

MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

T a g u n g s b e r i c h t 8/1981

Medizinische Statistik
8.2. - 14.2.1981

Die Tagung fand unter Leitung von Herrn J. Berger (Hamburg) und Herrn E. Walter (Freiburg) statt.

Wenn auch im Mittelpunkt der Vorträge Probleme der klinischen Prüfung standen, so reicht doch das Spektrum der Vorträge in alle Teilbereiche der medizinischen Statistik. Am Dienstag abend wurde in einer zusätzlichen Sitzung ausgehend von einem Referat von Herrn Immich ausführlich über klinische Prüfungen diskutiert. Am Donnerstag abend wurde in Kurzreferaten die Situation des Unterrichts in Biomathematik für Mediziner in den einzelnen Ländern dargestellt: Frankreich (Frau Hoang), USA (Mantel), Großbritannien (Anderson), Niederlande (Rümke), Österreich (Bauer), Schweiz (Riedwyl), Bundesrepublik (Abt). Dabei zeigte es sich, daß in allen Ländern mit Ausnahme der Bundesrepublik der Unterricht in den ersten vorklinischen Semestern stattfindet.

Vortragsauszüge

K. ABT:

On the non-validity of elementary null hypotheses rejected at the nominal significance level of the "Rüger-Test"

Rüger (Metrika 1978) proved that in testing N elementary null hypotheses H_{0i} , $i = 1, 2, \dots, N$, each at level α^* (special case of his theorem with levels α_i^* , $i=1, \dots, N$), a sharp upper bound for the probability α of rejecting the valid global null hypothesis H_0 {All N H_{0i} valid}, when testing H_0 against H_1 {Not all N H_{0i} valid} under the condition of $m^* \geq M$ of the N H_{0i} to show significance at α^* , is given by $\alpha = \min(\frac{N}{M} \alpha^*; 1)$. From this and with prechosen α , $\alpha^* = \frac{M}{N} \alpha$ results as nominal significance level for each of the N tests.

The experimenter likes to know whether the $m^* \geq M$ rejected H_{0i} are non-valid. Assuming that he will accept a prechosen number cN among the m^* rejected H_{0i} to be valid, a coefficient $q = q(N, M, c, \alpha, \pi_{\max})$ is derived such that the probability of rejecting H_0 at level $\alpha' = q\alpha \leq \alpha$ if m H_{0i} , $0 < m < M - cN$, are non-valid among the $m^* \geq M$ H_{0i} , each rejected at $q\alpha^* \leq \alpha^* = \frac{M}{N} \alpha$, is kept below a prechosen value π_{\max} , for ex. $\pi_{\max} = 0,10$.

J.A. ANDERSON:

Logistic methods for prediction in medicine

Logistic discrimination provides a fairly easy way of making clinical inferences or decisions. Suppose that there are two diseases (H_1 and H_2) and p variables $x^T = (x_1, x_2, \dots, x_p)$ under consideration as potential discriminators. The only assumption is that the conditional probability of H_1 given \tilde{x}

$$\text{pr}(H_1 | \tilde{x}) = \frac{e^z}{1 + e^z},$$

where $z = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$. Here $\beta_0, \beta_1, \dots, \beta_p$ are unknown parameters that are to be estimated from the data, using the likelihood principle. The prediction variables (x_j) may be continuous or discrete.

An example of the method is given for the pre-operative prediction of post-operative deep vein thrombosis in gynaecological surgery. This is used in a prospective study to aid the decision about preoperative, prophylactic treatment.

Extensions of the method to three or more disease groups and ordered disease categories are introduced and inferential problems are described which suggest the use of penalised maximum likelihood estimation.

P. ARMITAGE:

Cross-over Trials

Cross-over Trials with two treatments, and two periods of administration for each patient, are increasingly used for small trials of treatments for short-term alleviation of chronic diseases. The usual Gaussian model gives t-Tests for treatment (T) and period (P) effects based on within-subject variation, and a t-Test between subjects for the interaction (TP). The latter is not necessarily done to differences in carry-over effects. Equivalently, the data may be analysed as a split-plot experiment. Baseline readings, immediately before the first period and during the washout period between treatments, may be used as covariates (a) to improve sensitivity of the TP test, and (b) to improve sensitivity of a test between treatments based on the first period only (which may be necessary if TP exists). A short example illustrates the danger of removing from the analysis patients with incomplete readings. The consequent bias may be wrongly interpreted as a TP effect.

P. BAUER:

Zur Qualitätskontrolle in Beobachtungsreihen

Es soll geprüft werden, ob die sequentiell anfallenden Beobachtungen x_1, x_2, x_3, \dots unabhängige Realisierungen einer mit bekanntem Erwartungswert θ und bekannter Varianz σ^2 normalverteilten Zufallsvariablen sind. Eine Zusatzforderung ist dabei, daß das Auftreten von Abweichungen von diesen Voraussetzungen nach möglichst wenigen Beobachtungen erkannt werden soll.

Dieses Problem tritt etwa in der Qualitätskontrolle im klinischen Laboratorium oder in der laufenden Kontrolle von belastenden Umweltfaktoren auf.

