

MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

Tagungsbericht 11/1971

Medizinische Statistik

8. bis 13.3.1971

Die diesjährige Tagung über medizinische Statistik fand unter der Leitung von Professor S.Koller (Mainz) statt.

Der große Teilnehmerkreis umfaßte im wesentlichen Mathematiker und Mediziner, die überwiegend in Instituten für Medizinische Statistik arbeiten. Die Breite des Anwendungsbereiches statistischer Verfahren in der Medizin spiegelte sich in der unterschiedlichen Thematik der einzelnen Vorträge wider; lediglich einige Probleme der Biosignalverarbeitung (insbesondere EKG und Phonokardiogramm) wurden zusammenhängend an einem Tag in mehreren Vorträgen behandelt. Entsprechend den Arbeitsschwerpunkten der einzelnen Teilnehmer standen in den Referaten und Diskussionen methodologische oder Anwendungsgesichtspunkte im Vordergrund. Zusätzlich zu den nachstehend aufgeführten Vorträgen fanden noch Diskussionen über Fragen der Ausbildung in medizinischer Statistik (Unterrichtsgestaltung, Lernzielkatalog) und der biostatistischen Vaterschaftsbegutachtung statt.

Teilnehmer

H.Bloedhorn, Freiburg
W.Bühler, Heidelberg
H.O.Christl, Wiesbaden
R.Dannhorn, Hannover
J.Dudeck, Gießen
van Eimeren, Ulm
U.Feldmann, Hamburg
B.Friedel, Mainz
P.Friedrichs, Hannover
P.Gille, Hannover
G.Glettler, Gießen

E.B.Greiser, Hannover
E.Hansert, München
E.Hartmann, Berlin
F.Hehl, Heidelberg
H.Hermann, Hannover
G.W.Himmelmann, Marburg
W.Höbel, Gießen
V.Hodapp, Mainz
L.Horbach, Mainz
J.Hornung, Berlin
E.Hultsch, Hannover

P.IHM, Marburg
 H.J.JESDINSKY, Freiburg
 J.KRAUTH, Düsseldorf
 G.A.LIENERT, Düsseldorf
 U.LUDWIG, Tübingen
 J.MICHAELIS, Mainz
 D.NORGENSTERN, Freiburg
 G.OBERHOFFER, Bonn
 REPGES, Gießen
 H.RIEDWYL, Bern
 H.RUNDFELDT, Hannover
 D.SCHNEIDER, Hannover
 U.SCHULZ, Marburg
 I.STRASSER, Freiburg
 M.TÜRÜK, Hannover
 K.ÜBERLA, Ulm
 E.WALTER, Freiburg
 P.O.WHITE, London
 CH.ZYWIETZ, Hannover

H.IMMICH, Heidelberg
 H.KLINGER, Düsseldorf
 H.J.LANGE, München
 O.LUDWIG, Bad Nauheim
 S.MANDEL, Genf
 K.MIESKE, Heidelberg
 A.NEISS, München
 S.J.PÜPPEL, München
 G.REISSIG, Bochum
 CHR.L.RÜMKE, Amsterdam
 E.SCHEIDT, Mainz
 J.SCHULTE-MÖNTING, Freiburg
 R.STADLER, Bergen/Ffm
 H.THÖNI, Zürich
 M.TÜCKE, Marburg
 N.VICTOR, München
 G.WETTER, Mainz
 F.X.WOHLZOGEN, Wien

Vortragsauszüge

H.L.CHRISTL: Eine Erweiterung der Diskriminanzanalyse

Es werden n-dimensionale Zufallsgrößen $X(t)$ mit multivariabler Normalverteilung $N(\mu(t), \Sigma)$ betrachtet. $X(t)$ gehöre einer von zwei Populationen Π_1, Π_2 an, jede mit $N(\mu_i(t), \Sigma)$.

Bei unbekannter a-priori Verteilung werden die Klassifikationsgebiete mit dem Minimax-Prinzip bestimmt; d. h.: Bestimme ein $b(t)$, so daß

$$c(2|1) \int_{b(t)}^{\infty} dN(\beta_1(t), \sigma_1^2) = c(1|2) \int_{-\infty}^{b(t)} dN(\beta_1(t), \sigma_2^2)$$

wobei $E(D) = \beta_i \quad (X(t) \in \Pi_i) \quad i=1,2$
 $Var(D) = \sigma_i^2$

D = Diskriminanzfunktion

Es wird ein Weg gezeigt, wie über die Schätzung von $E(\hat{D})$, $Var(\hat{D})$ die Bestimmung von $b(t)$ erfolgen kann.

Vortragsauszüge

J. DUDECK: Computeranalyse von Phonokardiogrammen

Am gesunden Herzen kommt es während jeder Herzaktion zu charakteristischen Schallphänomenen, deren wichtigste als 1. und 2. Herzton bezeichnet werden. Diese entstehen in engem Zusammenhang mit dem Schluß der Herzklappen durch die durch den Klappenschluß bedingte Geschwindigkeitsänderung der während der Herzaktion bewegten Blutmenge. Die graphische Aufzeichnung der Schallphänomene wird als Phonokardiogramm bezeichnet. Die auftretenden Schwingungen sind sehr kurzfristig und wechseln in Abhängigkeit von der Phase der Herzaktionen. Zur Bestimmung der im Phonokardiogramm enthaltenen Frequenzen wurde die von Stumpff angegebene Methode der Phasenzeitdiagramme verwendet, die neben der Schätzung der Frequenz die Bestimmung der Phase und des Zeitpunktes des Auftretens erlaubt. In ersten Untersuchungen konnten enge Korrelationen zwischen der Frequenz von Herztonanteilen und Herzvolumina gezeigt werden. Ein entsprechendes Programm, das diesen Zusammenhang zur einfachen Schätzung von Herzvolumina ausnutzen soll, ist in Entwicklung.

Van EIMEREN: Ist eine Bedarfs- und Erfolgskontrolle im Fach "Medizinische Statistik" sinnvoll?

