^p$ (Andersson and Perlman, 1993). For a given lattice $\mathcal{K}$, the model is determined by the subspace $\mathcal{P}(\mathcal{K})$ of the cone of positive definite symmetric matrices such that $\Sigma \in \mathcal{P}(\mathcal{K})$. The cone of positive definite real symmetric matrices is the symmetric cone of the Jordan algebra of symmetric matrices on $I\!R$. We extend the definition of $\mathcal{P}(\mathcal{K})$ to all symmetric cones Ω in the following way: Let x be an element of the irreducible symmetric cone Ω of the Jordan algebra V. Then

$$\begin{aligned} x \in \mathcal{P}(\mathcal{K}) \quad \text{iff} \quad & \forall L, M \in \mathcal{K}\ \forall y \in V: \\ & \operatorname{tr}\, x_{L\cup M}^{-1} y_{L\cup M} = \operatorname{tr}\, x_L^{-1} y_L + \operatorname{tr}\, x_M^{-1} y_M - \operatorname{tr}\, x_{L\cap M}^{-1} y_{L\cap M}. \end{aligned}$$

We denote by $J(\mathcal{K})$ the set of all irreducible elements of the lattice $\mathcal{K}$ and given $L \in \mathcal{K}$, $J(\mathcal{K}_L) = \{K \in J(\mathcal{K}), \mathcal{K} \subset \mathcal{L}\}$. The cone $\mathcal{P}(\mathcal{K})$ can be characterized as follows

$$x \in \mathcal{P}(\mathcal{K}) \Leftrightarrow \forall L \in \mathcal{K} : x_L^{-1} = \sum_{K \in J(\mathcal{K}_L)} (x_K^{-1} - x_{<K>}^{-1}).$$

It also holds that

$$\det x_L = \prod_{K \in J(\mathcal{K}_L)} \frac{\det x_k}{\det x_{<K>}}.$$

From these two formulas, we can derive the estimates of the maximum likelihood estimate of $x \in \mathcal{P}(\mathcal{K})$ with formulas of the type

$$\hat{x}_{L \cup M}^{-1} = \hat{x}_L^{-1} + \hat{x}_M^{-1} - \hat{x}_{L \cap M}^{-1}$$

whenever $L \cap M$ separates L from M, as given by Lauritzen (1995) for decomposable models. We can also derive the moments of the likelihood ratio statistic Q for comparing two models with lattices $\mathcal{K}$ and $\mathcal{M}$ such that $\mathcal{M} \subset \mathcal{K}$. There, similar to what happens for decomposable models, $Q = \det V$ where V is the product of independent beta distributions.

František MATÚŠ

On conditional independence and polymatroids

Having a finite set N, a set-function $g : 2^N \to [0, \infty)$ is called probabilistically (p-) representable if there exist random variables $\xi = (\xi_i)_{i \in N}$ and a number $\epsilon > 0$ such that $\epsilon g(I)$ is equal to the Shannon entropy of $\xi_I = (\xi_i)_{i \in I}$ for all $I \subset N$. Such a family ξ is said to be p-representation of g. An immediate observation is that every p-representable g must be (a rank function) of a polymatroid. A set $\mathcal{L}$ of triples $(I, J|K)$ where $I, J, K \subset N$ is p–representable if for some ξ

$$(I, J|K) \in \mathcal{L} \Leftrightarrow \xi_I \text{ is conditionally independent of } \xi_J \text{ given } \xi_K.$$

The problem of which polymatroid is p–representable is shown to be more general than that of p–representability of sets $\mathcal{L}$.

Every linearly representable matroid is p–representable. All p–representations ξ of a connected matroid of rank at least two must satisfy: there exists an integer $u \geq 2$ such that ξ_I is uniformly distributed on a set of cardinality $u^{r(I)}$ for all $I \subset N$ (r is the rank function). All p–representations of this matroid can be constructed from finite groups and they are (up to a proper notion of isomorphism among them) in one-to-one correspondence with the classes of isomorphic groups.

