

MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

T a g u n g s b e r i c h t 8/1972

Medizinische Statistik

20.2. bis 26.2.1972

Die diesjährige Tagung über medizinische Statistik wurde von K. Überla und E. Walter geleitet. Neben zahlreichen Vorträgen über neuere Ergebnisse statistischer Methoden wurden biomathematische Modelle, Fragen der EDV in der Medizin und Probleme, die bei der Anwendung statistischer Methoden in der Praxis auftreten, behandelt.

Außerhalb der Vorträge wurde die Aufstellung des Lernzielkataloges für das in das Medizinstudium neu eingeführte Pflichtfach Biomathematik besprochen.

Herr Klotz (Mainz) führte Teile des vom ZDF produzierten Fernsehfilms "Integralrechnung" vor, woran sich lange Diskussionen über die Möglichkeiten, Vor- und Nachteile eines in ähnlicher Weise zu produzierenden Filmes "Statistische Methoden" anschlossen.

Diese Tagung bot wieder die Möglichkeit, daß die unterschiedlichen, auf dem Gebiet der medizinischen Statistik methodisch arbeitenden Wissenschaftler ihre Forschungsergebnisse vortragen und diskutieren konnten und förderte damit die Konsolidierung dieses Fachgebietes, wofür dem Mathematischen Forschungsinstitut gedankt sei.

Teilnehmer

J. Berger (Mainz)	H. Klinger (Düsseldorf)
E. Brunner (Aachen)	R. Klotz (Mainz)
W. Eberl (Düsseldorf)	J. Krauth (Düsseldorf)
W. van Eimeren (Ulm)	G. Lienert (Düsseldorf)
W. Gulbinat (Genf)	R.J. Lorenz (Tübingen)
F. Hampel (Zürich)	H. Maier (Berlin)
E. Hansert (München)	D. Morgenstern (Hannover)
W. Höbel (Gießen)	D. Plachky (Düsseldorf)
L. Horbach (Mainz)	S.J. Pöpl (München)
J. Hornung (Berlin)	A.J.Porth (Hannover)
H. Immich (Heidelberg)	H.-J. Reinhard (Hannover)

Chr. L. Rümke (Amsterdam)
H. Rundfeldt (Hannover)
L. Sachs (Kiel)
H.K. Selbmann (Ulm)
J. Siegerstetter (München)
E. Sonnemann (Heidelberg)
R. Stadler (Frankfurt)
W. Stucky (Darmstadt)
H. Thöni (Zürich)

K. Überla (Ulm)
N. Victor (München)
E. Walter (Freiburg)
H. Weiß (Berlin)
G. Wetter (Mainz)
F. Wingert (Hannover)
F. X. Wohlzogen (Wien)
G. Wolf (Ulm)
D. Wuppermann (Hannover)

VORTRAGSAUSZÜGE

D. MORGENSTERN : Zusammenfügung von Experimenten

Für das Alternativproblem bzw. Schätzproblem werden die Größen $I_{12}(f) = E_1 \left(\log \frac{f_1(X)}{f_2(X)} \right)$ bzw. $v(f) = E \left(\frac{\partial \log f(X,p)}{\partial p} \right)^2$ als Maß für die Einfachheit (= Leichtlösbarkeit) der betreffenden Experimente erläutert. Neben der Ungleichung $I_{12}(\varphi) \geq \text{Max} \{ I_{12}(f), I_{12}(g) \}$ für die gemeinsame Dichte φ von X und Y kann die vermutete Ungleichung $I_{12}(\varphi) \geq I_{12}(f) + I_{12}(g)$ durch Klassen von Gegenbeispielen widerlegt werden. Verallgemeinerungen auf Sequentialverfahren, mehrere Parameter. Allgemeine Empfehlung, Angaben über Auswahl der Population und den Anteil der Untersuchung bzw. seiner Auswertung zu machen.