"Medizinische Statistik" ist eine angewandte Wissenschaft. Ihre Aufgabe liegt in der optimalen Anwendung bzw. Entwicklung theoretischer statistischer Ansätze auf eine konkrete Forschungssituation in der Medizin.

Z. Zt. existieren keine vertrauenswürdigen Angaben über Stand und Bedeutung der "med. Stat." im Rahmen der Wissenschaftsentwicklung einzelner medizinischer Fachdisziplinen.

Man sollte sich die Informations-Grundlage schaffen, Aussagen über Sinn und Unsinn bestimmter Anstrengungen im Fachbereich nach Art einer cost-benefit-Analyse machen zu können.

Eine beurteilende Literaturanalyse der deutschen med. Fachliteratur könnte eine Informationsquelle dazu sein. Die Benutzung des science citation index könnte Aufschluß über die jeweiligen bevorzugten Zitatquellen geben.

U. FELDMANN: Linearisierung spezieller Anpassungsprobleme bei nichtlinearen Schätzungen

Insbesondere in der klinischen Chemie tritt das Problem auf, Schätzungen aufgrund einer vorgegebenen diskreten Stichprobe zu realisieren.

Als Schätzfunktion wird im allgemeinen eine Funktion

$$y(x) = g(a_1, a_2, \dots, a_m; x)$$

gewählt, deren Parameter durch die Bedingung, daß die Euklidische Norm der Differenz von Wirkungsvektor und Schätzvektor ein Minimum bildet, bestimmt werden.

Im allgemeinen führt diese Bedingung zu einem nichtlinearen Gleichungssystem in den Parametern a_k .

Es wird ein Verfahren angegeben, spezielle Schätzfunktionen, die unter einem gewissen Gesichtspunkt linearisierbar sind, iterativ zu approximieren.

P. FRIEDRICHS: Digitale Filterung zur Glättung von Biosignalen

Digitale Filterung ist ein Prozeß, der eine Menge von Eingabedaten x_t in eine Menge von Ausgabedaten y_t umwandelt durch die lineare Beziehung

$$y_t = \sum_{l=-m}^{\infty} c_l x_{t-l}$$

Es wird die Wirkungsweise eines digitalen Filters auf die verschiedenen Frequenzanteile eines Signals in Abhängigkeit von den Koeffizienten c_i gezeigt.

Eberhard GREISER: Zur Strategie der Bewertung von Arzneimittelnebenwirkungen

Wegen der Möglichkeit erheblicher sozialer und ökonomischer Folgen von Arzneimittelnebenwirkungen schwererer Art ist es erforderlich, schon vor einem endgültigen Urteil über die mögliche Kausalität zu Kriterien zu gelangen, die mögliche Entscheidungen (Zurückziehung des Mittels, Rezeptpflicht) stützen.

Zur Abschätzung der Bezugspopulation für Patienten mit vermueteter Nebenwirkung bieten sich Repräsentativerhebungen an. Am Beispiel der primär vaskulären pulmonalen Hypertonie als vermueteter Nebenwirkung des Appetitzüglers Aminorex fumarat (Menocil) wird die Brauchbarkeit dieser Methode demonstriert, um zu Anhaltspunkten über Altersabhängigkeit und Dosisabhängigkeit zu gelangen.

Es wird vorgeschlagen, in Anlehnung an die von H. Bradford Hill entwickelten Kriterien eine Terminologie zur Skalierung des Begriffes "Kausalität" zu entwickeln.

E. HANSERT: Cluster-Analyse

Cluster- und faktorenanalytische Methoden pflegen von paarweisen Ähnlichkeiten bzw. Distanzen auszugehen. Es ist zu fragen, ob man durch Einsetzen "globaler Distanzen" zwischen Tripeln, Quadrupeln usw. nicht befriedigendere Resultate erhalten kann. Als Axiome für globale Distanzmaße werden vorgeschlagen: seien x_1, x_2, \dots, x_k Elemente einer Beobachtungsskala, dann

1. $d^{(k)}(x_1, \dots, x_k) \geq 0$; das Gleichheitszeichen soll dann und nur dann gelten, wenn alle d_i im wesentlichen identisch sind.
2. $d^{(k)}(x_1, \dots, x_k)$ sei symmetrisch
3. Gilt für jedes $i = 1, 2, \dots, k - 1$, daß $d^{(2)}(x_i, x'_k) \geq d^{(2)}(x_i, x_k)$ ist, dann sei $d^{(k)}(x_1, x_2, \dots, x'_k) \geq d^{(k)}(x_1, x_2, \dots, x_k)$.

Konkretes Beispiel ist das Problem, bei Fragebögen mit alternativen Items typische Antwortstrukturen zu finden.

G.W.HIMMELMANN: Orthogonale Polynome und Normbereich am Beispiel des arteriellen Blutdrucks

Das verallgemeinerte Verfahren der Regression mittels orthogonaler Polynome kann zur Berechnung der Parameter mehrdimensionaler Normalverteilungen, die Funktionen dritter Größen sind, benutzt werden.

Die Lage eines mehrdimensionalen Maßpunktes in einem solchen Normbereich kann für einen gegebenen Wert der unabhängigen Größe durch den als Perzentil angegebenen Abstand vom Verteilungsschwerpunkt sowie den zugehörigen Winkel ausreichend beschrieben werden.

Am Beispiel des aufgrund von Literaturangaben errechneten zweidimensionalen Normbereichs des arteriellen Blutdrucks, der eine Funktion von Alter und Geschlecht ist, wird das Verfahren hinsichtlich seiner Aussagekraft für die Therapiekontrolle anhand von zwei praktischen Fällen demonstriert.