Chris MEEK

New methods for selecting graphical models

The problem of model selection in the case of graphical models (e.g. acyclic digraphs or chain graphs) is complicated by the fact that there is a combinatorial explosion in the number of graphs as a function of the number of vertices. In addition, most search operators on the space of graphical often require many operations to reach statistically similar models. This last problem arises because graphically dissimilar models can represent the same set of distributions (i.e. be Markov equivalent). A natural approach to addressing both of these problems is to consider a transformation of the search space from model space to the model space modulo an equivalence relation. In the case of acyclic digraphs this is accomplished by searching in pattern space; a pattern is a canonical mixed graphical representation of a Markov equivalence class. This transformation of the search space is justified by the fact that Markov equivalent models receive the same penalized likelihood score. Algorithms for finding a completed pattern (a unique representative of a Markov equivalence class which has all and only those edges common to members of that Markov equivalence class and which is a chain graph) and for finding a Markov equivalent acyclic digraph were presented in addition to some preliminary results of a particular algorithm for searching pattern space.

Kenneth NORDSTRÖM

Stochastic Schur-convexity properties of functions of noncentral χ^2-variables

Stochastic ordering properties of functions of independent noncentrally χ^2-distributed random variables will be discussed. In particular, a result will be presented which establishes the (multivariate) stochastic Schur-convexity of the vector of partial sums of independent noncentrally χ^2-distributed random variables as a function of the vector of noncentrality parameters. Connnections to the problem of determining the extrema of the probability content of a rotated ellipsoid will also be discussed. This is joint work with Thomas Mathew, University of Maryland Baltimore County.

Ingram OLKIN

A personal overview of the conference and *Correlation analysis of extreme observations from a multivariate normal distribution*

The theme of the conference might be called "Dependency Models", and the talks related to some parts of this topic. I outline the various components of dependency:

1. Independence, conditional independence, graphical structures
2. Time Series
3. Stochastic Processes
4. Markov Chains

5. Regression Type Models — latent variables, factor analyses, structural equations, hierarchical models, causal models

6. Correlation analysis — product moment correlation, partial correlation, multiple correlations, tetrachoric correlation, part correlations

7. Non-parametric methods

   (a) Families of bivariate distributions — Farlie, Gumbel-Morgenstern, Gumbel, Plackett, Marshall-Olkin, etc.

   (b) Inequalities — Stepian, Sidak, Total positivity, majorization, stochastic orderings, F-K-G, etc.

   (c) Characteristics — positive orthant dependent, associated, increasing failure rate, less concordant, etc.

8. Multivariate Normal Distribution

9. Multivariate Non-normal Distributions

   (a) Fitting — density estimation splines, etc.

   (b) Generation of distributions

      i. Bounds — Fréchet bounds, etc.

      ii. Methods of generating bivariate distributions
      That generate bivariate binomial, Poisson, geometric, gamma, exponential, etc.

         - Mixtures
         - Compounding
         - Statistical properties
         - Characterizations
         - physical models

In this survey we discuss how these components interact in trying to develop dependency systems.

Judea PEARL

Graphs, structural models, and counterfactuals

The paper presents a formal interpretation of Structural Equation models as used in econometrics and social science and the Neyman-Rubin-Holland-Robins model of counterfactuals (or potential response). We show that every causal-diagram (Pearl, 1995, forthcoming Biometrika) can be interpreted in the Rubins model and all causal conclusions derived by two axioms: composition and consistency.

Michael PERLMAN

Lattice conditional independence models and acyclic digraphs

Let $\mathcal{K}$ be a ring ($\equiv$ finite distributive lattice) of subsets of a finite index set I. A random vector $X \in \mathbb{R}^I$ satisfies the lattice conditional independence (LCI) property determined by $\mathcal{K}$ if $X_L \perp\!\!\!\perp X_M \mid X_{L \cap M}$ for all $L, M \in \mathcal{K}$, where $X_L := (X_i \mid i \in L)$. We show that the LCI models coincide with the graphical Markov models determined by transitive acyclic digraphs (TADGs). Furthermore, a graphical characterization is given for those ADGs that are Markov equivalent to some TADG, i.e., to some LCI model. Under the additional assumption of multivariate normality, it is shown that LCI models are suitable for the analysis of non-nested missing data patterns and seemingly unrelated regressions.