H. KLINGER : Wahrscheinlichkeitsmodelle bei linearen Regressionsansätzen

Ist T eine Zufallsvariable, die nicht direkt beobachtet werden kann, so hat man oft die Möglichkeit, $X = T + \epsilon$ mit ϵ zufällig, $E(\epsilon) = 0$ zu messen. $\hat{T} = \beta (X - E(X)) + E(T)$ mit $\beta = \frac{\text{cov}(T, X)}{\text{var}(X)}$ kann dann als "Schätzfunktion für T" benutzt werden. Es wurde diskutiert, welche Konsequenzen sich aus zusätzlichen Forderungen wie $E(\epsilon^2 | T) \equiv \sigma_\epsilon^2$, ϵ und T sind unabhängig, $E(T - \hat{T} | X) \equiv 0$ und $E((T - \hat{T})^2 | X) \equiv \sigma_R^2$ für die Klasse der dann noch zugelassenen Wkt-Verteilungen ergeben. Es wurde gezeigt, daß nur normalverteilte Zufallsvariable allen Bedingungen genügen.

E. SONNEMANN : Einige Bemerkungen über gleichmäßig beste Tests

Es sei \mathcal{P} eine Menge von Wahrscheinlichkeitsmaßen über einem meßbaren Raum (X, \mathfrak{B}) , $P_0 \in \mathcal{P}$ und $T: (X, \mathfrak{B}) \rightarrow (R_1, \mathfrak{B}_1)$ eine meßbare Abbildung. Definiert man \mathcal{P}_{\leq} als Gesamtheit aller $P \in \mathcal{P}$ mit

$$\forall c \in R_1 : P_0 (T \leq c) \leq P (T \leq c)$$

und $\mathcal{P}_{>}$ als Gesamtheit aller $P \in \mathcal{P} - \{P_0\}$, für die eine monoton nichtfallende Funktion $H: R_1 \rightarrow [0, \infty]$ existiert mit

$$dP/dP_0 = H \cdot T [P + P_0],$$

so gibt es für alle $\alpha \in [0, 1]$ einen gleichmäßig besten (UMP-) Test zum Niveau α für \mathcal{P}_{\leq} gegen $\mathcal{P}_{>}$.

Als Beispiel für diese Aussage werden der χ^2 -Test für die Varianz und der Gauss-Symmetrie-Test betrachtet.

J. KRAUTH : Ein Zweistichproben-Skalen-Test

Es wird ein lokal bester Zweistichproben-Rangtest für Skalenunterschiede angegeben. Der Test setzt bezüglich Nullsymmetrische Dichten für beide Stichproben voraus und ist ein Analogon zum Savage - Test (Savage AMS 27 (1956)). Im Gegensatz zu diesem beruht er auf den Rängen R_{ij}^+ , $j=1, \dots, n_i$, $i=1, 2$ der Absolutwerte der Beobachtungswerte in der vereinigten Stichprobe und basiert auf der Statistik

$$T_N = \sum_{i=1}^{n_1} \sum_{j=N-R_{1i}^+ + 1}^N j^{-1}.$$

Unter der Nullhypothese ist die Statistik

$$T_N^* = (T_N - ET_N) / (\text{Var } T_N)^{1/2}$$

asymptotisch $N(0,1)$ - verteilt mit

$$ET_N = n_1, \quad \text{Var } T_N = \frac{n_1 n_2}{N(N-1)} \left(N - \sum_{i=1}^N i^{-1} \right).$$

Die lokale Optimalität bezieht sich auf die Doppel-Exponentialverteilung.

E. BRUNNER : Eine Beziehung zwischen dem Holm-Test und dem Kolmogorov-Smirnov-Test

Für die Wahrscheinlichkeitspapiertests von Holm werden kurz die Prüfgrößen und ihre Verteilungen aufgeführt und erläutert. Durch eine Substitution gelingt es, diese Prüfgrößenverteilungen aus den bereits bekannten Polynomen von Birnbaum-Tingey für den Kolmogorov-Smirnov-Test zu berechnen. Umgekehrt läßt sich die Verteilung der Prüfgröße D_n für den zweiseitigen Kolmogorov-Smirnov-Test aus einem von Holm angegebenen Gleichungssystem für Wahrscheinlichkeitspapiertests berechnen. Die OC des Holm-Tests läßt sich für bestimmte Klassen von Alternativen einfach berechnen und wird kurz mit der des Kolmogorov-Smirnov-Tests verglichen.

F. WINGERT : Beiträge zur Sensitivitätsanalyse statistischer Tests (Student, F-Test, Chiquadrat-Test, Bartlett-Test)

Es wurde die Sensitivität der Tests anhand der Literatur und eigener Untersuchungen gegenüber Abweichungen von der Normaltheorie analysiert.