L.HORBACH: Analyse der Beziehung zwischen dem zeitlichen Einfall der Vorhofkontraktion und der Herzleistung bei dissoziierter Tätigkeit von Herzvorhöfen und -kammern

Beim Herzgesunden wird die Tätigkeit von Herzvorhöfen und -kammern durch den normalen Erregungsablauf im Reizleitungssystem koordiniert. Bei Patienten mit vollkommen unterbrochener Überleitung vom Vorhof auf die Kammern (Block 3.Grades), erhaltenem Sinusrhythmus der Vorhöfe und liegendem Schrittmacher in der rechten Kammer wurde ein Beobachtungsplan zur quantitativen Untersuchung der Reservoir- und Kontraktionsfunktion des linken Vorhofs aufgestellt. Die Frequenzen von Vorhof und Kammer wurden so abgestimmt, daß jeweils nur eine oder keine Vorhofaktion zwischen zwei Kammerystolen fiel. Bei 17 Patienten wurde dabei über eine bestimmte Zeit EKG, Phonokardiogramm und Pulsdruck in einer herznahen Arterie registriert. Es fand sich - in gewisser Analogie zur Schwebung bei physikalischen Wellen mit etwas unterschiedlicher Frequenz - eine in allen Fällen gleichartige Abhängigkeit zwischen Abstand von Vorhof- und Kammererregung

und der Pulsdruckamplitude als Maß der Herzleistung, die jeweils durch ein nach Tschebyscheff approximiertes Polynom dargestellt werden kann. Der Mittelwert der Pulsdrucke ohne vorausgehende Vorhofkontraktion, d.h. bei alleiniger ungestörter Reservoirfunktion, ergibt eine Beziehungslinie, die eine Abschätzung des Einflusses der Reservoirbildung einerseits und der Kontraktionsfunktion der Vorhöfe andererseits auf die Herzleistung ermöglicht. Die Abhängigkeit des optimalen Abstandes der Vorhof- von der Schrittmachererregung läßt sich als lineare Regression darstellen.

Diese Analyse wurde in Zusammenarbeit mit dem Kardiologen PD Dr. JUST und cand.med.K.SCHUSTER, II.Mediz.Klinik der Universität Mainz (Direktor: Prof.Dr.P.Schölmerich), durchgeführt.

Joachim HORNING: Über mehrdimensionale Konfidenzbereiche

Um Konfidenzbereiche für den BERNOULLI-Parameter p zu finden, invertiert man Annahmehbereiche für die stetige Zufallsvariable $y = x + R$, wobei $x \mathcal{L}(N, p)$ und R Rechteck $(0,1)$ -verteilt ist. Eine Funktion $\varphi_r(p|y) = r\varphi_0(p|y) + (1-r)\varphi_1(p|y)$, genannt Konfidenzverteilung, hat die Eigenschaft, daß für alle Konfidenzbereiche G gilt:

$$\int_G \varphi_r(p|y) dp = W(p \in G / p) = 1 - \alpha \quad (1)$$

Im Falle von 3 möglichen Versuchsergebnissen kann man wiederum eine Funktion $\varphi_r(p|y) = \sum r_i \varphi_i(p|y)$ mit der Eigenschaft (1) finden, wenn G zugelassener Konfidenzbereich in der p_0 - p_1 -Ebene ist. Dabei ist $p = (p_0, p_1, p_2)$, $\sum p_i = 1$, $x = (x_0, x_1, x_2)$, $\sum x_i = N$, $r = (r_0, r_1, r_2)$, $\sum r_i = 1$, $y = x + r$. Eine Projektion eines Gebiets G auf die p_0 -Achse ist in der Weise möglich, daß man die Randverteilung der 2-dimensionalen Konfidenzverteilung bildet. Dieses Verfahren liefert "beinahe" die eindimensionalen Konfidenzbereiche.

P. IHM: Zum Problem der Schilddrüsendiagnostik

Der Marburger Poliklinik für Nuklearmedizin werden von den Ärzten, die Serum zur PBJ- und T₃-Bestimmung einschicken, auf einem Fragebogen zusätzlich sieben klinische Merkmale angegeben. In den meisten Fällen werden die PBJ- und T₃-Bestimmungen zur Differentialdiagnose zwischen euthyreoter vegetativer Dystonie, Adenom und Hyperthyreose vorgenommen. Die Wahrscheinlichkeiten der drei Diagnosen können mittels des Bayes'schen Theorems berechnet werden. Die Praxis zeigt aber, daß die Ärzte die Patienten entsprechend dem Grad der Überfunktion der Schilddrüse in sechs Kategorien einteilen; die Anfertigung eines Szintigrammes wird bei allen Kategorien mit Ausnahme der ersten Kategorie ("sicher euthyreot") empfohlen, bei dieser ersten Kategorie wird die Szintigraphie in Abhängigkeit vom Vorliegen eines Tastbefundes vorgeschlagen. Es wird versucht, aufgrund eines mathematischen Ansatzes vorauszusagen, in welche Kategorie ein Patient eingeteilt wird.

P.IHM: Ein Verfahren zur chronologischen Reihung

Die Aufgabe bei der chronologischen Reihung besteht darin, Objekte, die durch einen von der Zeit t abhängigen Zufallsvektor gekennzeichnet sind, in die richtige zeitliche Reihenfolge zu bringen. Der Erwartungswert für das i-te Element des Vektors (der im vorgetragenen Beispiel aus den Häufigkeiten bestimmter Stiltypen-Index i - bei archäologischen Funden - Index k - besteht) soll durch die Funktion

$$y_i(t) = \frac{p_i}{\sqrt{2\pi}} e^{-\frac{(t-\mu_i)^2}{2}}$$

dargestellt werden. t bezeichnet die Zeit, μ_i den zeitlichen Mittelpunkt des Stiltyps i, p_i die Häufigkeit des Stiltyps i insgesamt. Für jedes Paar erhält man

$$\alpha_{ij} = 2 \log \int_{-\infty}^{\infty} y_i(t) y_j(t) dt = 2 \log p_i + 2 \log p_j - \log \pi - \frac{1}{2}(\mu_i - \mu_j)^2$$