Donald St. RICHARDS

Algebraic methods towards higher-order probability inequalities

In treating the anaylsis of variance with two criteria of classification, we are led to the probability inequality

$$P(V_1 \geq a_1, V_2 \geq a_2) \geq P(V_1 \geq a_1)P(V_2 \geq a_2) \tag{1}$$

where (V_1, V_2) are "positively dependent by mixture" (PDM); that is, $(V_1, V_2) = (W_1, W_2)/X$ where W_1, W_2 and X are mutually independent, positive random variables. In the ANOVA context, W_1, W_2 and X are chi-square distributed quadratic forms; and V_1 and V_2 are (constant multiples of) the likelihood ratio statistics for testing row and column effects, respectively.
Using the notion of "similarly ordered" functions, together with some simple algebraic identities, we derive a variety of generalizations of the inequality (1). One general result in this direction is the following: Given a random variable X taking values in a space $\mathcal{X}$, and similarly ordered functions $\phi_1, \ldots, \phi_m : \mathcal{X} \to \mathbb{R}_+$, then $\mathrm{E}[\prod \phi_i] \geq \prod \mathrm{E}[\phi_i]$, whenever all expectations exist. In another direction, it is well-known that the inequality (1) arises within the context of "total positivity of order 2" (TP_2). This motivates the problem of deriving probability inequalities for bivariate densities K which are totally positive of order m (TP_m). We derive a number of inequalities generalizing (1), the simplest of which is the following: If the random vector (X, Y) has a TP_m density K, then the $m \times m$ matrix with (i, j)–th entry $\mathrm{E}X^{i-1}Y^{j-1}$ has nonnegative determinant. Finally, we derive probability inequalities for densities which satisfy recently-defined notions of total positivity, including those notions involving finite reflection groups.

Dietrich von ROSEN
Multivariate linear models with certain covariance structures
Let
$$y=(y_0',y_1',y_2')'\sim N_{p^3}(0,\Phi^{-1})$$
where
$$\Phi=\begin{pmatrix}\Phi_0 & -\Phi_1 & -\Phi_2\\ \Phi_1 & \Phi_0 & 0\\ \Phi_2 & 0 & \Phi_0\end{pmatrix}.$$
Here $y_i : p\times 1, i=0,1,2, \Phi_0$ is positive definite, Φ_1 and Φ_2 are both skew symmetric and Φ is positive definite. The density of y equals
$$c\mid\Phi\mid^{1/2}e^{-1/2\mathrm{tr}(\Phi yy')},\qquad c=(2\pi)^{-1/2p^3}. \tag{1}$$
Let $z=y_0+iy_1+jy_2$ and $\Phi=\Phi_0+i\Phi_1+j\Phi_2$, where $i^2=j^2=-1$ and $ij=ji=0$. it follows that (1) is equivalent to
$$c\mid\Phi\mid\mid\Phi_0\mid^{1/2}e^{-1/2\mathrm{tr}(\Phi zz')},\qquad c=(2\pi)^{-1/2p^3}$$
where $\mid\Phi\mid$ is defined in a certain way. Thus, a density of z has been obtained.

Glenn SHAFER
Alternative causal interpretations of multivariate models
This lecture used the idea of probability trees to provide causal interpretations for some common statistical models. Emphasis was placed on three simple models that have been represented by directed acyclic graphs:

1. $X \qquad Y$ (X and Y are independent)
2. $X\longrightarrow Y$ (Y depends on X)
3. $X \begin{smallmatrix}\nearrow Y\\ \searrow Z\end{smallmatrix}$ (Z does not depend on X given Y)

In each case, several probability-tree representations were given, leading sometimes to the same and sometimes to different sample-space representation. The implication of these results is that claims of causal meaning for these statistical models should, in general, be made more precise.
The probability-tree interpretations lead away from the idea that variables on events are causes. Steps in natures tree are causes, and causal relations among variables are always statements about these variables common causes. For example, variables are independent when they have no common causes.

