

Mathematisches Forschungsinstitut

Oberwolfach

Tagungsbericht 5/68

Medizinische Statistik und Dokumentation

26. 2. bis 2. 3. 1968

Vom 26. Februar bis 2. März fand zum ersten Mal eine Tagung über medizinische Statistik in Oberwolfach statt. Die Leitung der Tagung hatte Professor Dr. P. IHM (Marburg). Es wurden zwei Gruppen von Vorträgen gehalten. Die erste Gruppe betraf statistische Methoden, die zweite, kleinere, behandelte sonstige Themen der Biomathematik.

Teilnehmer

| | |
|-----------------------------|---------------------------|
| Bambynek, G., Leverkusen | Jainz, M., Kiel |
| Bammert, J., Freiburg | Jesdinski, H., Freiburg |
| Beinhauer, R., Freiburg | Klinger, H., Düsseldorf |
| Bloedhorn, H., Freiburg | Koller, S., Mainz |
| Bühler, W., Heidelberg | Marksteiner, Wien |
| Christl, Hannover | Morgenstern, D., Freiburg |
| Dietz, K., z. Zt. Sheffield | Nash, S., Vancouver |
| Dudeck, J., Mainz | Neiß, A., Konstanz |
| Eberl, Düsseldorf | Parkel, F.T., Zürich |
| Gottschewski, J., Berlin | Pfander, R., Freiburg |
| Greiser, Hannover | Roßner, R., Freiburg |
| Ihm, P., Marburg | Sachs, Kiel |
| Immich, H., Heidelberg | Scheiber, Wien |

1093

14
10



| | |
|--------------------------|----------------------|
| Schneider, B., Hannover | Victor, N., Mainz |
| Schönemann, D., Hannover | Walter, E., Freiburg |
| Sinios, A., Hannover | Wetter, G., Mainz |
| Überla, K., Mainz | |

Vortragsauszüge

BAMBYNEK, G.: Darstellung einiger Überlagerungsschlüssel

Bei der Dokumentation ergeben sich Schwierigkeiten, wenn zur Beschreibung eines Gegenstandes eine im Vergleich zur Speicherkapazität große Anzahl möglicher Eigenschaften (charakterisiert durch vielstellige Codezahlen) verwendet wird. Eine Möglichkeit für die Dokumentation bieten dann Überlagerungsschlüssel. Die Codezahlen werden in Ziffernfolgen zerlegt und jede Ziffernfolge in das entsprechende Feld des Datenträgers abgelocht. Auf diese Weise können viele Codezahlen auf dem selben Datenträger überlagert werden. Bei der Abfrage werden außer den gewünschten Codezahlen noch weitere, sich zufällig aus den Ziffernfolgen ergebende Codezahlen selektiert. Die Abschätzung des Anteils dieser nicht-gewünschten Codezahlen und die Überlagerungsfähigkeit des Schlüssels erfolgt über die "Ballast"-Wahrscheinlichkeit, die für gleichverteilte Belegung des Datenträgers hergeleitet wurde. Es wurden Überlagerungsschlüssel mit 2 und 4 Feldern (zu je 100 Lochstellen) bei Überlappung und bei Nichtüberlappung der Codezahlen miteinander verglichen.

BAMMERT, J.: Modelle zur Verschaltung visueller Neurone

Laterale Rückinhibition verwendeten Hartline (1956), Reichardt (1962) u. a. zur Deutung visueller Kontrastphänomene und schneller Nachbilder. Dies befriedigt in einfachen Fällen, wie bei Limulus.