Die Zeilen- und Gesamtmittel der α_{ij} seien mit $\bar{\alpha}_i$ und $\bar{\alpha}$ bezeichnet, dann gilt:

$$\beta_{ij} = \alpha_{ij} - \bar{\alpha}_i - \bar{\alpha}_j + \bar{\alpha} = (\mu_i - \bar{\mu})(\mu_j - \bar{\mu})$$

wobei $\bar{\mu}$ das Mittel der μ_i ist.

m sei der dominante Eigenvektor der Matrix $B = (\beta_{ij})$, dann kann man f_k^m als Schätzwert für die Zeit verwenden, aus der der beobachtete Fund stammt. Bisher wurden zur chronologischen Reihung u.a. Verfahren verwendet, bei denen mit $\alpha_{ij} = f_i' \cdot f_j$ gerechnet wurde. f variiert bei diesem Ansatz auf einer hufeisenförmigen Kurve. Durch die Verwendung von $\alpha_{ij} = \log f_i' \cdot f_j$ wird die Variation auf einer Geraden erreicht.

und der Pulsdruckamplitude als Maß der Herzleistung, die jeweils durch ein nach Tschebyscheff approximiertes Polynom dargestellt werden kann. Der Mittelwert der Pulsdrucke ohne vorausgehende Vorhofkontraktion, d.h. bei alleiniger ungestörter Reservoirfunktion, ergibt eine Beziehungslinie, die eine Abschätzung des Einflusses der Reservoirbildung einerseits und der Kontraktionsfunktion der Vorhöfe andererseits auf die Herzleistung ermöglicht. Die Abhängigkeit des optimalen Abstandes der Vorhof- von der Schrittmachererregung läßt sich als lineare Regression darstellen.

Diese Analyse wurde in Zusammenarbeit mit dem Kardiologen PD Dr. JUST und cand.med.K.SCHUSTER, II.Mediz.Klinik der Universität Mainz (Direktor: Prof.Dr.P.Schölmerich), durchgeführt.

Joachim HORNING: Über mehrdimensionale Konfidenzbereiche

Um Konfidenzbereiche für den BERNOULLI-Parameter p zu finden, invertiert man Annahmebereiche für die stetige Zufallsvariable $y = x + R$, wobei $x \mathcal{L}(N, p)$ und R Rechteck $(0,1)$ -verteilt ist. Eine Funktion $\varphi_r(p|y) = r \varphi_0(p|y) + (1-r) \varphi_1(p|y)$, genannt Konfidenzverteilung, hat die Eigenschaft, daß für alle Konfidenzbereiche G gilt:

$$\int_G \varphi_r(p|y) dp = W(p \in G / p) = 1 - \alpha \quad (1)$$

Im Falle von 3 möglichen Versuchsergebnissen kann man wiederum eine Funktion $\varphi_r(p|y) = \sum r_i \varphi_i(p|y)$ mit der Eigenschaft (1) finden, wenn G zugelassener Konfidenzbereich in der p_0 - p_1 -Ebene ist. Dabei ist $p = (p_0, p_1, p_2)$, $\sum p_i = 1$, $x = (x_0, x_1, x_2)$, $\sum x_i = N$, $r = (r_0, r_1, r_2)$, $\sum r_i = 1$, $y = x + r$. Eine Projektion eines Gebiets G auf die p_0 -Achse ist in der Weise möglich, daß man die Randverteilung der 2-dimensionalen Konfidenzverteilung bildet. Dieses Verfahren liefert "beinahe" die eindimensionalen Konfidenzbereiche.

P. IHM: Zum Problem der Schilddrüsendiagnostik

Der Marburger Poliklinik für Nuklearmedizin werden von den Ärzten, die Serum zur PBJ- und T₃-Bestimmung einschicken, auf einem Fragebogen zusätzlich sieben klinische Merkmale angegeben. In den meisten Fällen werden die PBJ- und T₃-Bestimmungen zur Differentialdiagnose zwischen euthyreoter vegetativer Dystonie, Adenom und Hyperthyreose vorgenommen. Die Wahrscheinlichkeiten der drei Diagnosen können mittels des Bayes'schen Theorems berechnet werden. Die Praxis zeigt aber, daß die Ärzte die Patienten entsprechend dem Grad der Überfunktion der Schilddrüse in sechs Kategorien einteilen; die Anfertigung eines Szintigrammes wird bei allen Kategorien mit Ausnahme der ersten Kategorie ("sicher euthyreot") empfohlen, bei dieser ersten Kategorie wird die Szintigraphie in Abhängigkeit vom Vorliegen eines Tastbefundes vorgeschlagen. Es wird versucht, aufgrund eines mathematischen Ansatzes vorauszusagen, in welche Kategorie ein Patient eingeteilt wird.

P.IHM: Ein Verfahren zur chronologischen Reihung

Die Aufgabe bei der chronologischen Reihung besteht darin, Objekte, die durch einen von der Zeit t abhängigen Zufallsvektor gekennzeichnet sind, in die richtige zeitliche Reihenfolge zu bringen. Der Erwartungswert für das i-te Element des Vektors (der im vorgetragenen Beispiel aus den Häufigkeiten bestimmter Stiltypen-Index i - bei archäologischen Funden - Index k - besteht) soll durch die Funktion

$$\psi_i(t) = \frac{p_i}{\sqrt{2\pi}} e^{-\frac{(t-\mu_i)^2}{2}}$$

dargestellt werden. t bezeichnet die Zeit, μ_i den zeitlichen Mittelpunkt des Stiltyps i, p_i die Häufigkeit des Stiltyps i insgesamt. Für jedes Paar erhält man

$$\alpha_{ij} = 2 \log \int_{-\infty}^{\infty} \psi_i(t) \psi_j(t) dt = 2 \log p_i + 2 \log p_j - \log \pi - \frac{1}{2}(\mu_i - \mu_j)^2$$

Die Zeilen- und Gesamtmittel der α_{ij} seien mit $\bar{\alpha}_i$ und $\bar{\alpha}$ bezeichnet, dann gilt:

$$\beta_{ij} = \alpha_{ij} - \bar{\alpha}_i - \bar{\alpha}_j + \bar{\alpha} = (\mu_i - \bar{\mu})(\mu_j - \bar{\mu})$$

wobei $\bar{\mu}$ das Mittel der μ_i ist.

m sei der dominante Eigenvektor der Matrix $B = (\beta_{ij})$, dann kann man f_k^m als Schätzwert für die Zeit verwenden, aus der der beobachtete Fund stammt. Bisher wurden zur chronologischen Reihung u.a. Verfahren verwendet, bei denen mit $\alpha_{ij} = f_i' \cdot f_j$ gerechnet wurde. f variiert bei diesem Ansatz auf einer hufeisenförmigen Kurve. Durch die Verwendung von $\alpha_{ij} = \log f_i' f_j$ wird die Variation auf einer Geraden erreicht.