Peter SPIRTES
Marginalization, conditionalization and causal inference
The variables in each DAG are partitioned into 3 sets, O, the observed, L, the latent. and S the conditioned one. It is assumed that the DAG generates a probability distribution, but all we observe is $P(O\mid S)$. A sense of equivalence for DAGs with variables partitioned not this way is defined,

and a polynomial algorithm for deciding equivalence is decribed. A generalization of patterns, named POIPGs is described. Each POIPG represents an equivalence class of DAGs with the variable partitioned into O, S and L. The features of algorithms for constructing POIPGS are described, and it is shown how to draw conclusions about the effect of interventions from POIPGs.

Muni S. SRIVASTAVA

Admissibility of the inverse regression estimator in the multi-univariate calibration model

For simplicity, we consider the unidimensional linear calibration model in which the calibration experiment can be represented as

$$y_i = \alpha + \beta x_i + \epsilon_i, \qquad i = 1, \ldots, n$$

and the calibration experiment as

$$y_{0j} = \alpha + \beta x_0 + \epsilon_{0j}, \qquad j = 1, \ldots, m$$

where y_i, y_{0j} are all independent random p-vectors and α and β are p-vectors of unknown parameters. To estimate x_0, two estimators have been proposed in the literature, classical and inverse regression. We show that the inverse regression estimator is admissible when $\epsilon_i \sim N_p(0, \Sigma)$ and $\epsilon_{0j} \sim N_p(0, \Sigma)$ and Σ unknown.

Milan STUDENÝ

Chain graphs as models of probabilistic conditional independence structures

Chain graphs are known as graphical models of conditional independence (CI) structure which generalize both undirected graphs and directed acyclic graphs. They define through a "moralization" procedure certain independency models. Nevertheless, the principle question whether it is really a model of probabilistic CI-structure (i.e. whether for each chain graph there exists a probability measure complying exactly with independencies written in the graph through the moralization criterion) is open. The contribution should report on an attempt to prove the completeness of the criterion. It is made in a similar way as it was done for undirected acyclic graphs: through developping an equivalent "separation" criterion, choosing properly a subgraph for every graphical dependency and finding a suitable construction of a probability measure for such a subgraph.

Akimichi TAKEMURA

Independence Bartlett correction of nested likelihood ratio tests

It is well known that the likelihood ratio statistic is Bartlett correctable. Here we decompose a likelihood ratio statistic into one degree of freedom likelihood ratio statistics considering nested hypotheses. Let nested hypotheses be given as

$$\begin{aligned}
H_0 : &\quad \theta_1 = \theta_{10}, \ldots, \theta_p = \theta_{p0} \\
H_0 : &\quad \theta_1 : \text{ free }, \theta_2 = \theta_{20}, \ldots, \theta_p = \theta_{p0} \\
H_0 : &\quad \theta_1, \theta_2 : \text{ free }, \theta_3 = \theta_{30}, \ldots, \theta_p = \theta_{p0} \\
&\quad \ldots \\
H_0 : &\quad \theta_1, \ldots, \theta_{p-1} : \text{ free }, , \theta_p = \theta_{p0} \\
H_0 : &\quad \theta_1, \ldots, \theta_p : \text{ free.}
\end{aligned}$$

Let λ_j be the likelihood ratio test for testing H_{j-1} vs. H_j. Our result is as follows. Under H_0 the λ_js are independent up to the order $O(1/n)$ and they are independently Bartlett correctable.

Nanny WERMUTH

Applications of graphical chain models

Variables in four larger empirical investigations are presented with first orderings in terms of pure responses, intermediate variables and pure explanatory variables. Specific features of the research questions and the data sets and analyses in terms of graphical chain models are pointed out. In a second part of the lecture solutions to the following questions one can ask for a given generating process of the data are answered. Suppose the process can be described in terms of a directional acyclic graph, when does this generating process imply

1. $Y_i \not\perp\!\!\!\perp Y_j \mid Y_C$ for an arbitrarily selected variable pair (Y_i, Y_j) and a selected conditionary set of variables Y_C

2. an edge in the covariance graph is present, i.e. $Y_i \not\perp\!\!\!\perp Y_j$ and

3. an edge in the implied concentration graph is present, i.e. $Y_i \not\perp\!\!\!\perp Y_j \mid$ all remaining variables.