Ein modifiziertes Modell, bei dem an gewissen Schaltinstanzen hemmende und erregende Eingänge gegeneinander vertauscht sind, hat Eigenschaften, die dem Vorkommen von on- bzw. off-center-Neuronen in der Wirbeltierretina entsprechen. In qualitativer Übereinstimmung mit den Experimenten sind auch am Modell die Spontanaktivitäten der off-Neurone größer, mit sinkender Erregungsintensität (bei zur Peripherie des rezeptiven Feldes hin wandernden Reizen) wächst die Latenzzeit, die Einschwingvorgänge (Nacherregungen) sind mit höherer Reizintensität oder niedriger Spontanaktivität stärker gedämpft, und off-Neurone zeigen langdauernde totale Hemmphasen. Jedoch ist die Abhängigkeit der Schwingungsperiodik von Reizdauer und -intensität (Baumgartner) im Modell nicht erfaßt.

BEINHAUER, R.: Nomogramm zur Bestimmung des Stichprobenumfangs

In einem Bestand von N Tieren mögen sich m kranke befinden. Untersucht man eine Stichprobe von n Tieren, so ist die Wahrscheinlichkeit, kein krankes zu finden

$$P_0 = \frac{(N-n)! (M-m)!}{(M-m-n)! N!}$$

Gibt man sich P_0 und den (vermuteten) Krankenstand $m/N = p$ vor, so ist also der Stichprobenumfang aus der obigen Gleichung zu bestimmen.

Mit Hilfe der Stirlingschen Formel wird P_0 in einfacher Weise durch p und den Anteil $q = n/N$ der untersuchten Tiere ausgedrückt:

$$\log P_0 \approx \frac{1}{2} A(p, q) + N B(p, q).$$

A und B sind symmetrische Funktionen in p und q . Für die Näherung

$$\log P_0 \approx N B(p, q)$$

wird ein Nomogramm, das den Zusammenhang zwischen N , p , q und P_0 ausdrückt.

BLOEDHORN, H.: Über die Varianz des Pipettenfehlers beim Verdünnen von Virussuspensionen

Beim Ausverdünnen von Virussuspensionen werden bei jedem Verdünnungsvorgang u ml Virussuspension mit w ml Verdünnungsflüssigkeit gemischt. Ist die Ausgangsverdünnung gleich V_o mit Varianz $\sigma_{V_o}^2$, so ist die Verdünnung nach n Verdünnungsschritten gleich

$$V_o \prod_{i=1}^n \frac{u_i}{u_i + w_i}$$

Faßt man u_i, w_i als Realisationen von $2n$ unabhängigen, normalverteilten Zufallsvariablen $U_i, W_i, (i=1, \dots, n)$ auf, so findet man nach dem Fehlerfortpflanzungsgesetz für den Logarithmus der n -ten Verdünnungsstufe Mittelwert und Varianz zu

$$E \ln(V_o \prod_{i=1}^n \frac{U_i}{U_i + W_i}) = \ln(V_o (\frac{1}{k})^n) + n(\frac{1}{2} - 1)v_u^2 + n(1 - \frac{1}{k})^2 v_w^2$$

$$\begin{aligned} \text{Var} \ln(V_o \prod_{i=1}^n \frac{U_i}{U_i + W_i}) &= n(1 - \frac{1}{k})^2 (v_u^2 + v_w^2) + \{ \frac{n}{2} [(\frac{1}{2} - 1)^2 v_u^4 + (1 - \frac{1}{k})^4 v_w^4] - \\ &\quad - \frac{1}{2} (\frac{n}{2}) (\frac{1}{2} - 1) v_u^4 - \frac{n^2}{2} (\frac{1}{2} - 1) (1 - \frac{1}{k})^2 v_u^2 v_w^2 - \frac{1}{2} (\frac{n}{2}) \cdot \\ &\quad \cdot (1 - \frac{1}{k})^4 v_w^4 \} + \sigma_{V_o}^2 \cdot \frac{1}{V_o^2} \end{aligned}$$

Hierbei war $E(U_i) = \mu_u, E(W_i) = \mu_w, i=1, \dots, n$, ferner

$$k = \frac{\mu_u + \mu_w}{\mu_u} \text{ und } v_u = \frac{\sigma_u}{\mu_u}, \text{ sowie } v_w = \frac{\sigma_w}{\mu_w} \text{ mit}$$

$$\text{Var}(U_i) = \sigma_u, \text{ Var}(W_i) = \sigma_w, i=1, \dots, n.$$

Bei der Varianz werden allerdings die Glieder in $\{ \dots \}$ im allgemeinen so klein sein, daß man sie vernachlässigen darf. In vielen Fällen wird außerdem $\sigma_{V_o}^2 = 0$ sein.

