

T a g u n g s b e r i c h t 9/1975

Medizinische Statistik

24.2. bis 28.2.1975

Die 8. Tagung über Medizinische Statistik wurde von H. Klinger und R. Repges geleitet. Die Vorträge hatten größtenteils methodischen Inhalt. Dabei stand die Anwendbarkeit dieser Methoden in der Medizin stets im Vordergrund. Andere Vorträge brachten interessante Anwendungen in Medizin und Biologie. Besonders fruchtbare Diskussionen lieferten die Beiträge, die sich kritisch mit angewendeten Methoden auseinandersetzten, bzw. die Fälle aus der Praxis vorstellten, zu denen es (noch) keine Lösungen gibt.

An zwei Abenden wurde in Kurzvorträgen und zahlreichen Diskussionsbeiträgen die Situation der Ausbildung in "Biomathematik" im ersten klinischen Teil des Medizinstudiums an den verschiedenen Universitäten vorgestellt. Dabei zeigte sich erneut die Diskrepanz zwischen einheitlicher Prüfung und nicht einheitlicher Ausbildung. Die Frage, wie verbindlich und umfassend der Stoff des Lernzielkatalogs von den Hochschullehrern gebracht werden sollte, konnte während der Tagung nicht ausdiskutiert werden. Dagegen lag ein großer Gewinn der Tagung im Austausch der Erfahrungen auf didaktischem Gebiet.

Das während der Tagung anhaltend schöne Wetter erlaubte die Durchführung mancher Diskussion in kleinen Gruppen während der Spaziergänge. Die traditionelle Wanderung zum Brandenkopf fand in einer neuen Variante statt: von den Nillhöfen bei Haslach ausgehend.

Teilnehmer

K. Abt, Frankfurt

J. Bammert, Freiburg

M.P. Bauer, Bonn

J. Berger, Mainz

P.D. Bezemer, Amsterdam (NL)
E. Brunner, Aachen
G. Deichsel, Tübingen
U. Feldmann, Hannover
J.D.F. Habbema, Leiden (NL)
J. Hermans, Leiden (NL)
R. Hilgers, Göttingen
H. Hirche, Essen
G. Hommel, Erlangen
J. Hornung, Berlin
H. Immich, Heidelberg
H. Klinger, Düsseldorf
J. Krauth, Düsseldorf
U. Kuß, Mainz
D. Morgenstern, Hannover
B. Nguyen, Hannover
H. Nowak, Aachen
G. Oberhoffer, Bonn
R. Repges, Aachen
H. Riedwyl, Bern (CH)
C.L. Rümke, Amsterdam (NL)
H. Rundfeldt, Hannover
L. Sachs, Kiel
W. Schaafsma, Groningen (NL)
S. Schach, Dortmund
T. Schäfer, Dortmund
W. Stucky, Darmstadt
H. Thöni, Stuttgart
N. Victor, Giessen
E. Walter, Freiburg
S. Wellek, Mainz
H. Witting, Freiburg
F.X. Wohlzogen, Wien (A)
R. Zentgraf, Giessen

VORTRAGSAUSZÜGE

K. ABT: Zur Frage der Homogenität der Kovarianzmatrizen im Parallelgruppenversuch

Im linearen Modell des Parallelgruppenversuchs ist

$y_{ijv} = \mu_{ij} + \epsilon_{ijv}$ ($i = 1, \dots, k$ Behandlungen; $j = 1, \dots, p$ Zeitpunkte; $v = 1, \dots, n_i$ Wiederholungen an verschiedenen Probanden). Statt einer

zweifaktoriellen univariaten Auswertung läßt sich eine einfaktorielle p -variante Auswertung durchführen. Dabei ist die Nullhypothese

$H_0 : \underline{\mu}_1 = \dots = \underline{\mu}_k$ mit $\underline{\mu}_i = (\mu_{i1}, \dots, \mu_{ip})'$. Durch eine geeignete

Zerlegung $\epsilon_{ijv} = \beta_{ijv} + \delta_{ijv}$ mit einer Patientenkomponente

$\beta_{ijv} \sim N(0, \underline{\Sigma}_\beta^{(i)})$ und einem Rest $\delta_{ijv} \sim N(0, \sigma_0^2 \cdot \underline{I})$ gelingt es,

die Korrelation über die Zeit zu trennen. Mit einem rekursiven Ansatz

$\mu_{ij} + \beta_{ijv} = \alpha_{jj'} \cdot (\mu_{ij'} + \beta_{ij'v}) + \tau_{ij}$, $j = j' + 1$ lassen sich

Varianzen auf den Anfangs-Zeitpunkt 1 zurückrechnen. Außerdem werden

die Varianzen für alle k Gruppen als gleich angenommen (Randomisierung).