Der zweiseitige t-Test ist sehr robust. Der einseitige t-Test ist dagegen stark von der Schiefe der Population abhängig. Vergleicht man mehrere Mittelwerte, dann wächst der Fehler gegen die Normaltheorie, wenn die Freiheitsgrade verschieden sind.

Der Varianzanalysen-Test ist ebenfalls sehr robust. Der Fehler gegen die Normaltheorie steigt mit der Variation der Varianzen und der Variation der Gruppenumfänge. Der Test besitzt in diesen Fällen einen Bias. Verschiedene Korrekturen werden in der Literatur vorgeschlagen (Modifikation der Freiheitsgrade, Anpassung von PEARSON-Verteilungen).

Tests über unabhängige Varianzen sind sehr sensitiv. Die Verteilung der Stichprobenvarianz besitzt auch für große Freiheitsgrade einen vom Exzeß der Population abhängigen Bias. Als Alternativen werden besonders für den BARTLETT-Test die Modifikation der Freiheitsgrade oder die Anpassung von Standardverteilungen vorgeschlagen.

H. WEISS : Auswirkung von Randomisierungsbeschränkungen auf die varianzanalytische Auswertung von Versuchen

Auf der Grundlage des von FISHER eingeführten Randomisierungsprinzips sind u.a. von NEYMAN et al. Randomisierungsmodellansätze für die Planung und Durchführung von Versuchen aufgestellt worden. Hinsichtlich der Auswertung haben PITMAN, WELCH und KEMPTHORNE gezeigt, daß häufig anstelle der exakten Randomisierungstests in guter Näherung die Tests der Varianzanalyse unter der "Normaltheorie" herangezogen werden können. Die Beurteilung darf sich jedoch lediglich auf die Untersuchungsfaktoreffekte erstrecken sowie auf die alternativen Zuordnungsmöglichkeiten bezogen werden, wenn der Versuchsumfang keine Zufallsstichprobe aus einer größeren Gesamtheit darstellt.

Eine analoge Beurteilung der Gruppierungseffekte ist nicht statthaft, wie wegen der Randomisierungsbeschränkungen nur unter der "Randomisierungstheorie" ersichtlich wird. Erfolgt jedoch die vereinfachende Auswertung unter der "Normaltheorie", so muß durch Einführung nichtschätzbarer Beschränkungsfehler diesem Sachverhalt Rechnung getragen werden.

F.X. WOHLZOGEN : Neue Entwicklungen auf dem Gebiet sequentieller Gruppierungstests in der klinischen Forschung

Praktischen Bedürfnissen der klinischen Forschung entsprechend wurden sequentielle Gruppierungstests entwickelt. Sie unterscheiden sich von den orthodoxen sequentiellen Paarvergleichen dadurch, daß die Paarbewertungen nicht einzeln, sondern in Gruppen vorgenommen werden.

Zwei Typen von Testplänen werden vorgestellt: solche, die maximal drei Gruppen und solche, die maximal vier Gruppen vorsehen (3- und 4-Stufenpläne). Wesentlich ist, daß auf jeder Stufe eine Entscheidung möglich ist, die zur Beendigung des Tests führt. Folgende Entscheidungen bezüglich der relativen

Überlegenheit zweier im Paarvergleich zu prüfenden Verfahren (A, B) wurden berücksichtigt:

1. 3-Entscheidungstests, mit den möglichen Entscheidungen:
 - A besser als B
 - B besser als A
 - kein Unterschied zwischen A und B
2. 2-Entscheidungstests, mit den möglichen Entscheidungen:
 - A besser als B
 - A nicht besser als B

Anhand der OC- und ASN-Funktionen dieser Pläne werden ihre statistischen Eigenschaften und ihr zweckmäßiger Einsatz in der Praxis der klinischen Forschung diskutiert.

D. PLACHKY : Wahrscheinlichkeiten großer Abweichungen

Ein Resultat über Wahrscheinlichkeiten großer Abweichungen für eine beliebige Folge von Zufallsgrößen wird vorgestellt und zur Bestimmung der Konvergenzgeschwindigkeit beim schwachen Gesetz der großen Zahlen benutzt. Dabei bewirkt das Randomisieren von Summen eine Veränderung der Konvergenzgeschwindigkeit in beiden Richtungen, während das Glätten mit Hilfe bedingter Erwartungswerte niemals zur Verschlechterung der Konvergenzgeschwindigkeit führt.