BÜHLER, W.: Stochastische Modelle der Carcinogenese

Diese Modelle gehen alle von der Mutationstheorie der Krebsentstehung aus, d.h. von der Vorstellung, gewisse Zellen erfahren Veränderungen, die bei Zellteilung auf die Tochterzellen übergehen. Solche Veränderungen können spontan erfolgen oder unter dem Einfluß eines Carcinogens. Die Ein-Treffer-Modelle von Arley-Iversen und von Neyman konnten Versuche mit fraktionierten Dosen eines Carcinogens nicht erklären, daher wurde später von Neyman und Scott ein Zwei-Treffer-Modell konstruiert, bei dem die Umwandlungsraten dosisabhängig sind und sich "graue" und "schwarze" Zellen nach Geburts- und Todesprozessen vermehren.

Es wurde am Beispiel des Modells von Stocks gezeigt, daß es oft durch Spielen mit den Parametern des Modells möglich ist, auf mehr als eine Weise eine Anpassung an vorhandene Daten zu finden.

Das Modell der normalen Mäusehaut, das von Bjerknes und Iversen auf dem Computer simuliert wurde, wurde als interessanter Ansatz erwähnt.

DIETZ, K.: Ausdünnung von Erneuerungsprozessen

Den durch einen Erneuerungsprozeß bestimmten Punkten werden nach einer bestimmten Regel die Marken 0 und 1 ("auslöschen" oder "stehenbleiben") zugeordnet. Aus vorgegebenen Eigenschaften des Erneuerungsprozesses und der Markierungsregel sollen die Eigenschaften des Prozesses bestimmt werden, der durch die Punkte mit der gleichen Marke (z.B. 1) gebildet wird. Es wird ein Überblick über bisher in der Literatur betrachtete Markierungsregeln gegeben. Neue Resultate werden für die beiden folgenden Regeln mitgeteilt:

1. Die Folge der Marken bildet eine Markovkette.
2. Die Wahrscheinlichkeit des Auslöschens ist eine exponentiell fal-

lende Funktion vom Abstand zum jeweils letzten stehengebliebenen Punkt.

Die Eigenschaften der dabei auftretenden multimodalen Verteilungen werden untersucht.

GOTTSCHESKI, J.: Ein einfaches Stoffwechselmodell

Es wird ein Modell für die Ausscheidung einer Substanz aus dem Körper angegeben. Der Körper wird durch 2 Kompartimente dargestellt, in denen die Substanz jeweils gleichmäßig verteilt ist. Der Austausch zwischen den beiden Kompartimenten wird in beiden Richtungen zugelassen, die Ausscheidung aus dem Gesamtsystem ist nur aus einem der beiden Kompartimente möglich. Die in den Kompartimenten vorhandenen Mengen der Substanz in Abhängigkeit von der Zeit werden durch ein lineares Differentialgleichungssystem zweiter Ordnung beschrieben.

Unter der Annahme, daß Beobachtungen lediglich an einem der beiden Kompartimente vorgenommen werden können, werden verschiedene Abschätzungen für die das System beschreibenden Parameter vorgenommen.

IHM, P.: Subjektivistische Interpretation von Konfidenzmengen im Falle lokal kompakter topologischer Gruppen

Bei Wahrscheinlichkeitsverteilungen auf lokal kompakten topologischen Gruppen und Dichten $f(\vartheta^{-1}x)$ bezüglich des linksinvarianten Haarschen Maßes μ sind bestimmte Konfidenzmengen γ gleichzeitig " γ Quantile" der subjektiven Wahrscheinlichkeitsverteilung, wenn als a-priori- (improper) Wahrscheinlichkeitsverteilung das entsprechende rechtsinvariante Haarsche Maß angenommen wird.