Es wird eine Übersicht über Methoden zur Entdeckung von Abweichungen im Erwartungswert θ auf der Basis von Mittelwertskarten, gleitenden Summen, geometrisch gewichteten kumulativen Summen sowie kumulativen Summen gegeben und auf einige jüngere Entwicklungen eingegangen.

J. BERGER:

Analysis of categorial data by linear models

For analyzing contingency tables log-linear models are now "in". However there exist two disadvantages of this technique, one is that the models are limited to a hierarchical procedure and on the other hand the interpretation of the estimated parameters is difficult to understand by the consumers. An alternative approach, specially if you are mainly interested to find out how a dependent variable is influenced by a given set of other variables, is the modified weighted least square regression model (Grizzle, Starmer, Koch 1969)

$$M^{-1}F = M^{-1}Xb + M^{-1}E$$

with F vector of the function of the dependent variable, X the design matrix, b the coefficient vector, E the vector of the error terms and M a weight matrix with the characteristic $M = M^T$ and $MM = V(F)$, $V(F)$ the covariance matrix of F.

If a suitable coding for the categories of the independent variables is used the estimated coefficients are submitted to an effect interpretation comparable to the analysis of variance and interaction terms can be formulated in the design matrix as conditional main effects, so that the consumer has no difficulties in understanding the results of the analysis. The feature is illustrated on data where the log linear model didn't led to a simplification of the underlying structure.

E. BRUNNER:

A three-factor mixed model for ordinal data

Gegeben seien n Objekte (unabhängige Versuchswiederholungen), je Objekt p verschiedene Behandlungen, die jeweils von denselben m Prüfern in eine Rangfolge gebracht werden. Die Faktoren "Behandlung" und "Prüfer" seien fest; eine Wechselwirkung zwischen Behandlung und Prüfer wird nicht angenommen. Man erhält das Modell:

$$Y_{ijk} = \mu + \alpha_k + \beta_j + \pi_i + P_{ik} + W_{ij} + \epsilon_{ijk},$$

für jede Stufenkombination (i,k) werden die Ränge von 1 bis p vergeben:

$$R_{ijk} = \sum_{j'=1}^p c(Y_{ijk} - Y_{ij'k}), \quad c(x) = 0 \text{ für } x < 0 \\ \text{und } 1 \text{ für } x \geq 0.$$

Zum Testen der Nullhypothese $H_0: \beta_1 = \dots = \beta_p$ wird die Statistik $R_{.j}$ gewählt, und man erhält unter einigen Regularitätsbedingungen an das Modell, daß die quadratische Form

$$K_n = n^2 (p-1) \frac{\sum_{j=1}^p (\bar{R}_{.j} - \bar{r})^2}{\sum_{i=1}^n \sum_{j=1}^p (\bar{R}_{ij} - \bar{r})^2} \underset{n \rightarrow \infty}{\overset{H_0}{\rightsquigarrow}} \chi^2_{p-1} \text{ unter } H_0$$

verteilt ist. K_n ist nur asymptotisch verteilungsfrei, da die Korrelationen der Ränge zwischen den Prüfern unbekannt sind und aus dem Experiment konsistent geschätzt werden müssen.

K. DIETZ:

The Shanghai model for vivax malaria

During a course on Mathematical Models for Infectious Diseases which was organised by the World Health Organisation from 13-24 October 1980 in Shanghai the participants developed a new model for the transmission and control of vivax malaria. The parameters have been estimated partly from historical data on artificial infections in volunteers, partly from present day records of mosquito density. The different threshold theorems for constant and for periodic

contact rates hold. The model takes into account nonhomogeneous participation of the population at subsequent drug applications.

L. EDLER:

Median und arithmetisches Mittel zensierter Beobachtungen

Liegen zensierte Daten, z.B. zensierte Überlebenszeiten, vor, so sind Median oder arithmetisches Mittel sowohl aller als auch nur der unzensierten Beobachtungen verzerrte Schätzungen.

Betrachtet man den Median M oder das Mittel μ als Funktional der Überlebensfunktion $S: M = S^{-1}(\frac{1}{2})$ bzw. $\mu = \int_0^{\infty} S(x) dx$ so erhält man Schätzer \hat{M} bzw. $\hat{\mu}$ indem man S z.B. durch den Produkt-Limit Schätzer \hat{S} ersetzt. Mittels der simultanen Konfidenzbänder von Hall und Wellner (1979) für S lassen sich leicht auch Konfidenzschranken für M und weiterer Quantile ausrechnen. Eine Schätzung von μ über die empirische Überlebensfunktion ist schwieriger. Ansätze von Sander (1975) und Susarla und van Ryzin (1980) werden skizziert. Eine andere Möglichkeit einer unverzerrten Schätzung von μ besteht darin, die Beobachtungen so zu transformieren, daß die Verzerrung verschwindet. Mit Hilfe der Restlebenserwartung kann man so zu einem plausiblen Schätzer kommen, über dessen Eigenschaften allerdings wenig bekannt ist.