F. HAMPEL : Einige Beispiele für robuste Schätzungen

Zunächst werden Gründe für die Notwendigkeit robuster Schätzungen (in gewissen Situationen) angegeben. Dann wird eine kurze Liste von klassischen Beispielen für nicht-robuste Schätzungen aufgestellt, sowie von behelfsmäßigen, robusten Ersatzmethoden. Einige klassische und neuere robuste Schätzungen für Lokation und Dispersion von ungefähr normalverteilten Daten werden diskutiert (einschließlich der "median deviation"), mit kurzen Hinweisen auf ein robustes Analogon zum 1-Stichproben-t-Test und auf robuste Regression. Zum Abschluß wird eine kleine Tabelle mit Monte-Carlo-Varianzen einiger Schätzungen interpretiert.

H. THÖNI : Ein graphisches Verfahren zur raschen Beurteilung von Häufigkeitsverteilungen diskreter Zufallsvariabler

Ist y eine diskrete Zufallsvariable ($y=0,1,\dots$) und $f(y)$ die Häufigkeit, mit der y in einer Stichprobe vom Umfang $N = \sum f(y)$ beobachtet wird, so läßt die Funktion $g(y) = (y+1) f(y+1)/f(y)$ Rückschlüsse auf die Verteilungsfunktion von y zu. Bei Vorliegen einer Binomialverteilung, Poisson-Verteilung, negativen Binomial- oder logarithmischen Reihenverteilung müssen die Punkte $g(y)$, als Funktion von y aufgetragen, entlang einer Geraden liegen, die bei Vorliegen einer Binomialverteilung eine negative Steigung, bei negativer Binomialverteilung und logarithmischer Reihenverteilung eine positive Steigung aufweist und bei Poisson-Verteilung horizontal verläuft. Abweichungen vom geradlinigem Verlauf lassen Unregelmäßigkeiten in der Stichprobe erkennen. Aus den Regressionsformeln für die Berechnung der Wahrscheinlichkeitsfunktion lassen sich Parameter der Geraden $g(y)$ angeben.

E. HANSERT : Ein Ansatz zur Cluster-Analyse von Alternativmerkmalen

Gegeben sei eine $m \times n$ -Matrix aus Nullen und Einsen als Resultat der Beobachtung von m Alternativmerkmalen an n Objekten. Gefragt sei nach Teilmengen untereinander "ähnlicher" Zeilen (Merkmals-Cluster). Der Ansatz besteht darin, die Zeilenhäufigkeiten der "1" als gegeben zu denken und alle zufälligen Anordnungen der Einsen in das Rechteckschema unter dieser Vorgabe zu betrachten. Greift man eine bestimmte Teilmenge der Zeilen heraus und verwendet als "Ähnlichkeit" der Elemente dieser Teilmenge die Anzahl S der übereinstimmenden Spalten, so kann deren Wahrscheinlichkeitsverteilung unter der Zufallshypothese bestimmt werden. Hat man nun nach einem Suchverfahren ein Merkmals-Cluster gefunden, so daß der Wert von S für deren Elemente größer ist, als nach Zufall zu erwarten, so verwende man dieses Cluster in hypothetischer Weise (Überprüfung durch ein zweites, gezieltes Experiment).

J. PÖPPL : Spektraldarstellungen stationärer Zeitreihen durch Walshtransformationen

Es wird untersucht, wie man Walshfunktionen in diskreter Form zu Spektraldarstellungen stationärer Zeitreihen verwenden kann. Der Aufbau der Transformationsmatrix aus Kroneckerprodukten sowie eine weitergehende Faktorisierung nach GOOD ermöglicht einen Algorithmus, der im Gegensatz zur FFT (Fast-Fourier-Transform) nur $N \log N$ Additionen benötigt, um ein Array der Größe N zu transformieren. Es wird ein 'Walsh-Periodogramm' definiert, seine Entstehung und Berechnung wird aufgezeigt, ebenso der Begriff Frequenz in diesem Zusammenhang definiert. Schließlich werden Schätzverfahren angegeben.