(based on joint work with D.R. Cox)

Henry P. WYNN

Some algebraic topology in statistics: the Voronoi method for inclusion-exclusion inequalities

The classical inclusion-exclusion lemma for n sets $\{A_i\}_{i=1}^{n}$ takes the form

$$I(\bigcup A_i) = \sum_{i=1}^{n} I(A_i) - \sum_{i<j} I(A_i \cap A_j) + \ldots (-1)^n I(\bigcap_{i=1}^{n} A_i)$$

where $I(A)$ is the indicator function of A.

In Naiman and Wynn (Ann. Stat. 1992) we show that for A_i = Ball centered x_i, radius r, in $I\!R^d$ there is a "depth" $m = d + 1$ version (compared to the depth n, in general). The new version takes the form

$$\begin{aligned} I(\bigcup A_i) &= \sum_{i=1}^{n} I(A_i) - \sum_{(i,j)\in K_2} I(A_i \cap A_j) + \sum_{(i,j,k)\in K_3} I(A_i \cap A_j \cap A_k) \\ &\quad - \ldots (-1)^m \sum_{(i_1,i_2,\ldots,i_k)\in K_m} I(A_{i_1} \ldots \cap_{i=1}^{n} A_{i_k}) \quad \forall r \end{aligned}$$

where summation is over the subsimplicies of the Delaunay simplicial complex, the dual of the Voronoi relaxation of $I\!R^d$ using x_i (the center of the balls).

More recent work shows that the upper and lower bounds obtained by truncating the classical IE (generalized Bonferroni bounds) can be generalised to the above equation. All the theory is based on the socalled DISCRETE TUBE property: the subsimplicial complex for $\{A_i\}$ $A_i \ni x$ has Euler characteristic = 1; or more strongly is simply connected. Many collections have the discrete tube property: balls, hyperplanes, halfspaces etc. For all of these the theory works. A particularly important case is half-spaces. In that case $\bigcup A_i \equiv$ complement of a convex set. It is expected that in this case there will be application to multiple comparison and reliability.

Berichterstatter: Adalbert Wilhelm (Augsburg)

MONDAY, 3 July 1995

Chair: Michael PERLMAN

09:00–09:15 Friedrich PUKELSHEIM: Welcome

09:15–10:15 Steffen LAURITZEN: On the statistical theory of graphical models

10:25–10:55 Judea PEARL: Graphs, structural models, and counterfactuals

Chair: Eckart SONNEMANN

11:15–11:45 Dan GEIGER: A characterization of the Dirichlet, Normal and Wishart distribution, with applications to learning Bayesian networks

11:55–12:25 Henry WYNN: Some algebraic topology in statistics: The Voronoi method for inclusion–exclusion inequalities

Chair: Steen ANDERSSON

16:00–16:30 Muriel CASALIS: The Lukacs–Olkin–Rubin characterization of the Wishart distribution on symmetric cones

16:35–17:05 Hélène MASSAM: Lattice conditional independence models on symmetric cones

Chair: Ted ANDERSON

17:20–17:50 Akimichi TAKEMURA: A (frequentist) proof of independent Bartlett correction of nested likelihood ratio tests

17:55–18:25 Kai Tai FANG: On multivariate vertical density representation

TUESDAY, 4 July 1995

Chair: Joe WHITTAKER

09:00–10:00 David COX: General remarks on graphical models

Chair: Antony UNWIN

10:30–11:20 Nanny WERMUTH: Applications of graphical chain models

11:30–12:20 Glenn SHAFER: Alternative causal interpretations of multivariate models

Chair: Hélène MASSAM:

16:00–16:45 Jesper MADSEN: Models combining group symmetry and conditional independence in a multivariate normal distribution