Eine Reihe von Widersprüchen bezüglich der Anwendung des Konfidenz-

schlusses lassen sich durch diese Interpretation klären.

IMMICH, H.: Einfluß der Ausgangsdaten auf die Ergebnisse statistischer Tests

Bei sogenannten verbundenen Stichproben ist es üblich, die Differenzen zwischen beispielsweise einem Meßwert vor und einem Meßwert nach einer Behandlung als Variable einzuführen, die den Behandlungseffekt messen. Diese Methode kann jedoch nur dann benutzt werden, wenn die Differenzen von den Meßwerten vor der Behandlung unabhängig sind. An einem Beispiel wird gezeigt, wie sich die Verwendung "abhängiger" Differenzen bei einer einfachen Varianzanalyse auswirkt.

Bei Tierexperimenten verwendet man häufig Quotienten oder Indizes, ohne vorher zu prüfen, ob die Meßwerte im Zähler von den Meßwerten im Nenner unabhängig sind. Am Beispiel einer Varianzanalyse mit dreifachem Eingang wird demonstriert, wie sich die Verwendung derartiger Quotienten auf das Ergebnis auswirken kann.

JESDINSKY, H.: Versuchspläne zur Schätzung der Nichtadditivität von Direkt- und Nachwirkungen

Bei medizinischen Experimenten werden oft denselben Individuen nacheinander verschiedene Behandlungen zugeteilt. Dabei kann es vorkommen, daß der Effekt einer Behandlung noch nicht abgeklungen ist, wenn die nächste Behandlung erfolgt. Es gibt Fragestellungen, bei denen nicht nur die Schätzung der Direkt- und Nachwirkungen, sondern auch der Nichtadditivität beider Effekte wichtig ist.

Seien die Intervalle zwischen zwei Untersuchungen gleich groß und r die Anzahl solcher Intervalle, in denen Nachwirkungen einer Behandlung möglich sind. Dann gewinnt man Versuchspläne zur Schät-

zung der Nichtadditivität von Direkt- und Nachwirkungen durch Lösung folgender Aufgabe:

Es sollen n verschiedene Elemente, die je in einer Anzahl von n^{r-1} zur Verfügung stehen, so auf n^r Plätze verteilt werden, daß vor jedem Element alle $(r-1)$ -Permutationen der n (nicht notwendig verschiedenen) Elemente vorkommen. Dabei soll dem n^r -ten Platz der erste Platz folgen, also mit den Platznummern $\text{mod } n^r$ gerechnet werden.

Die Existenz solcher Pläne für $r=2$ wird bewiesen für beliebige n . Mit einer Anordnung hat man für $n > 2$ sogleich eine Klasse von $2n^2 n!$ Anordnungen gefunden. Eine vollständige Aufzählung aller möglichen Anordnungen kann noch nicht angegeben werden.

KLINGER, H.: Regressionsansätze mit zufälligen Parametern

Bei Experimenten an biologischen Objekten erweist sich der Ansatz einer für alle Objekte einheitlichen Regressionsgeraden oft als unzureichend. Entspricht die Zufallsvariable Y_{ij} der j -ten Beobachtung ($1 \leq j \leq r$) am i -ten Objekt ($1 \leq i \leq n$), so können "individuelle" Regressionsgeraden durch den Ansatz $Y_{ij} = U_i + B_i + B_i x_{ij} + e_{ij}$ in adäquater Weise berücksichtigt werden. Sind die x_{ij} manipulierbar, so daß

$$\sum_{j=1}^r x_{ij} = \bar{x} \quad \text{und} \quad \sum_{j=1}^r (x_{ij} - \bar{x})^2 = S_{xx}$$