H. IMMICH: Auswertung nichtdefinierter Grundgesamtheiten

In einer Doppelblindstudie sollte die Wirkung eines Verums und eines Placebos auf klimakterische Störungen untersucht werden. Streng zufällige Zuteilung, fixe Dosen und fixer Behandlungszeitraum waren festgelegt. Die Kriterien hingen weitgehend ab vom Ermessen des Prüfarztes. Der vorzeitige Versuchsabbruch wurde registriert. Die Studie erfaßte nur ambulante Patientinnen. Störfaktoren waren auf den Protokollen berücksichtigt.

Nach Abschluß des Versuchs lagen 517 auswertbare Protokolle vor. Alle Störfaktoren waren unter Verum und Placebo gleich häufig, der Altersdurchschnitt gleich. Der Versuch war vorzeitig abgebrochen bei 61 von 259 Pat. unter Verum, bei 74 von 258 Pat. unter Placebo. Ein Wirkungsunterschied ließ sich demnach nicht nachweisen.

Das Ergebnis ließ sich nicht verallgemeinern, weil der Auswahlmodus nicht festlag

- 1.) der Ärzte, die sich an der Studie beteiligt hatten
- 2.) der Patientinnen, welche die Ärzte als für die Studie geeignet ausgesucht hatten.

H.J.JESDINSKY: Hierarchische Pläne

Die Aufgabe, aus Beobachtungen

$$y_{i_1 \dots i_k} = \mu + U_{i_1}^{(1)} + U_{i_1 i_2}^{(2)} + \dots + U_{i_1 i_2 \dots i_k}^{(k)}$$

mit $E(U_{i_1}^{(1)}) = \dots = E(U_{i_1 \dots i_k}^{(k)}) = 0$, $E(U_{i_1 \dots i_m}^{(m)2}) = \sigma_m^2$; $m=1, \dots, k$

die σ_m^2 mit gleicher Genauigkeit zu schätzen, führt - in Abhängigkeit von den Verhältnissen $\sigma_1^2 := \sigma_2^2 : \dots : \sigma_k^2$ - i.a. zu unausgewogenen Plänen, d. h. Plänen,

in denen die Laufgrenzen der mit i_2, \dots, i_k indizierten Faktoren von den Werten der "höheren" Indizes abhängen. Mit Hilfe der Sätze über die Varianz und Kovarianzen quadratischer Formen werden für unausgewogene Pläne eines bestimmten Typs mit $k=4$

die Vorteile gegenüber den zugehörigen, ausgewogenen Plänen gezeigt. Zu einer vollständigeren Lösung der Aufgabe werden Simulationen vorgeschlagen. Diese hätten auch den Vorteil, die Stichprobenverteilungen der Schätzungen $\hat{\beta}_m^2$ genauer kennenzulernen.

J. KRAUTH: Bindungen bei Rangkorrelationen

Es wird ein lokal bester unverfälschter Rangtest zum Niveau α für ein einseitiges Korrelationsproblem bei Vorliegen diskreter Verteilungen angegeben. Der Test wird bedingt unter der jeweiligen Bindungskonstellation angewandt, wobei den kleinsten Beobachtungswerten die Rangzahl 1, den zweitkleinsten die Rangzahl 2 usw. zugeordnet wird. Liegen keine Bindungen vor, so ergibt sich gerade der Spearman-Test. Ebenso konvergieren die Alternativen, für die der Test lokal optimal ist, gegen entsprechende stetige Alternativen für den Spearman-Test, falls die zugrundeliegende Stützstellenmenge immer dichter wird. Es wird die (bedingte) asymptotische Verteilung unter der Hypothese angegeben und eine asymptotische Optimalitätseigenschaft diskutiert.

G.A.LIENERT: Hierarchische Konfigurationsfrequenzanalyse

Die KFA basiert auf dem Vergleich der beobachteten Frequenzen mit den erwarteten Frequenzen einer i -dimensionalen Kontingenztafel und deren Gesamt χ^2 -Wert $\chi^2 = \sum [(b-e)^2 / e]$. Die 2^i Konfigurationen von i Alternativ-Merkmalen ergeben $2^i \chi^2$ -Komponenten. Überschreitet eine χ^2 -Komponente eine vereinbarte Schranke, dann konstituiert die betreffende Konfiguration einen "Typus" oder ein "Syndrom". Um die Zahl der typenkonstituierenden Merkmale auf ein Minimum zu reduzieren, wird vorgeschlagen, zunächst die KFA mit allen i -Merkmalen, dann alle

$\binom{i}{i-1}$ KFA, alle $\binom{i}{i-2}$ KFA usw. bis alle $\binom{i}{2}$ KFA zu rechnen und sich für jene zu entscheiden, die den höchstsignifikanten χ^2 -Wert für $2^i - i - 1$ Freiheitsgrade liefert bzw. deren überfrequentierte Konfiguration zu interpretieren. Im Fall näherungsweise gleich großer Gesamt- χ^2 -Werte soll nach sachlogischen Aspekten entschieden werden.