U. FELDMANN:

On a fishery problem discussed in Biometrics

P. Sprent and G.R. Dolby (Biometrics 1980) discussed the use of lawlike relationships in fishery research.

They stated that the Maximum-Likelihood-Estimator in the case of linear functional relationship has poor asymptotic properties (inconsistency) and special empirical parameters (geometric mean slope, coefficient of correlation) have no statistical interpretation in the functional model.

In this paper, one was reminded of the fact, that the linear structural relationship approach doesn't have this disadvantage.

W. GUNSELMANN:

Schätzung der "crude survival probability" bei (abhängigen) konkurrierenden Risiken in Anlehnung an Kaplan und Meier

Auf eine Population mögen $k \geq 2$ Todesursachen einwirken. Im Zufallsvariablenvektor (T, J) sei T die Überlebenszeit und J die Todesursache. Es wird das Problem behandelt, $P(T < t, J = j)$ für alle $t \geq 0$ und alle $j = 1, 2, \dots, k$ auch im Fall von "censoring" zu schätzen.

Unter gewissen Unabhängigkeitsannahmen für das censoring werden Schätzer der für $t \geq 0$ definierten Funktionen $P(T < t, J = j)$ für jedes $j = 1, 2, \dots, k$ vorgeschlagen mit der Eigenschaft, daß ihre Summe über j zu jedem Zeitpunkt t dem Kaplan-Meier-Schätzwert für $P(T < t, \cdot)$ entspricht. Der vorgeschlagene Schätzer erweist sich als Maximum/Likelihood Schätzer im Sinn von Kaplan und Meier.

J. HERMANS:

A calibration problem in microbiology

In clinical microbiology one has to determine the susceptibility to antibiotics of strains of bacteria. Two quantitative methods are available. A dilution method, leading to a minimum inhibitory concentration (MIC), and a diffusion method, leading to an inhibition zone diameter (ZD). In practice one has to relate the ZD (routine determination) to the MIC (reference determination). To the MIC values are associated break-points (usually two), indicating a susceptible category (low values), a resistant category (high values) and an intermediate category (doubt whether the drug is effective).

Relating ZD to MIC is commonly done with linear regression analysis. Several advantages and disadvantages of the use of regression analysis in this context will be discussed.

P. IHM:

Clusteranalyse bei Beobachtungsdaten

Dateien von Beobachtungsdaten geht im allgemeinen die häufig bei Datenanalysen vorausgesetzte Rechtecksgestalt ab.

Ein Beispiel stellen Krankheitsverläufe bei Patienten dar, die zu verschiedenen Zeiten auf mehr oder weniger verschiedene Weise untersucht wurden. Es besteht kein Zweifel, daß ein Arzt die Ähnlichkeit von Verläufen beurteilen kann. Wie soll man diese Ähnlichkeit aber messen? Relativ einfach ist die Lösung bei Vorliegen einer Faktorenstruktur. Eine Methode wird vorgestellt.

W. KÖPCKE:

Repeated significance tests with censored survival data

Legal and ethical considerations lead to the necessity of interim analysis and early stopping rules in clinical trials. It can be shown that six teststatistics developed for censored survival data (Armitage (1960), Breslow (1972), Canner (1977), Gehan (1965), Peto (1979), Pocock (1977)) are very similar under certain conditions. According to a suggestion by Pocock (1980) to perform interim analysis after an equal number of observed deaths it can be shown by simulation that the best strategy (α and β error, mean stopping time) is to test at least five to ten times.

W. LEHMACHER:

Recent developments in Configural Frequency Analysis

A short outline of Configural Frequency Analysis (CFA) of Krauth and Lienert is given. Then, the relationship between CFA and the log-linear model is shown. CFA can be looked at as an analysis of residuals in the model of total independence of all variables. Strategies are suggested, which allow for analysing the structure of dependencies, which causes an outlier (type) detected by the CFA. In this manner, the advantages of both methods are combined to an efficient data analysis. Then, an improvement of the binomial tests of CFA is given; these alternative test are based on the generalized hypergeometrical distribution of a cell frequency and are not conservative. Finally, it is suggested to work with the Holm method instead of the Bonferroni method.

N. MANTEL:

Cautions on the use of medical data bases

A survey of striking examples is given which show a bias if the relative risk is estimated from existing data bases. If one does not take into account the selection mechanisms leading to the inclusion of the data into the study even contradictory interpretations can suggest itself depending on the model used.

J. MICHAELIS:

Prospective study on pregnancy and child development (evaluation problems and recent results)

From 1964 until 1972 15 000 pregnancies were recruited for the study by 21 cooperating obstetric and paediatric clinics. The study may be described as a prospective cohort study with some retrospective elements. Evaluation problems are due to the length of the study, lack of hypotheses at the beginning and methodological difficulties. Evaluation strategy consists out of exploratory and confirmative data analysis. Results are demonstrated for the analysis of teratogenicity of sexual hormones and antinauseant drugs taken within the early period of pregnancy. Calculation of corresponding odds ratios and appropriate confidence intervals based on matched pairs gives evidence of the safety of those drugs.