für alle Versuchsobjekte gleich sind, so ist unter der Annahme, daß die Größen U_i , B_i , e_{ij} normalverteilte und abgesehen von den Paaren U_i , B_i unabhängige Zufallsvariable mit den Erwartungswerten U_i bzw. B_i bzw. 0 u. den Varianzen σ_U^2 bzw. σ_B^2 bzw. σ_e^2 und der Kovarianz σ_{UB}^2 sind, die Konstruktion von exakten Tests und Konfidenzbereichen bezüglich der Parameter μ, β mit Hilfe von Hotellings T^2 -Verteilung möglich. Auf Verallgemeinerungen hinsichtlich der ent-

sprechenden Zwei- und Mehrstichprobenprobleme sowie auf Ansätze mit komplizierteren Regressionsgleichungen wurde hingewiesen.

KOLLER, S.: Anwendung nichtrepräsentativer Untersuchungsreihen bei ätiologischen Problemen

Verallgemeinernde Schlüsse aus Beobachtungsreihen beruhen meist auf der Annahme, daß die Reihe eine Stichprobe aus einer Grundgesamtheit ist. Die meisten Medizinischen Reihen werden aber nicht als Stichproben gewonnen, sondern sind in ihrer Struktur vorgegebene Gruppen. Eine Verallgemeinerung kann auf diejenige fiktive Grundgesamtheit erfolgen, die sich ergäbe, wenn man unter denselben Bedingungen sehr große Patientenzahlen gleicher Art untersuchen würde. - Bei ätiologischen Untersuchungen geht man retrospektiv oder prospektiv vor. Zentralpunkt der Materialgewinnung ist stets die Gegenüberstellung "vergleichbarer" Reihen, was durch Strukturgleichheit, Beobachtungsgleichheit und Repräsentationsgleichheit definiert wird, es kann im extremen Fall durch matched pairs erreicht werden. Vergleichbarkeit kann nur durch Einengung erzielt werden, damit wird die induktive Basis für Verallgemeinerungen jedoch reduziert. Zur Verbesserung der Verallgemeinerungsfähigkeit müssen die Einengungen schrittweise überprüft und ggf. aufgehoben werden. Dies erfolgt entweder durch Gegenüberstellung von Untergruppen des Materials oder durch neue Reihen an anderen Orten, durch andere Untersucher, mit anderen Patientengruppen und möglichst auch anderem Problemansatz.

In der Medizin ist anstelle des einfachen, verallgemeinernden Rückschlusses von der Stichprobe auf die Grundgesamtheit ein Prozeß anzusetzen, der schrittweise zwischen Vergleichbarkeit und Verallgemeinerung pendelt.

MORGENSTERN, D.: Doppelte Korrelationsanalyse

Zur mathematischen Behandlung einer verbreiteten Fragestellung, Kovarianzen zu bestimmen, wenn die Beobachtungsreihen selbst in unbekannter Weise korreliert sind, werden zwei Verfahren angegeben. Das eine ist ein Iterationsverfahren, welches auf in Spezialfällen plausiblen Formeln beruht, während das andere ein mathematisches Modell mit Normalverteilung benutzt.

Die Fragestellung und Methoden sind nicht nur selbst wichtig, sondern zugleich eine Vorstufe für ein zweifaches Klassifikationsproblem, das ebenfalls in den verschiedenen Anwendungsbereichen häufig auftritt.

ROSSNER, R. und E. WALTER: Schätzung von Parametern im verteilungsunabhängigen Falle

Jedem Test zur Prüfung eines Parameters η auf Grund einer Stichprobe r mit einem Prüfmaß $T_{\eta}(r)$ ist ein Konfidenzintervall und ein mediantreuer Schätzwert für η zugeordnet, wenn $T_{\eta}(r)$ in η monoton ist. In dieser Weise werden Schätzwerte und Konfidenzintervalle, die folgenden verteilungsunabhängigen Problemen entsprechen, betrachtet:

Prüfung, ob eine Verteilung symmetrisch bezüglich eines Punktes η ist.