O.LUDWIG: Mathematische und statistische Überlegungen zur zur BAK-Rückrechnung

In der Rechtsmedizin muß man häufig aus dem Ergebnis einer Blutalkoholbestimmung auf die Blutalkoholkonzentration zur Zeit eines oft mehrere Stunden vor der Blutentnahme liegenden Vorfalls schließen. Die Grundlage dafür bilden die auf den Untersuchungen von Widmark und anderen basierenden theoretischen Überlegungen über den zeitlichen Verlauf der Blutalkoholkonzentration. In der sogenannten Eliminationsphase wird dabei ein geradliniger Abfall der Konzentration vorausgesetzt. Es wird gezeigt, wie man die empirisch gefundene schwache Konzentrationsabhängigkeit theoretisch begründen kann. Statistische Untersuchungen betreffen individuelle Unterschiede der Abbaukurven, kurzfristige Schwankungen derselben und zufällige Bestimmungsungenauigkeiten bei der Blutprobe. Auf die Schwierigkeit, allen statistischen, medizinischen und juristischen Aspekten des Problems gerecht zu werden, wird eingegangen.

J.MICHAELIS: Probleme der automatischen EKG-Analyse

Nach einer Darstellung des elektrischen Erregungsablaufs am Herzen und des daraus abgeleiteten Modells des "Summen"- bzw. "Informationsvektors" wird das Prinzip der korrigierten orthogonalen Ableitungssysteme beschrieben.

Es wird der Ablauf der automatischen EKG-Analyse dargestellt, der sich in folgende Stufen gliedert:

Analogsignalerfassung und -speicherung

Analog-Digital-Umwandlung

Wellenidentifizierung

Vermessung der identifizierten EKG-Abschnitte und

Elimination von Störungen

Diagnostische Beurteilung

Auf Einzelprobleme der Wellenidentifizierung, Diagnostik und des Leistungsvergleichs verschiedener Programmsysteme wird näher eingegangen.

D. MORGENSTERN: Graphische Behandlung des 3-Alternativen-Problems.

Das Entscheidungsproblem, zwischen 3 bekannten, möglichen Dichten für die Beobachtung einer Zufallsvariablen zu wählen, führt sowohl nach "BAYES" wie nach NEYMAN-PEARSON (zwei Fehlerwahrscheinlichkeiten fest, die dritte minimieren) auf dieselbe Klasse:

Wenn $f_1/f_3 \leq k_1$ und $f_2/f_3 \leq k_2$, soll " f_3 " gewählt werden, andernfalls " f_1 " (wenn $f_1/f_2 > k_1/k_2$) oder " f_2 " (wenn $f_1/f_2 \leq k_1/k_2$).

Die Abbildung der Beobachtung durch vektorielle Addition von $\log f_i$ in den Richtungen 0° , 120° und 240° führt auf gerade, "sternförmige" Grenzen zwischen den Annahmebereichen. Die Verallgemeinerung bei detaillierten Fehlerarten α_{ij} (Fehlerwahrscheinlichkeiten, f_i zu wählen, wenn f_j vorliegt) ergibt gewisse Kurven als Trennlinien; weitere Hinweise auf Zulassung von Antworten vom Typ " f_1 oder f_2 " und geeignete Formulierung werden gegeben.

S. J. PÖPPEL: Grundlagen der automatischen Verarbeitung elektro-physiologischer Kurven (spez. EEG)

Zunächst werden einige elektrophysiologische Grundlagen (u.a.: Ruhepotential, Aktionspotential bei Nervenfasern und Muskelfasern dargestellt. Die an der Körperoberfläche kontinuierlich registrierten Biosignale müssen zur Verarbeitung auf einer EDV-Anlage diskretisiert werden (Analog-Digital-Umwandlung). Die zur AD-Wandlung erforderlichen Frequenzen sind in Abhängigkeit von den Frequenzen der registrierten Signale durch das Abtasttheorem bestimmt. Einige Typen digitaler Filter, die bei Störungen der Signale angewendet werden können, werden beschrieben. Für die Auswertung von Hirnstromkurven (EEG) sind bisher überwiegend spektralanalytische Verfahren eingesetzt worden. Ein nichtlinearer Ansatz zur Zeitreihenanalyse sowie ein statistisches Modell zur Erzeugung von Zeitreihen, die dem EEG ähnlich sind, werden vorgestellt.

G.REISSIG: Einige Ergebnisse von Krankenstandsuntersuchungen

Unter Krankenstand (KS) versteht man den prozentualen Anteil der durch Krankschreibung ausgefallenen Arbeitstage bezogen auf die Soll - Arbeitstage. In der DDR werden jährlich etwa 7 Millionen Arbeitsunfähigkeitsfälle im Rechenzentrum der Sozialversicherung registriert; bei einer Auswertung dieses Materials zeigte sich:

1. Die jahreszeitliche Verteilung weist bei nahezu allen Diagnosen einen Frühjahrs- und einen Herbstgipfel auf und verläuft entgegengesetzt zur Urlaubskurve.
2. Die Häufigkeit der Krankschreibungen nimmt mit zunehmendem Lebensalter ab, die Dauer dagegen zu, der Krankenstand bleibt unverändert.

Gleichartige Ergebnisse wurden bei der Auswertung von Personalunterlagen langjähriger Mitarbeiter eines Großbetriebs in der Bundesrepublik gefunden.

H.RIEDWYL: Anschauliches zu einigen nicht parametrischen Tests

Die Definition neuer nichtparametrischer Testgrößen schafft neue Verbindungen zwischen den bekannten Tests von Wilcoxon, Cramér-von Mises und Kolmogoroff-Smirnoff. An einfachen Beispielen sollen diese Beziehungen geometrisch veranschaulicht werden.

Bei dem 1. Test werden folgende Annahmen gemacht:

X und Y seien stetige Zufallsvariablen;

x_1, x_2, \dots, x_n seien n Realisierungen von X.