D. MORGENSTERN:

Der χ^2 -Test für die Multinomialverteilung bei fehlenden Beobachtungen mit Anwendung auf Kontingenztafeln

Wenn N_1, \dots, N_S multinomial $(n, p_1 \dots p_S)$ verteilt sind, soll $p_i = \pi_i(\theta)$ geprüft werden, ohne (oBdA) N_1 (und n) zu benutzen. Die bedingte Verteilung von N_2, \dots, N_S bei festgehaltener Summe und die gewöhnliche χ^2 -Test-Theorie führen auf die übliche Testgröße und Maximum-Likelihood-Schätzgleichung, wenn die Schätzwerte $\hat{N}_1 = \frac{\pi_1}{1-\pi_1} \left(\sum_2^S N_i \right)$ und

$\hat{N} = \hat{N}_1 + \sum_{i=1}^s N_i$ eingesetzt werden; Freiheitsgrad = $s-2-r$. Der Nichtzentralitätsparameter wird angegeben und Anwendungen auf die $a \times b$ -Tafel, wo der bekannte Victor-Schätzer entsteht und wo auch zusammengefaßte Felder sein dürfen, werden besprochen.

A. NEISS:

Remarks on sample size determination for clinical trials

The problems which arise in sample size determination for clinical trials is demonstrated by a study in which the author was involved as the statistical consultant. The following topics are discussed: the choice of the criterion to measure the effect of the therapies, the formulation of the mathematical model and the hypotheses to be tested, the choice of the statistical test, the derivation of the (asymptotic) power function and the formula for the sample size as a function of the power $1-\beta$, the significance level α , the difference Δ of the effects which is of medical importance and the probability p of success for one of the both treatments. It is shown that the main problems are not mathematical ones but problems like how to choose appropriate values for α and β . A decision theoretic approach to this problem is discussed.

S.T. POCKOCK (vorgelesen von P. Armitage):

Interim analyses for clinical trials

Interim analyses of long term clinical trials are useful ways of improving the adherence to protocol and methods of data processing, and they permit early detection of large differences between treatment effects. However, if repeated significance tests of a null hypothesis are done at different stages of the accumulation of data, the 'overall' significance level for the whole procedure will exceed the nominal level for each test. For given power requirements, and a given number of interim analyses, 'group sequential plans' are available, which use the same 'nominal' level throughout. These are based on a

Gaussian distribution of response, with known variance, but are robust for other situations. Plans which use small probability levels for early stages may be attractive, but have larger average sample number against alternatives for which the power is high.

P. ROEBRUCK:

Zur Varianzkomponentenschätzung in regulären linearen Modellen

Das Problem der nicht-negativen Schätzung von Varianzkomponenten wurde von J. Hartung kürzlich in recht befriedigender Weise gelöst, indem er die Existenz und Eindeutigkeit der Minimum-Norm-Minimum-Bias-nichtnegativen-quadratischen Schätzer bewies. Die Berechnung dieser Schätzer ist jedoch nicht ganz einfach; sie führt selbst für balancierte Modelle auf nichtlineare Gleichungssysteme. A. Infante hat Lösungen für eine Reihe von balancierten Modellen explizit angegeben. Hier wird gezeigt, wie man die oben genannten Schätzer in der größeren Klasse regulärer Modelle meist durch einfache Überlegungen erhalten kann. Auf dem Weg dahin fällt noch eine einfache Darstellung der Minimum-Norm-erwartungstreuen-quadratischen Schätzer für diese Modelle ab.

CHR.L. RÜMKE:

Comparison of two proportions in two phases

Outlines are given of a procedure for comparing two probabilities p_1 and p_2 ($p_1 \leq p_2$) in two phases. At first two samples are compared with $n=20$ each if the anticipated difference between p_2 and p_1 is ca. 0.30, or $n=40$ each if it is ca. 0.20. The trial is only completed with the required sample size (given $p_1, p_2, \alpha=0.05$ one-sided, and $\beta=0.05$) if there are more successes in the second sample than in the first one. Some simulations showed that with this procedure and with the usual test, the influence upon α is small, and the influence upon β is acceptable. This procedure has the advantage that ca. 55 % of the

trials will be stopped at the end of the first phase if $p_2 = p_1$, whereas the loss of power is $< 5\%$. Especially for clinical trials this may be an attractive property.

Preliminary results show that for a two-sided test with anticipated $|p_2 - p_1| \geq 0.30$ in the first phase two samples with $n=30$ each might be a good suggestion, and with $|p_2 - p_1| \geq 0.20, n=55$. In this case trials are to be completed only if the numbers of successes in the two samples differ by more than two.

S. SCHACH:

A method of analysis for a random coefficient of regression model

We consider a model $y_{ij} = \sum_{k=1}^p b_{ik} \varphi_k(t_{ij}) + e_{ij}$ ($i = 1, \dots, n$; $j = 1, \dots, m_i$), where $\varphi_1, \dots, \varphi_p$ are known functions, $b_i := (b_{i1}, \dots, b_{ip})$ are i.i.d. and the e_{ij} satisfy the usual assumptions on error terms. We propose estimators for $\beta := E b_i$, $\Delta := \text{cov}(b_i)$ and the error variance σ^2 which are unbiased and consistent. Furthermore, $\hat{\beta}$ is asymptotically normally distributed as $n \rightarrow \infty$, with an estimable covariance matrix. The application of the method to a particular medical problem is discussed.