16:50–17:35 Søren Tolver JENSEN: Classification of full exponential transformation models

17:40–18:25 Jim MALLEY: Statistical applications of Jordan algebras

WEDNESDAY, 5 July 1995

Chair: Nanny WERMUTH

09:00–09:40 Göran KAUERMANN: On a dualization of graphical Gaussian models

09:50–10:30 David MADIGAN: Bayesian analysis of Markov equivalence classes of acyclic digraphs

Chair: Dan GEIGER

10:55–11:35 Chris MEEK: New methods for selecting graphical models

11:45–12:25 Michael PERLMAN: Lattice conditional independence models and acyclic digraphs

THURSDAY, 6 July 1995

Chair: Rosemary BAILEY

09:00–10:00 Phil DAWID: The analysis of distributive block structures for experimental layouts

Chair: Joe EATON

10:30–11:20 John MARDEN: Multivariate ranks and their application to invariant normal models

11:30–12:20 František MATÚŠ: On conditional independence and polymatroids

Chair: Morten FRYDENBERG

16:00–16:45 Milan STUDENÝ: Chain graphs as models of probabilistic conditional independence structures

16:50–17:35 Jan KOSTER: Markov properties of non-recursive

causal models

17:40–18:25 Peter SPIRTES: Generalized graphical representations of marginalization, conditionalization and feedback

FRIDAY, 7 July 1995

Chair: Hans BRØNS

09:00–09:40 Dietrich von ROSEN: Multivariate linear models with certain covariance structures

09:50–10:30 Joe EATON: Invariance and strong inconsistency

Chair: Robb MUIRHEAD

10:55–11:35 Muni SRIVASTAVA: Admissibility of the inverse regression estimator in the multi–univariate linear calibration model

11:45–12:25 Kenneth NORDSTRÖM: Stochastic Schur convexity

Chair: Ludger RÜSCHENDORF

13:30–14:10 Don RICHARDS: Algebraic methods towards higher-order probability inequalities

14:20–15:00 Ingram OLKIN: Correlation analysis of extreme observations from a multivariate normal distribution

15:00 Michael PERLMAN: Farewell

E-mail adresses

Ted Anderson	twa@playfair.stanford.edu
Steen Arne Andersson	standers@ucs.indiana.edu
Rosemary A. Bailey	r.a.bailey@qmw.ac.uk
Hans K. Brøns	hbrons@math.ku.dk
Muriel Casalis	casalis@cict.fr
David R. Cox	drcox@vax.ox.ac.uk
Phil Dawid	dawid@stats.ucl.ac.uk
Morris L. Eaton	eaton@stat.umn.edu
Kai Tai Fang	KTFANG@FISHER.HKBU.EDU.HK
Morten Frydenberg	morten@biostat.aau.dk
Dan Geiger	dang@cs.technion.ac.il
Herbert Heyer	heyer@uni-tuebingen.de
Søren Tolver Jensen	soerent@math.ku.dk
Göran Kauermann	kauerman@cs.tu-berlin.de
Jan Koster	koster@soc.fsw.eur.nl
Steffen Lauritzen	steffen@iesd.auc.dk
David Madigan	madigan@stat.washington.edu
Jesper Madsen	jmadsen@math.ku.dk
James Malley	jmalley@helix.nih.gov
John I. Marden	j-marden@uiuc.edu
Hélène Massam	massamh@vm1.yorku.ca
František Matúš	matus@utia.cas.cz
Chris Meek	cm1x@andrew.cmu.edu
Robb J. Muirhead	robb@umich.edu
Kenneth Nordström	nordstro@cc.helsinki.fi
Ingram Olkin	iolkin@playfair.stanford.edu
Judea Pearl	judea@cs.ucla.edu
Michael Perlman	michael@stat.washington.edu
Friedrich Pukelsheim	Pukelsheim@Uni-Augsburg.De
Don Richards	drichards@virginia.edu
Dietrich von Rosen	dietrich@math.uu.se
Chris A. Rowley	c.a.rowley@open.ac.uk
Ludger Rüschendorf	ruschen@buffon.mathematik.uni-freiburg.de
Glenn Shafer	gshafer@andromeda.rutgers.edu
Eckart Sonnemann	elenz@uni-trier.de
Peter Spirtes	ps7z@andrew.cmu.edu
Muni S. Srivastava	srivastava@utstat.toronto.edu
Milan Studený	studeny@utia.cas.cz
Akimichi Takemura	takemura@e.u-tokyo.ac.jp
Antony Unwin	unwin@uni-augsburg.de
Nanny Wermuth	wermuth@animal.sowi.uni-mainz.de
Joe Whittaker	joe.whittaker@lancaster.ac.uk
Adalbert Wilhelm	wilhelm@uni-augsburg.de
Graham Wood	g.wood@cqu.edu.au
Henry P. Wynn	h.p.wynn@city.ac.uk