Die Homogenität der Kovarianzmatrizen ergibt sich somit direkt

aus dem Modellansatz.

J. BMMERT: Ein Modellbeispiel zur Evolution von Langtag-Kurztags-Verhalten

Viele Insekten schlüpfen nach der letzten Larvenhäutung oder gehen

in Diapause (Winterruhe) - je nachdem, ob die Tageslänge größer oder kleiner als ein genotypisch festgelegter Schwellenwert τ ist.

Unter einer vom Langtag- oder Kurztags-typ abhängigen Verteilungsannahme für τ ($P(\text{Diapause bei Typ } i) = p_i$, $i = 1, 2$), unter

der Voraussetzung der Dominanz des Langtagtypes und unter der Modell-

annahme einer jahreszeitlichen Überlebensrate $s(x)$, die die Umwelt-

bedingungen für beide Typen wiedergibt, läßt sich die Populations-

größe beider Typen berechnen. Man erhält, daß unter anhaltenden

extremen Umweltbedingungen jeweils ein Typ aussterben muß. Der beobachtete balancierte Polymorphismus läßt sich mit diesem Modell nicht erklären; er ist offensichtlich auf eine Arten-ausgleichende Migration aus Nachbarregionen zurückzuführen.

M.P. BAUR : Die Beziehung zwischen genotypischen und phänotypischen Vaterschaftsplaussibilitäten

Wendet man bei gegebener phänotypischer Kind-Mutter-Putativvater Konstellation die Formel von OKAJIMA auf die genotypisch möglichen Fälle an, so erhält man für jede dieser Möglichkeiten einen genotypischen W-Wert. Die diskrete Verteilungsfunktion, der Erwartungswert und die Varianz dieser genotypischen W-Werte lassen sich bestimmen, und es zeigt sich, daß der phänotypische W-Wert gleich dem Erwartungswert der genotypischen W-Werte ist.

J. BERGER : Schwellenwert für die Infektionsausbreitung in sich nicht homogen mischenden Populationen

Die Änderung der Infizierten in einer sich homogen mischenden Population läßt sich durch die Differentialgleichung $\dot{z}(t) = r(S_0 - z(t)) \cdot (z(t-1) - z(t-d))$ beschreiben. Dabei sind r die Infektionsrate, S_0 der Anteil der Empfänglichen zur Zeit $t = 0$, 1 die Latenzzeit und d die Dauer der infektiösen Periode. Für das Entstehen einer Epidemie muß $S_0 > 1/rd$ sein. Eine räumliche Anordnung und Aufteilung der Population in Untergruppen muß sich in einer Änderung von r widerspiegeln. In einem Simulationsmodell wird die Infektionsausbreitung in Abhängigkeit von der Populationsdichte erzeugt. Dann werden durch Anpassung an die Beobachtungswerte r geschätzt, der für die Verhinderung einer Epidemie benötigte Immunisierungsgrad berechnet und nach Ausführung der Vaccination im Modell der Erfolg überprüft. Dabei zeigt sich, daß der durch die Strukturierung bedingte "Bremseffekt" eine Funktion der Zeit ist und somit auch r von t abhängt. Durch Eingehen mit einem gemittelten r in die genannte Ungleichung läßt sich eine wirksame Abschätzung für den notwendigen Immunisierungsgrad in Abhängigkeit von der Populationsanordnung und Dichte gewinnen.

P.D. BEZEMER : Discriminant analysis for two populations with binary variables

Three distance measures for two populations with binary variables are investigated : MAHALANOBIS' distance δ^2 , coefficient of determination (from regression with equal sizes) R^2 , and the information divergence I. Relationships between these measures are given in the case of one variable. These relations are lost with more than one binary variable. Assuming equal prior probabilities and equal costs of misallocation BAYES' rule yields that there is a tendency to select variables with asymmetrical probabilities of misallocation. Finally it is shown that all measures have good and bad aspects for selecting variables.