B. SCHNEIDER:

Multivariate analysis of cross-over designs

The results of cross-over designs can be interpreted as realisations of a general multivariate linear model (GMLM) by looking at the p results y_{ij1}, \dots, y_{ijp} obtained from the patient j in the group i at the p different treatment periods as realisations of a response vector \vec{y}_{ij} and assuming the linear model:

$$\vec{y}_{ij} = \begin{pmatrix} y_{ij1} & \mu_{i1} \\ \vdots & \vdots \\ y_{ijp} & \mu_{ip} \end{pmatrix} + \vec{e}_{ij} = X B + \vec{e}$$

with $\vec{e}_{ij} \sim N(\vec{0}, \downarrow)$ independent for different i, j .

As "group" is defined the "arrangement group", characterised by the $p!$ possible arrangements of the p treatments to the p periods.

In practice one is less interested in the behaviour of the mean values $\bar{\mu}_i$ but in certain contrasts representing "treatment effects", "period effects" etc. Such contrasts are estimable functions of the parameters $\bar{\mu}_i : \psi_r = \sum_{i,k} a_{rik} \mu_{ik}$ estimated by $\hat{\psi}_r = \sum_{i,k} a_{rik} \bar{y}_{i.k}$

The elements of the estimation space form the total space of the possible estimation functions. For $p=2$ this total space is represented by 4 orthogonal coefficient vectors (a_{ik}):

		group		1		2	
		period		1	2	1	2
		treatment		A	B	B	A
ψ_1	general mean	1	1	1	1	1	1
ψ_2	treatment	1	-1	-1	1	1	1
ψ_3	period	1	-1	1	-1	-1	-1
ψ_4	interaction	1	1	-1	-1	-1	-1

In the case of more than 2 periods there must be inducted constraints to the parameters μ_{ik} .

E. SONNEMANN:

The GMANOVA-model with constraints for the analysis of growth curves

The GMANOVA-model introduced by Potthoff and Roy (1964) cannot be applicated directly to many growth curve problems in praxis, because it does not allow any constraints given in the parameter space. Thus this model

$$X = A\mu B + e, e \sim N(0, I \otimes \Sigma), (\mu, \Sigma) \in R^{m \times q} \times R^{p \times p}$$

$n \times p$

has to be extended by adding constraints concerning the space of the μ 's which may be given in the form $G\mu H = 0$. If one wants to test the null hypothesis $H_0 : K\mu L = 0$, then under reasonable regularity conditions for the known matrices G, H, K, L one can show that the canonical form of this generalized GMANOVA-problem is as follows:

$Z \sim N(\theta, I \circ \dagger)$, $V \sim W_p(\dagger, n-a)$ independent,
axp pxp

$$\theta = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & 0 \\ \theta_{21} & \theta_{22} & 0 & 0 \\ \theta_{31} & \theta_{32} & 0 & 0 \end{pmatrix}, \text{ and } H_0 : \theta_{32} = 0.$$

Recently Banken (1980) has pointed out that the results due to Kariya (1978) can be extended to this problem, and he has found a unique locally best invariant test.

R.E. TARONE:

Test statistics and estimators based on derivatives of the log-likelihood function

In many inference problems involving complicated likelihood functions, simple test statistics may be constructed from derivatives of the log-likelihood function evaluated under the null hypothesis. Tests based on these statistics, often called score statistics, will be asymptotically locally optimal provided certain regularity conditions are satisfied. Simple estimators, closely related to both the score test statistics and maximum likelihood estimators, may be easily constructed. These methods are illustrated in examples involving the analysis of censored survival data using Cox's proportional hazards model, the analysis of stratified data from retrospective or prospective epidemiological studies, and the use of historical control information in testing for a trend in proportions.

P. TAUTU:

Ein graphentheoretisches Modell für den Biotransformationsprozeß chemischer Karzinogene

Das Problem der Biotransformation chemischer Präkarzinogene wird im Rahmen der Signalflußgraphentheorie (Mason 1953, 1956) behandelt.

Zwischen linearen algebraischen Gleichungen und einem Masongraphen G_α besteht folgender Zusammenhang: Die Unbekannten der Gleichungen entsprechen den Punkten des Graphen, die linearen Relationen zwischen ihnen den gerichteten Kanten, mit einer gegebenen $n \times n$ Koeffizientenmatrix A kann ein gerichteter Graph $G_\alpha = G(V, E, \alpha)$ konstruiert werden, wobei V die Menge der Punkte, E die Menge der Kanten und α eine Numerierung der Punkte von 1 bis n ist. Jedem Punkt v_i entspricht eine Reihe i von A und jeder Kante (v_i, v_j) ein von Null verschiedenes Element a_{ij} von A .