Prüfung, ob sich zwei Verteilungen um eine Differenz η in der Lokalisation oder um einen Faktor η in der Skala unterscheiden (d.h. ob die Verteilungen der Zufallsvariablen X und $Z = Y - \eta$ bzw. X und $Z = \eta(Y - y_0) + x_0$ gleich sind).

Prüfung des Faktors η in einer allgemeinen Regressionsbeziehung (d.h. ob X und $Z = Y - \eta f(X)$ unkorreliert sind).

Bedingungen für die Monotonie von $T_{\eta}(r)$ und Verfahren der graphi-

schen Konstruktion der Schätzwerte und Konfidenzintervalle werden diskutiert.

SCHNEIDER, B.: Das kybernetische Prinzip in der Biometrie

Das Referat soll zeigen, daß die Kybernetik eine formale Theorie der biometrischen Modelle und die biometrischen Modelle Interpretationen des abstrakten, kybernetischen Systems darstellen. Als das Prinzip der kybernetischen Methode wird die kanonische Analyse herausgestellt, bei der aus den Input- und Outputgrößen des Systems geeignete, charakteristische Größen berechnet werden. Aus diesen Größen kann in einfacher Form das gesamte Input- und Outputverhalten des Systems dargestellt werden (Prediction-Problem). Außerdem drücken diese Größen besonders wichtige, die Gesetzmäßigkeit des Systems kennzeichnende Eigenschaften (kanonische Eigenschaften) aus. Es wird gezeigt, daß dieses kybernetische Prinzip bei den wichtigsten biometrischen Methoden und Fragestellungen eine entscheidende Rolle spielt.

ÜBERLA, K.: Modelluntersuchungen zur Aussagegenauigkeit der Faktorenanalyse

Das übliche rechnerische Vorgehen der Faktorenanalyse wird dargestellt. Ausgehend von der Unmöglichkeit, die Genauigkeit dieses Verfahrens allgemein anzugeben, wird ein Simulationsprogramm beschrieben. Dabei wird für eine Grundgesamtheit ein Faktorenmuster A festgelegt, eine davon zufällig abweichende Stichprobe ermittelt und deren Faktorenwerte P als normalverteilte Zufallsgrößen vorgegeben. Nach der Bestimmung der Datenmatrix wird eine Faktorenanalyse durchgeführt und die Faktorenwerte geschätzt. $\gamma_{p\hat{p}}$ ist das Gütekriterium. Es werden jeweils 50 Stichproben gezogen und die

1
2
3
4



Grundgesamtheit wird systematisch variiert. Die Ergebnisse zeigen, daß die Genauigkeit der Faktorenanalyse in Abhängigkeit von Fallzahl, Variablenzahl und Höhe der Korrelationen wie zu erwarten variiert. Die Genauigkeit ist bei vorliegender klarer Struktur hoch und unterscheidet sich nur geringfügig von der der multiplen Korrelationsrechnung, bei der man die Zielgröße kennt, während man sie in der Faktorenanalyse schätzen muß.

VICTOR, N.: "Wechselwirkungen" in mehrdimensionalen Kontingenztafeln

Die wichtigsten Unabhängigkeitshypothesen in mehrdimensionalen Kontingenztafeln und ihre Zusammenhänge wurden angegeben. Die beiden verschiedenen auf Bartlett und Lancaster zurückgehenden Definitionen des Begriffs "zweifache Wechselwirkung in der 3-dimensionalen Kontingenztafel" sowie die dazugehörigen Tests auf Verschwinden dieser Wechselwirkungen wurden gegenübergestellt. Als Anwendung wurden beide Verfahren zur Untersuchung der Zusammenhänge zwischen Fettleber, Lungenfettembolie und Herzinsuffizienz herangezogen.

J. Bammert, Freiburg