E. BRUNNER : Ein nichtparametrischer Test für übervollständige Blockpläne

Bei einem einfaktoriellen Blockplan wird angenommen, daß pro Block mehr Versuchseinheiten vorhanden sind, als Behandlungen untersucht werden sollen. Diese Versuchseinheiten sollen zur (teilweisen) Wiederholung des Versuchs verwendet werden. Man erhält so einen Versuchsplan, der aus einem vollständigen Plan und aus einem unvollständigen zusammengesetzt ist, wobei der unvollständige Plan und die Anzahl der Blöcke so gewählt werden, daß der unvollständige Plan balanciert ist. Solche Pläne heißen "übervollständige balancierte " Pläne und wurden für den parametrischen Fall von JOHN (1962) und TRAIL, WEEKS (1973) bereits untersucht. Eine analoge nichtparametrische Teststatistik wird aufgestellt, und ihre asymptotische Verteilung wird hergeleitet. Ferner wird gezeigt, daß andere bekannte nichtparametrische Tests Sonderfälle dieses Tests sind.

U. FELDMANN : Ein Kompartiment-Modell zur Populationskinetik und seine Anwendung auf CHO T 71 Wachstum

Modelle mit homogenen Kompartimenten werden durch ein System gewöhnlicher Differentialgleichungen beschrieben; sie haben sich

bei Anwendungen in der Pharmakokinetik und Tracerkinetik bewährt. Bei einer neuen Anwendung in der Populationskinetik werden altersabhängige Kompartimente durch parametrische Übergangsfunktionen definiert und eine allgemeine Vorwärtsgleichung für das Kompartimentmodell hergeleitet. Im Grenzfall des altersunabhängigen Wachstums können bekannte Modelle abgeleitet werden, etwa das logistische Wachstum und Jäger-Beute-Modelle.

Das altersabhängige Wachstum wird am Beispiel der Synchronisation von CHO T 71 Zellen demonstriert.

J. D. F. HABBEMA : A statistical model for the classification of human chromosomes

Human cells contain 23 types of chromosomes, which are partitioned into 7 groups according to the DENVER system. Group B consists of 2 types, each type existing twice. The problem of allocating the 4 chromosomes into the two types can be solved by "discriminant analysis under restrictions". The usual and the restricted model are applied to experimental data. The allocation rule "with doubt" is also tested. The restricted model seems to yield superior results.

J. HERMANS : Nonlinear regression applied to MICHAELIS-MENTEN-kinetics

For a set of experimental data the fitting of the two following models is discussed

$$(1) \quad y = \frac{\alpha x}{\beta + x} \quad , \quad (2) \quad y = \frac{\alpha_1 x}{\beta_1 + x} + \frac{\alpha_2 x}{\beta_2 + x} \quad .$$

Three types of rearrangements of (1) lead to linear regression equation. Including the two nonlinear regressions (1), (2) one can compare 5 procedures. Varying the weights w_i in the sum of squares $\sum w_i (y_i - \hat{y}_i)^2$ (e.g. $w_i = 1$; $1/y_i$; $1/y_i^2$), the results will differ. The following problems are discussed: the choice of the criterion, available algorithm and programs, choice of weights, statistical evaluation of estimates, and design problems.

G. HOMMEL : Abgrenzung kritischer Bereiche bei zweidimensionalen Kontingenztafeln mit Hilfe der exakten Wahrscheinlichkeiten

Sind bei einer $(r \times s)$ - Kontingenztafel mit festen Randsummen die Erwartungswerte klein, so ist der χ^2 - Test nicht mehr anwendbar. Die Menge aller möglichen Kontingenztafeln A mit denselben Randsummen unterteilt man in

$$\pi_\alpha = \{A \mid a_{ij} \geq \alpha\}.$$

Die Verteilung der Prüfgrößen $\chi^2(A)$, $A \in \pi_0$ stimmt im Falle $(r,s) \neq (2,2)$ mit der exakten χ^2 - Verteilung bis auf Diskretisierungsfehler überein, solange $\beta(\chi^2) \geq \beta_0$ ist mit

$$\beta(x) = \sup \{ \alpha \mid \chi^2(A) \leq x \rightarrow A \in \pi_\alpha \}.$$

Hierbei hängt $\beta_0 < 0$ im wesentlichen nur von r, s und dem kleinsten Erwartungswert ab. Kritische Bereiche lassen sich aus den exakten Wahrscheinlichkeiten der Kontingenztafeln mit den größten Prüfgrößen bestimmen. Somit ist eine Entscheidungshilfe gegeben, ob der χ^2 - Test anwendbar ist. In manchen Fällen lassen sich χ^2 - Test und exakter Test kombinieren.