Ein Grund der Anwendung des Masongraphen G_2 ist, den Zusammenhang zwischen Eingang (Punkt v_1) und Ausgang (Punkt v_n) zu finden. Hier wird dies als die Beziehung zwischen aufgenommenem Präkarzinogen und eigentlichem Karzinogen, das eine konvalente Bindung mit dem DNS-Molekül eingehen kann, gedeutet. Die dem Punkt v_n zugeordnete reelle Zahl ist ein Maß für das in der Zelle wirksame Karzinogen.

Lineare Graphen haben bei der Analyse chemischer Netzwerke den Vorzug, daß der zugehörige gerichtete Graph ohne die zugehörigen Gleichungen direkt gebildet werden kann. Da eine Menge der hierbei auftretenden Vorzeichenmatrizen einem Netzwerk entspricht, ist das Problem der Vorzeichenstabilität eng mit der Netzwerkstabilität verknüpft.

H. THÖNI:

Graphische Darstellung eines Ellipsoids

Für die graphische Darstellung eines Ellipsoids $x'Ax = 0$ als Ellipse $\xi'B\xi = 0$ in einer Zeichenebene (ξ, η) werden die Elemente einer Matrix C zur Berechnung der Koeffizientenmatrix $B = CAC'$ explizit angegeben.

Es gilt

$$C = - \frac{1}{d_1 Ad_1} \begin{pmatrix} d_1 A (d_2 d_1 - d_1 d_2) \\ d_1 A (d_3 d_1 - d_1 d_3) \\ d_1 A (d_4 d_1 - d_1 d_4) \end{pmatrix}$$

d_1, d_2, d_3, d_4 sind die Spalten der Matrix $D = T^{-1}$, welche als

Inverse aus der Beziehung

$$\begin{pmatrix} t \\ \xi \\ \eta \\ 1 \end{pmatrix} = \begin{pmatrix} t_1 & t_2 & t_3 & 0 \\ \lambda \cos \alpha & \mu \cos \beta & \nu \cos \gamma & 0 \\ \lambda \sin \alpha & \mu \sin \beta & \nu \sin \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = T x$$

hervorgeht. Die Hilfsgröße t wird eliminiert und fällt aus den Elementen von C heraus.

(Die Ergebnisse entstanden in Zusammenarbeit mit Prof. Dr. G. Bach, Hohenheim.)

J. TIMM:

Some problems concerning short biological time series

A biomathematical model concerning the oscillations of biological subsystems such as special cells is deduced in order to obtain a valid basis for theoretical and empirical approaches to the phenomenon of their stability. The model takes account for possible feedback and white noise superimposing an "internal clock"-oscillation. As the experimental observations in vitro are often restricted to a short discrete time scale, the following mixed regression-autoregression model seems to be appropriate:

$$\varphi(B) y_i = f(t_i) + \epsilon_i$$

φ being a polynomial function on the algebra generated by the backshift operator B , f a deterministic periodic function and ϵ_i a white noise process.

Some statistical problems arising with short asynchronous time series under specialisations of this model are treated and examples (fluorescence activity in gland cells as well as acetylcholinesterase activity of erythrocytes) are given.

D.M. TITTERINGTON:

Comparison of discrimination techniques applied to data from head-injured patients

A report will be presented of a collaborative project involving statisticians from Glasgow, Nottingham and Rotterdam.

Several techniques for discriminant analysis are applied to a set of data from patients with severe head injuries, for the purpose of prognosis. The data are such that multidimensionality, continuous, binary and ordered categorical variables and missing data must be coped with. The various methods are compared using criteria of prognostic success and reliability. In general, performance varies more with choice of the set of predictor variables than with that of the discriminant rule.

H.J. TRAMPISCH:

Der mittlere Fehler von Zuordnungsregeln

Es sei $Z = (X, I)$ eine Zufallsvariable, $I: \Omega \rightarrow \{1, 2\}$, $X: \Omega \rightarrow H = \{1, \dots, s\}$. Gegeben seien die Wahrscheinlichkeiten $P(I = i) = 0,5$ und $P(X = x | I = i) = p_i(x)$. Eine Zuordnungsregel ist eine Partition von $H: D = \langle D_0, D_1, D_2 \rangle$. Der aktuelle Fehler von D ist gegeben durch:

$$t(D) = 0,5 \sum_H (p_2(x) I_{D_1}(x) + p_1(x) I_{D_2}(x) + 0,5 (p_1(x) + p_2(x)) I_{D_0}(x)),$$

wobei $I_{D_i}(x)$ die Indikator-Funktion von D_i bezeichnet. Es seien Stichproben $(x_1, 1), \dots, (x_n, 1)$ und $(x_1, 2), \dots, (x_n, 2)$ (i.i.d.) jeweils vom Umfang n gegeben. Mit $N_i(x) = \{v | (x_v, i_v) = (x, i)\}$ ist die multinomiale Regel $\hat{M} = \langle \hat{M}_0, \hat{M}_1, \hat{M}_2 \rangle$ gegeben durch:

$$x \in \hat{M}_1 \Leftrightarrow N_1(x) > N_2(x), x \in \hat{M}_2 \Leftrightarrow N_2(x) > N_1(x), x \in \hat{M}_0 \Leftrightarrow N_1(x) = N_2(x).$$