Tagungsteilnehmer

Prof.Dr. Theodore W. Anderson
Department of Statistics
Stanford University
Sequoia Hall

Stanford , CA 94305-4065
USA

Prof.Dr. Arthur Cohen
Dept. of Statistics
Rutgers University
Hill Center, Busch Campus

New Brunswick , NJ 08903
USA

Prof.Dr. Steen Arne Andersson
Dept. of Statistics
Indiana University

Bloomington , IN 47405
USA

Prof.Dr. David Cox
Nuffield College
Oxford University

GB-Oxford OX1 1NF

Prof.Dr. Rosemary A. Bailey
School of Mathematical Sciences
Queen Mary & Westfield College
Mile End Rd.

GB-London E1 4NS

Prof.Dr. A. Philip Dawid
Dept. of Statistical Sciences
University College
Gower Street

GB-London WC1E 6BT

Prof.Dr. Hans Kristian Brons
Inst. of Mathematical Statistics
University of Copenhagen
5 Universitetsparken

DK-2100 Copenhagen

Prof.Dr. Morris L. Eaton
Dept. of Statistics
University of Minnesota
270 Vincent Hall
206 Church Street S. E.

Minneapolis , MN 55 455
USA

Muriel Casalis
Laboratoire de
Statistiques et Probabilites
Universite Paul Sabatier
118 route de Narbonne

F-31062 Toulouse Cedex

Prof.Dr. Kai-Tai Fang
Department of Mathematics
Hong Kong Baptist College
224 Waterloo Road

Kowloon
HONG KONG

Prof.Dr. Morten Frydenberg
Department of Biostatistics
University of Aarhus
Building 2 C
Norrebrogade 44

DK-8000 Aarhus C

Prof.Dr. Jan T.A. Koster
Department of Sociology
Erasmus University Rotterdam
P.O.Box 1738

NL-3000 DR Rotterdam

Prof.Dr. Dan Geiger
Computer Science Department
TECHNION
Israel Institute of Technology

Haifa 32000
ISRAEL

Prof.Dr. Steffen Lilholt Lauri[illegible]n
Institut for elektroniske syst[illegible]r
Aalborg Universitetscenter
Fredrik Bajers Vej 7E

DK-9220 Aalborg 0

Prof.Dr. Herbert Heyer
Mathematisches Institut
Universität Tübingen
Auf der Morgenstelle 10

72076 Tübingen

Prof.Dr. David Madigan
Department of Statistics
University of Washington
Box 35 43 22

Seattle , WA 98195-4322
USA

Prof.Dr. Soren Tolver Jensen
Institut for Matematisk Statistik
Kobenhavns Universitet
Universitetsparken 5

DK-2100 Kobenhavn

Jesper Madsen
Institut for Matematisk Statistik
Kobenhavns Universitet
Universitetsparken 5

DK-2100 Kobenhavn

Dr. Göran Kauermann
Institut für Quantitative Methoden
Fachbereich Informatik
TU Berlin
Franklinstr. 28/29

10587 Berlin

Prof.Dr. James D. Malley
National Institute of Health
Division of Computer Research and
Technology
Bldg. 12A / Rm. 3039

Bethesda , MD 20892 5656
USA

Prof.Dr. John I. Marden
Dept. of Statistics
University of Illinois
725 S Wright St.