G. LIENERT : χ^2 -Zerlegung mehrdimensionaler Vierfeldertafeln nach Lancaster

Bei einer Untersuchung von Aphasikern mit 5 psychologischen Tests wurde die resultierende pentavariante Korrelationstafel durch Mediandichotomisierung jedes einzelnen Tests in eine fünfdimensionale Vierfeldertafel transformiert. Die χ^2 -Zerlegung dieser Tafel nach H.O. Lancaster (1969) ergab neben Assoziationen erster Ordnung in den zweidimensionalen Randtafeln auch Assoziationen höherer Ordnung in den drei- und vierdimensionalen Randtafeln sowie in der fünfdimensionalen Gesamttafel, die nicht durch Assoziationen von jeweils niedrigerer Ordnung aufgeklärt werden konnten. Es wird daher empfohlen, bei Assoziationen höherer Ordnung auf eine Faktor- oder Clusteranalyse zu verzichten.

CHR.L. RÜMKE : Das Problem der Normalwerte im Blutaussstrich

Es wird gezeigt, daß die üblichen Normalwerte für die prozentuale Zusammensetzung des weißen Blutbildes zu klein sind, so daß zu oft vermutet wird, daß die prozentuale Zusammensetzung pathologisch ist. Deshalb soll bei den Ergebnissen

eines Blutausstriches immer die Anzahl der differenzierten Zellen angegeben werden.

L. SACHS : Stichprobenumfang und Power für den Vergleich zweier Prozentsätze

An dem in der Medizin häufigen Vergleich zweier Prozentsätze wird an einigen Beispielen gezeigt, daß es dann, wenn die Stichprobenumfänge klein sind, in bestimmten Fällen sinnvoll sein sollte, auf dem 10 %-Niveau zu prüfen, vorausgesetzt, die Power liegt bei oder oberhalb von 70 %.

H. IMMICH : Statistische Probleme bei abhängigen Variablen

An einem Beispiel wird gezeigt, daß die Korrelationen zwischen Einfluß- und Zielgröße bei medizinischen Daten zu Fehlinterpretationen führen können.

H. MAIER : Eine Faktorenanalyse an umfangreichem statistischen Material

Es wird eine Faktorenanalyse beschrieben, die für das Problem: "Quantitative Messung des Lebensstandards in sechs EWG-Ländern" durchgeführt wurde.

J. SIEGERSTETTER : Ein Verfahren zur Überprüfung von Indizes

Die Eigenschaften dreier aus Körpergewicht und Größe hergeleiteter Indizes zur Erkennung von Adipositas werden mit Hilfe von Fourierreihenentwicklungen kritisch betrachtet.

J. BERGER : Einfluß der Strukturierung einer Population auf die Infektionsausbreitung - Modelluntersuchungen -

Es wird ein Simulationsmodell vorgestellt, das eine be-

liebige Anordnung der Population in Untergruppen gestattet und mit dem Vaccinationspläne studiert werden können.

Innerhalb der ersten Tage nach Infektionsausbruch unterscheidet sich der Epidemieumfang einer strukturierten von einer unstrukturierten Population nicht. Bei einer reihenartigen Anordnung der Gruppen wird eine Verzögerung der Infektionsausbreitung festgestellt ; mit zunehmendem Abstand vom Ursprung der Infektionseinschleppung gleichen sich die Epidemieumfänge immer mehr an.

Nach Ausbruch einer Infektion und Vaccination der noch nicht Infizierten kommt es, unabhängig ob es sich um eine strukturierte oder unstrukturierte Population handelt, noch zu Neuinfektionen, ehe die Infektionsausbreitung sistiert.

G. WOLF : Modelluntersuchungen zum Verhalten von Lymphocytenanzahlen beim offenen ductus thoracicus

Mit bestimmten einfachen Modellannahmen kann das Verhalten der Lymphocytenanzahlen für verschiedene Versuchsanordnungen beschrieben werden. Dabei wird gezeigt, daß sich aus dem Modell Folgerungen ergeben, die sehr viel weiter gehen, als nach der einfachen Formulierung des Modells angenommen wird. Von Interesse sind dabei die Schlußfolgerungen, die in neuen Experimenten biologisch überprüft werden können.