Der mittlere Fehler von \hat{M} ist der Erwartungswert von $t(\hat{M})$:

$$E_n(t(\hat{M})) = 0,5 \sum_H (p_2(x) P_n(x \in \hat{M}_1) + p_1(x) P_n(x \in \hat{M}_2) + 0,5 (p_1(x) + p_2(x)) P_n(x \in \hat{M}_0)).$$

Es wird gezeigt, daß gilt: $E_{n+1}(t(\hat{M})) \leq E_n(t(\hat{M}))$.

An einem Gegenbeispiel wird demonstriert, daß der mittlere Fehler nicht für alle Zuordnungsregeln und für alle zugrunde liegenden Verteilung eine monoton fallende Funktion in n ist (für Unabhängigkeitsregel gilt dies nicht).

J. WAHRENDORF:

Survival analysis with genetic covariates

Information on a genetic system of individuals can be analysed by using genetic models such as the Hardy-Weinberg model of random mating. Maximum likelihood procedures are available for this purpose. For the problem of analysing the dependency of censored survival data from a genetic system it is proposed to estimate at each failure time the parameters of the genetic model from all individuals at risk. This allows to calculate expected numbers of deaths at each failure time and compare them with the observed numbers along the classical lines of survival analysis.

H.E. WICHMANN:

About the circadian rhythm

A cellular structure with a circadian rhythm (CR) can be treated mathematically as a disturbed system with a periodic perturbation.

For the undisturbed system Kendall (1948) and Takahashi (1966 and 1968) have proposed to subdivide the cell cycle into n age classes where aging is understood as a Poisson process. This leads to a system of n linear differential equations,

$$\dot{\vec{y}} = A \vec{y}.$$

Assuming equal transition probabilities within all sub-compartments of the cell phases (G_1, S, G_2, M), all model parameters can be identified by measured transit times and their variances.

The perturbation caused by CR modifies the system to $\dot{\vec{y}} = A(t) \vec{y}$ with a periodic matrix $A(t) = A(t+T)$. Using data from mouse epithelium, it can be shown that it is sufficient to restrict the influence of CR to only one parameter, $\tau_S(t)$, the S-phase transit time.

In a second step the role of the labelling of mouse epithelium cells by $^3\text{HTdR}$ is analysed. It can be shown that 3 assumptions about the influence of the label are sufficient

to understand the data. These are (1) a short block of labelled cells in the S-phase (or the death of some of them) (2) a shift of unlabelled cells from G_1 (or G_0) into the S-phase without changing their total cell cycle time (3) the loss and reutilization of the label of some of the labelled cells (which may die or not). These hypotheses can be read from one part of the data while in addition a second part is available for testing them.

Berichterstatter: J. Bammert

Tagungsteilnehmer

- Herrn
Prof. Dr. E. Brunner
Lehrstuhl f. Med. Statistik
d. Univ. Göttingen
Windausweg 2
3400 Göttingen
- Herrn
Prof. Dr. K. Abt
Klinikum der Universität
Zentrum f. Med. Informat.
Abt. f. Biomathematik
Theodor-Stern-Kai 7
6000 Frankfurt 70
- Herrn
Dr. J.A. Anderson
Dept. of Statistics
Univ. of Newcastle-upon-Tyne
Newcastle-upon-Tyne
- Herrn
Prof. P. Armitage
Dept. of Biomathematics
Univ. of Oxford
Pusey St.
Oxford OX12JZ
- Herrn
Doz. Dr. P. Bauer
Institut f. Med. Statistik
Preindlg. 18/21
1090 Wien
- Herrn
Prof. Dr. J. Berger
UKE, Abt. f. Math. i.d. Medizin
Martinistraße 52
2000 Hamburg 20
- Herrn
Prof. Dr. H.-H. Bock
Institut f. Statistik u.
Wirtschaftsmath. d. RWTH
Pontstr. 51/Postf. 112
5100 Aachen
- Herrn
Prof. Dr. K. Dietz
Universität Tübingen
Institut f. Med. Biometrie
Hallstattstr. 6
7400 Tübingen 1
- Herrn
Dr. L. Edler
DKFZ Inst. f. Dokumentation,
Information u. Statistik
Postfach 10 19 49
6900 Heidelberg 1
- Herrn
Prof. Dr. U. Feldmann
Universität Heidelberg
Med. Statistik, Biomath.
u. Informationsverarb.
Theodor-Kutzer-Ufer
6800 Mannheim 1
- Herrn
Dr. W. Gunselmann
Institut f. Med. Statistik
und Dokumentation
Waldstraße 6
8520 Erlangen
- Herrn
Dr. J. Hermans
Universität Leiden
Abt. Medizinische Statistik
Wassenaarseweg 80
Leiden/Niederlande
- Frau
Thi Minh Thu Hoang
INSERM, U 170 (Statistique)
16^{bis}, avenue Paul-Vaillant-
Couturier
94800 Villejuif, France