Champaign , IL 61820
USA

Dr. Kenneth Nordström
Dept. of Mathematics and Statistics
University of Maryland
Baltimore County·Campus

Baltimore , MD 21228-5398
USA

Prof.Dr. Helene Massam
Laboratoire de Statistique et
Probabilites
Universite Paul Sabatier
118, route de Narbonne

F-31062 Toulouse Cedex

Prof.Dr. Ingram Olkin
Department of Statistics
Stanford University
Sequoia Hall

Stanford , CA 94305-4065
USA

Prof.Dr. Frantisek Matus
Inst. of Info. Theory & Automation
Academy of Sciences of the Czech R.
Pod Vodarenskou vezi 4
P.O.Box 18

182 08 Prague 8
CZECH REPUBLIC

Prof.Dr. Judea Pearl
Dept. of Mathematics
University of California
405 Hilgard Avenue

Los Angeles , CA 90024-1555
USA

Prof.Dr. Chris Meek
BH 135
Carnegie Mellon University

Pittsburgh , PA 15213
USA

Prof.Dr. Michael D. Perlman
Department of Statistics
University of Washington
Box 35 43 22

Seattle , WA 98195-4322
USA

Prof.Dr. Robb John Muirhead
Department of Statistics
The University of Michigan
1447 Mason Hall

Ann Arbor , MI 48109-1027
USA

Prof.Dr. Friedrich Pukelsheim
Lehrstuhl für Stochastik und ihre
Anwendungen
Institut für Mathematik der Univ.

86135 Augsburg

Prof.Dr. Donald St. P. Richards
Division of Statistics
University of Virginia
Cabell Drive
Halsey Hall

Charlottesville , VA 22903
USA

Prof.Dr. Eckart Sonnemann
Fachbereich IV
Abteilung Mathematik
Universität Trier

54286 Trier

Prof.Dr. Dietrich von Rosen
Department of Mathematics
University of Uppsala
P.O. Box 480

S-75106 Uppsala

Prof.Dr. Peter Spirtes
Department of Philosophy
Carnegie Mellon University
Schenley Park

Pittsburgh , PA 15213-3890
USA

Prof.Dr. Chris A. Rowley
The Open University
Parsifal College
527 Finchley Road

GB-London NW3 7BG

Prof.Dr. M.S. Srivastava
Dept. of Statistics
University of Toronto
100 St. George Street

Toronto ONT M5S 1A1
CANADA

Prof.Dr. Ludger Rüschendorf
Institut für Mathematische
Stochastik
Universität Freiburg
Eckerstr. 1

79104 Freiburg

Prof.Dr. Milan Studeny
Inst. of Info. Theory & Automation
Academy of Sciences of the Czech R.
Pod Vodarenskou vezi 4
P.O.Box 18

182 08 Prague 8
CZECH REPUBLIC

Prof.Dr. Glenn Shafer
Dept. of Account. & Inform. Systems
Faculty of Management
Rutgers University
180 University Avenue

Newark , NJ 07102
USA

Prof.Dr. Akimichi Takemura
Faculty of Economics
University of Tokyo
Bunkyo-ku

Tokyo 113
JAPAN

Anthony R. Unwin
School Systems
Data Studies
Trinity College

Dublin
IRELAND

Adalbert Wilhelm
Lehrstuhl für Rechnerorientierte
Statistik und Datenanalyse
Institut für Mathematik der Uni.
Augsburg

86135 Augsburg

Prof.Dr. Nanny Wermuth
Psychologisches Institut
Universität Mainz

55099 Mainz

Dr. Graham R. Wood
Dept. of Mathematics
University of Canterbury
Private Bag 4800

Christchurch 1
NEW ZEALAND

Dr. Joe C. Whittaker
Mathematics Department
School of Engineering, Computing
and Mathematical Sciences
Lancaster University

GB-Lancaster LA1 4YF

Prof.Dr. Henry P. Wynn
School of Mathematics
The City University
Northampton Square

GB-London , ECIV OHB