Die 0-Hypothese lautet: X und Y seien identisch verteilt (insbes. mit gleichen Mittelwerten). Es werden die Quantile $y_{i/n}$ für

$i = 1, 2, \dots, n-1$ bestimmt. Die $(n-1)$ Quantile und die n Realisierungen werden gemeinsam geordnet, die Ränge der $(n-1)$ Quantile zur Summe R_Q addiert. Dann ist die Testgröße

$$V = R_Q - n(n-1)$$

Für kleine n wurde eine Tabelle der Quantile von V, für große n eine Approximation an die Normalverteilung angegeben.

Für den 2. Test werden folgende Annahmen gemacht:

X und Y seien stetige Zufallsvariablen; x_1, x_2, \dots, x_{n_1} und y_1, y_2, \dots, y_{n_2} seien Realisierungen von X bzw. Y.

Es soll die 0-Hypothese geprüft werden, ob X und Y identisch verteilt sind (mit gleichen Mittelwerten). Zur Berechnung der Testgröße werden alle Beobachtungen unabhängig von der Stichprobenzugehörigkeit geordnet und mit den Rangzahlen $1, 2, \dots, n_1 + n_2$ versehen. Die Summe der Ränge der Stichprobe vom Umfang n_1 sei R_1 . Die Testgröße lautet:

$$W = 2 R_1 - n_1(n_1 + n_2 + 1)$$

Auch für W werden exakte Quantile und eine Approximation an die Normalverteilung angegeben.

CHR.L.RÜMKE und P.D.BEZEMER: Bestimmung von Normalwerten mittels Toleranzfaktoren eines ziemlich unbekanntens Typs

Im klinischen Labor möchte man über Normalwerte verfügen. Wenn man diese aus sehr vielen Beobachtungen an Normalen bestimmen kann ($N > 200$), kann man $2s$ -Grenzen berechnen. Bei kleinerem N kann $\bar{x} \pm 2s$ jedoch weniger als 95 % der Populationswerte umfassen! Für Beschreibungszwecke kann man dann die üblichen ("weiten") Toleranzgrenzen berechnen (z.B. wenigstens 95 % der Populationswerte zwischen $\bar{x} \pm ks$; Vertrauenswahrscheinlichkeit 95 %). In der klinischen Praxis könnte man besser "enge" Toleranzgrenzen berechnen, welche wenigstens (z.B.) $2 \frac{1}{2}$ oder 5 % der normalen Populationswerte in jeder der beiden Extremen mit einer bestimmten Wahrscheinlichkeit garantieren. Mit diesen engen "Alarmgrenzen" wird der Arzt in wenigstens 5 % der Fälle umsonst alarmiert werden, und zwar um so weniger, je mehr Beobachtungen an Normalen er gemacht hat. "Enge" "Factors for one-sided tolerance limits" findet man bei D.B.Owen (Scandia Corporation Monograph, March 1963). Verteilungsfreie enge und weite Grenzen lassen sich leicht errechnen, wenn die Beobachtungen keine Normalverteilung besitzen, wie im vorhergehenden unterstellt wurde. Immer bestimmt man Konfidenzbereiche für Perzentile.

B.SCHNEIDER: Mathematische Modelle für den Erregungsablauf am Herzen

Es wird das elektro-physiologische Konzept des Summenvektors dargestellt. Der Summenvektor als Funktion der Zeit ist Ausdruck des Erregungsablaufes und soll durch ein mathematisches Modell beschrieben werden. Da die Erregung sich nicht stetig, sondern sprunghaft von Zelle zu Zelle ausbreitet, kann man den Vorgang z.B. durch einen stochastischen Prozeß $N(t)$ beschreiben, wobei $N(t)$ die Zahl der erregten Zellen bis zur Zeit t angibt. $N(t)$ kann z.B. ein Poissonprozeß mit Mittelwert λt sein.

Statt der Zahlen $N(t)$ kann man auch die Folge der Zeitintervalle zwischen den Erregungssprüngen betrachten. Ein Modell für den Summenvektor erhält man, indem man den einzelnen Intervallen zufällige Vektorgrößen zuordnet. Die Repolarisation kann durch die Einführung eines 2. stochastischen Vektorprozesses oder mit Hilfe einer "impuls response" function dargestellt werden.

Diese Funktionen können für die einzelnen Intervalle verschieden sein. Ein Vergleich mit physiologischen Daten zeigt, daß über die Dauer eines Herzzyklus der Zahlprozeß $N(t)$ nicht stationär ist. Der stochastische Anteil des Zahlprozesses $N(t)$ und der Zufallsfolgen des Vektorprozesses macht sich nur dann bemerkbar, wenn pathologische Verhältnisse vorliegen, wie z.B. beim Kammerflimmern.

U.SCHULZ: Einige Anmerkungen zu einer Verallgemeinerung der multi-dimensionalen Skalierung

Die von J.Caroll und J. Chang, 1970 vorgeschlagene Berücksichtigung individueller Gewichte bei der Schätzung von Distanzen in einem Euklidischen Raum führt unter den üblichen Orthogonalitätsanforderungen an die Matrix der "Faktorladungen" zur folgenden notwendigen und hinreichenden Bedingung für die Existenz einer individuellen Gewichtungsstruktur: Die Eigenvektorsysteme der zu den Distanzmatrizen gehörenden Skalarproduktmatrizen müssen übereinstimmen. Ist diese Bedingung erfüllt, so kann man die Eigenvektoren aus der Summe der Skalarproduktmatrizen bestimmen und anschließend die Gewichte auf einfache Art berechnen.

H.THÖNI: Kriterien zur Wahl der "besten" Variablen-Transformation zur Auswertung von Versuchsdaten

Bei den sog. "linearisierenden" Transformationen (z.B. logarithmische Transformation bei exponentiellen oder Potenzfunktionen) ist der "Erfolg" der Transformation auch für den Laien leicht erkennbar. Schwieriger erscheint es, bei der Stabilisierung der Streuung dienenden Transformationen den "Gewinn" gegenüber einer Auswertung mit untransformierten Daten auszuweisen.