Herrn
Prof. Dr. P. Ihm
Inst. f. Med. Biol. Statistik
u. Dokumentation
Ernst-Giller-Str. 20
3550 Marburg/Lahn

Herrn
Prof. Dr. H. Immich
Institut f. Med. Dokumentation,
Statistik u. Datenverarbeitung
Im Neuenheimer Feld 325
6900 Heidelberg 1

Herrn
Prof. Dr. H.J. Jesdinsky
Med. Einrichtung der Univ.
Institut f. Med. Statistik
u. Biomathematik
Moorenstraße 5
4000 Düsseldorf 1

Herrn
Prof. Dr. H. Klinger
Institut f. Statistik u.
Dokumentation d. Univ.
Universitätsstr. 1
4000 Düsseldorf 1

Herrn
Dr. W. Koepcke
Universität München
ISB
Marchionistr. 16
8000 München 70

Herrn
Dr. W. Lehmacher
Institut f. Med. Informatik
u. Systemforschung
(MEDIS) d. GSF
Arabellastr. 4
8000 München 81

Herrn
Prof. Dr. N. Mantel
George Washington University
Biostatistics Center
7979 Old Georgetown Rd.
Bethesda, MD 20014 USA

Herrn
Prof. Dr. J. Michaelis
Institut f. Med. Statistik
u. Dokumentation d. Univ.
Langenbeckstr. 1/Postf. 3960
6500 Mainz 1

Herrn
Prof. Dr. D. Morgenstern
Universität Hannover
Im Dorffeld 58
3005 Hemmingen

Herrn
Prof. Dr. Dr. A. Neiß
Institut f. Med. Statistik
u. Epidemiologie d. TU
Sternwartstr. 2
8000 München 80

Herrn
Dr. H. Nowak
Fa. Goedecke AG, Abt. Biometrie
Postfach 569
7800 Freiburg i.Br.

Herrn
Prof. Dr. H. Riedwyl
Universität Bern
Institut f. Mathem. Statistik
Sidlergasse 5
CH-3012 Bern/Schweiz

Herrn
Dr. P. Roebuck
Abt. Med. Statistik u.
Dokumentation der RWTH
Goethestr. 23
5100 Aachen 1

Herrn
Prof. Dr. Chr. L. Rümke
Afdeling Medische Statistiek
der Vrije Universiteit
Van der Boechorststraat 7
Postbus 7161
Amsterdam/Niederlande.

Herrn
Prof. Dr. S. Schach
Lehrstuhl Math. Statistik I
d. Universität Dortmund
Postfach 50 05 00

4600 Dortmund 50

Herrn
Dr. H. Scheurlen
Institut f. Med. Dokumentation,
Statistik u. Datenverarb.
Im Neuenheimer Feld 325
6900 Heidelberg 1

Herrn
Prof. Dr. B. Schneider
Med. Hochschule Hannover
Abt. u. Lehrstuhl f. Biometrie
K. Wiechert-Allee 9/Postf. 610180
3000 Hannover 61 (Kleefeld)

Herrn
Prof. Dr. E. Sonnemann
Universität Dortmund
Abteilung Statistik
Postfach 50 05 00

4600 Dortmund 50

Herrn
Dr. R.E. Tarone
Math. Stat. und
Applied Math. Sec.
National Cancer Inst.
Bethesda/Maryland
U.S.A. 20014

Herrn
Prof. Dr. P. Tautu
Dt. Kz. Inst. f. Dokumentation,
Information u. Statistik
Im Neuenheimer Feld 280
6900 Heidelberg 1

Herrn
Prof. Dr. H.-P. Thöni
Universität Hohenheim
Institut f. Angew. Mathem.
u. Statistik
Postfach 106
7000 Stuttgart 70

Herrn
Prof. Dr. J. Timm
Universität Bremen
Mathematik und Informatik
Achterstraße, Postf. 330 440

2800 Bremen

Herrn
D. M. Titterington
Dept. of Statistics
Univers. of Glasgow

Glasgow G12 8QW

Herrn
Dr. H.J. Trampisch
Med. Einrichtung d. Univ.
Institut f. Med. Statistik
u. Biomathematik
Moorenstraße 5
4000 Düsseldorf 1

Herrn
Prof. Dr. N. Victor
Abtl. Biomathematik
d. Universität Gießen
Frankfurter Str. 100

6300 Lahn-Gießen 1

Herrn
Dr. J. Wahrendorf
Int. Agency f. Res. on Cancer
Unit of Biostatistics WHO
150, Cours Albert Thomas

F-69372 Lyon-Cedex 2

Herrn
Prof. Dr. E. Walter
Universität Freiburg
Inst. f. Med. Dokumentation
und Statistik
Stefan-Meier-Str. 26
7800 Freiburg i.Br.

Herrn.
Prof. Dr. H.E. Wichmann
Univ.-Klinikum LFI 02
Med. Universitätsklinik
Joseph-Stelzmann-Str. 9

5000 Köln 41