D. WUPPERMANN : Probleme der Datenerfassung

Verschiedene Methoden zur Datenerfassung (Maschinenlochkarte, Port-a-Punch-Karte, Mark-Sensing-Karte und Markierungsbeleg) wurden im Labor und im Feldversuch experimentell miteinander verglichen. Geeignete Datenstruktur und -organisation vorausgesetzt, ist es empfehlenswert, vorzugsweise Port-a-Punch-Karten und Markierungsbelege zu verwenden, da diese Datenträger bei den Untersuchungen die geringste Fehleranfälligkeit zeigten.

G. WETTER : Datenstrukturen und Speicherstrukturen aus der Sicht des Statistikers

Ein Überblick über Speicherstrukturen wird anhand des Begriffes des Plexen von D.T. Ross (1961) gegeben, als Überblick über Datenstrukturen wird das graphentheoretische bzw. mengentheoretische (Relationen-) Modell vorgestellt (E.F. CODD 1970). Auf die Datenstrukturen von Algol 68 wird verwiesen.

H.K. SELBMANN : Probleme medizinischer Datenbanken

Am Beispiel einer allgemeinen Vorsorgeuntersuchung im Land Baden-Württemberg lassen sich einige der beim Aufbau dem Bedarf angepaßter Informationssysteme auftretenden Probleme demonstrieren. Anforderungen von Seiten der Benutzer waren in diesem Spezialfall: die Verarbeitung der großen Datenmenge (4,5 Millionen Zeichen), ein offener, und durch Fragetypen beschriebener Fragenkatalog (höchste Priorität: Kontingenztafelberechnung) und eine Dialogauswertung (Antwortzeiten < 10 sec). Diese Anforderungen lassen sich mit einer dreiteiligen Speicherstruktur optimal erfüllen, bestehend aus einem Merkmalskatalog, der den Benutzer den Zugang zu den Daten vermittelt, einer invertierten Datei, die pro Merkmalsausprägung eine Bitkette (pro Patient 1 bit) enthält und einer satzorientierten Datei, in der die nicht kategorisierbaren Daten gespeichert sind.

A.J. PORTH : Zur Datenverarbeitung im klinisch-chemischen Laboratorium

Anhand von drei on-line Datenverarbeitungssystemen im klinisch-chemischen Laboratorium wird die Problematik der computerunterstützten Datenerfassung dargestellt. Insbesondere wird die Schwierigkeit diskutiert, mit den verfügbaren Mitteln (hard-ware, soft-ware, "man"-ware), die anfallenden Datenmengen sinnvoll zu reduzieren, geschickt darzustellen, um einen besseren Informationsgehalt zu erzielen und eine Basis für die Datenverwertung zu schaffen.

I. VAN EIMEREN : Zur Qualitätskontrolle im klinischen Labor

Werden mehrere Merkmale im klinischen Labor in ihrer Qualität kontrolliert, so ist es sinnvoll, anstelle einer ganzen Batterie von Kontrollkarten, wie in der industriellen Qualitätskontrolle mit je einer Kontrollkarte den Mittelwertvektor und die Kovarianzenmatrix zu erfassen, die bei entsprechendem Zugang zu einem Großrechner mit mehrvariablen Methoden ausgewertet werden können.

H. RUNDFELDT : Kombinierter Einsatz von Videorecordern und Datenverarbeitungsanlagen für den Hochschulunterricht

Mit Hilfe des Videorecorders kann man den Hochschulunterricht didaktisch verbessern. Nach unseren Erfahrungen sollten jedoch nicht mehr als 35 % der Gesamtzeit reine Fernsehzeit sein. Daneben sind direkter Vortrag, Diskussionen und Übungen wichtig. Für die Übungen zur Biomathematik haben wir die Datenverarbeitungsanlage eingesetzt. Es werden Aufgaben gestellt, bei denen Zufallszahlen verwendet werden, die deshalb für jeden Studenten verschieden sind. Die Studenten geben ihre Ergebnisse über Mark-Sensing-Karten in die Datenverarbeitungsanlage. Dort werden sie geprüft und, falls falsch, wird ein Fehlerblatt ausgegeben.

G. KLOTZ : Studium im Medienverbund

Nach einem Überblick über Aufgaben, Absichten und Möglichkeiten wird die Testproduktion Mathematik: "Integralrechnung" dargestellt.

E. Walter (Freiburg)