Die Mutmaßlichkeitsfunktion, wie sie von BOX und COX (1964) für ihre Transformation vorgeschlagen wurde, kann auch für andere Transformationen verwendet werden, um den Unterschied zwischen einer Streuungszersetzung vor und nach Datentransformation zu quantifizieren.

An Hand einiger Beispiele mit simulierten binomial verteilten Variablen wird gezeigt, daß dieser Gewinn z.B. bei kleinen Werten von p groß, bei mittleren Werten von p dagegen klein ist.

M.TÜCKE: Die Bewertung von Nebenwirkungen bei psychopharmakologischen Experimenten

Bei einer psychopharmakologischen Therapie ist es notwendig, die subjektive Beeinträchtigung der Patienten durch Nebenwirkungen zu quantifizieren. Am Beispiel eines Experimentes mit Noxiptilin, einem Antidepressivum, wird gezeigt, wie unter Verwendung psychologischer Skalierungsverfahren die möglichen Begleiterscheinungen gewichtet und zur Beurteilung des individuellen Therapieerfolgs mit herangezogen werden können. Als Maß für die globale Beurteilung der Begleiterscheinungen wird als "Beeinträchtigungsindex" die gewichtete Summe der Nebenwirkungen vorgeschlagen und auf ihre Brauchbarkeit geprüft.

N.VICTOR: "Perfekte und semiperfekte Kontingenztafeln"

Genügt eine dreidimensionale Kontingenztafel den Bedingungen

$$(1) \quad p_{ijk} = \frac{p_{ij} \cdot p_{i.k} \cdot p_{.jk}}{p_{i..} \cdot p_{.j.} \cdot p_{..k}} \quad \forall ijk$$

so bezeichnen wir sie als perfekt;

erfüllt sie die Bedingungen

$$\sum_i \frac{P_{ij} \cdot P_{i..k}}{P_{i..}} = P_{.j} \cdot P_{..k}, \quad \sum_j \frac{P_{ij} \cdot P_{.jk}}{P_{.j.}} = P_{i..} \cdot P_{..k} \quad \sum_k \frac{P_{i.k} \cdot P_{.j.k}}{P_{..k}} = P_{i..} \cdot P_{.j.}$$

so bezeichnen wir sie als semiperfekt. Natürlich ist jede perfekte Tafel auch semiperfekt. Da es bisher nicht gelungen ist, ML-Schätzungen unter (1) zu bestimmen, wird folgendes Vorgehen zum Testen der Perfektheit vorgeschlagen:

Die Semiperfektheit wird durch die Testgröße $\sum_{ijk} \frac{(\hat{m}_{ijk} - \hat{\hat{m}}_{ijk})^2}{\hat{m}_{ijk}}$ oder $\sum_{ijk} \frac{(\hat{m}_{ijk} - \hat{\hat{m}}_{ijk})^2}{\hat{m}_{ijk}}$ geprüft; dabei sind die ijk die ML-Schätzungen unter der BRK-Bedingung nach Roy-Kastenbaum und die

$$\hat{\hat{m}}_{ijk} = \frac{n_{ij} \cdot n_{i.k} \cdot n_{.jk}}{n_{i..} \cdot n_{.j.} \cdot n_{..k}}$$

Wird diese Hypothese abgelehnt, so kann gleichzeitig die Hypothese der Perfektheit abgelehnt werden, andernfalls entscheidet der anschließend durchzuführende Roy-Kastenbaum-Test.

P.O.WHITE: Ein Computermodell für das individuelle Lernen

Es wird ein Computerprogramm beschrieben, in dem aus psychologischen Untersuchungen für eine Versuchsperson ermittelte Parameter, die den Lernprozeß beeinflussen, als Modellparameter Simulationsuntersuchungen zugrunde gelegt werden. (z. B.: Verhaltensfestigkeit, Energie, Reaktionsschwelle, Reaktionshemmung, bedingte Hemmung, Vorgänge der Verstärkung, Auslöschung und Festigung). In den Simulationen wird die Auswirkung systematischer Veränderungen dieser Parameter untersucht. Es wird geprüft, wie gut die beobachteten Lernergebnisse mit den Ergebnissen übereinstimmen, die sich aufgrund der eingegebenen Parameter aus den Simulationsuntersuchungen ergeben.

F.X.WOHLZOGEN: Probleme der Anwendung sequentieller Testverfahren in der Medizinischen Forschung

Es werden Probleme erörtert, die sich bei der praktischen Anwendung sequentieller Methoden in der experimentellen Medizin ergeben. Im einzelnen werden folgende Verfahren vorgestellt, die zur Lösung einiger dieser Probleme beitragen können:

- 1) 3-Stufen-Stichprobenpläne anstelle des orthodoxen sequentiellen Paarvergleiches. Durch diese Pläne wird, bei praktisch unveränderter OC-Funktion, die Testzeit erheblich verkürzt.
- 2) Das Auftreten von "ties" und eine Möglichkeit ihrer Berücksichtigung in diesen Plänen wurde untersucht (Computer-simulation).
- 3) Ein neues Verfahren zur sequentiellen Parameterschätzung von Binomenalverteilungen.

CH.ZYWIETZ: Der Aufbau eines computerunterstützten EKG Auswertungssystems an der Medizinischen Hochschule Hannover

Es werden Ziele und Aufbau des aus Bundesmitteln aufgebauten Systems beschrieben. Zur Verfügung stehen neben den üblichen Registriergeräten Ampex-Analogbandgeräte sowie ein hybrides Rechnersystem CII 10020 - DO 720. Die Kennzeichnung aller Aufnahmen enthält neben der Patientenidentifikation Informationen über die Aufnahmeklinik, Datum usw. Die Auswertung beschränkt sich z. Zt. auf die Vermessung von Vektorkardiogrammen (Frank). Es wird mit einem modifizierten Pipberger Programm gearbeitet. Im Zusammenhang mit der Entwicklung von Diagnostikprogrammen soll eine über-regionale EKG Datenbank aufgebaut werden.

J.Michaelis (Mainz)