J. HORNING : Kritik der Signifikanztests

Es wird die Frage aufgeworfen, welche Aussagekraft die in der Medizin so häufig verwendeten Signifikanztests haben, mit Rücksicht auf zu unterstellende inhaltliche Fragestellungen, und gemessen an den jeweils besten bekannten Methoden. Handelt es sich darum, möglichst viel Information über einen Parameter θ zu geben, so erscheinen Konfidenzbereiche geeigneter. Im Falle einer Entscheidung zwischen Hypothesen $\theta = \theta_0$ und $\theta \neq \theta_0$ mit konstanter Verlustfunktion werden eine Abschätzung mit Hilfe eines Maximum-Likelihood-Verfahrens und eine Minimax-Methode mit einer Vorbewertung der Hypothesen diskutiert, und es wird gezeigt, daß eine willkürliche Wahl von N und α beim Signifikanztest rationalen Kriterien nicht genügen kann. Für den Fall einer nicht konstanten Verlustfunktion wird eine Entscheidungsregel aufgrund einer aus Konfidenzbereichen konstruierten Quasi-a-posteriori-Verteilung von θ , die proportional der Likelihoodfunktion ist, vorgeschlagen.

H. IMMICH : Wie kann man Arzneimittel-Nebenwirkungen testen?

Durch eine groß angelegte Studie soll die Frage geklärt werden, ob eine Assoziation zwischen der Neuroleptika-Behandlung und den Sterbefällen an Krebs (einschließlich Leukämie) besteht? Folgende Methoden werden verwendet:

1. Relatives Risiko (CORNFIELD 1951)
2. 4- und 3-dimensionale Kontingenztafeln
3. Summen- χ^2 -Test (MANTEL und HAENSZEL 1959)

Alle Methoden bestätigen: Die Krebsrate liegt unter Neuroleptika-Behandlung niedriger als unter Nicht-Neuroleptika-Behandlung. Jedoch ergibt eine weitere Analyse: Die mit Neuroleptika Behandelten sind hinsichtlich ihrer Krebssterberate und ihrer Aufenthaltsdauer in stationärer Behandlung nicht vergleichbar mit Patienten, die nicht mit Neuroleptika behandelt worden sind.

Man kann Arzneimittelnebenwirkungen also testen dann und nur dann, wenn die durch die Medikamenteneinnahme definierte Gruppe voll vergleichbar ist mit der Gruppe, welche das Medikament unter Prüfung nicht einnimmt.

J. KRAUTH : Bemerkungen zu einer Klasse von nichtparametrischen Trendtests

Ausgehend von einem Ergebnis von RÉNYI (1962) über sequentielle Ränge bewies KHURSHEED ALAM (1974) die Unverfälschtheit für eine Klasse von Tests, die auf sequentiellen Rängen beruhen. Ein Spezialfall ist dabei der Korrelationskoeffizient von KENDALL, der äquivalent ist mit der Summe der sequentiellen Ränge. Ebenso wie der SPEARMAN-Korrelationskoeffizient zu einem lokal besten Rangtest für nichtparametrische bzw. parametrische logistische Trendalternativen führt, so ergibt der auf dem KENDALL-Korrelationskoeffizienten beruhende Test einen lokal besten "Sequential"-Rangtest für die obigen Alternativen. Weiterhin sind die beiden Tests für die nichtparametrischen und beliebige parametrische Trendalternativen der obigen Form asymptotisch äquivalent.

U. KUSS : Ein neues statistisches Entscheidungsverfahren mit Robustheitseigenschaften

Eine Entscheidungsfunktion d_0 erfülle

$$\int_{\Theta} [R(d_0, \theta) - R(d, \theta)] \cdot f(\theta) \cdot \mu(d\theta) \leq c \quad \forall d \in D, \forall f \in F.$$

Hierbei sind $R(d, \theta)$ die Risikofunktion bei der Entscheidung d in Abhängigkeit vom Parameter θ , $c > 0$ eine vorgegebene Risikoschranke und D , bzw. F Klassen von Entscheidungsfunktionen, bzw. Dichten. In μ soll eine Vorbewertung im Parameterraum Θ gegeben sein, z. B. Gleichverteilung. F soll nun so beschaffen sein, daß $f d\mu$ möglichst viele interessante Verteilungen in Θ approximiert, u.a. auch die empirische Verteilung. Die Robustheitseigenschaft liegt darin, daß f auch Fehler bei der Verteilungsannahme ausgleicht. Die gesuchte Entscheidungsfunktion d^* heißt "c-zulässig", wenn F reduziert wird auf

$$F^* = \{ f \mid |f(\theta_1) - f(\theta_2)| \leq s \cdot h(\theta_1, \theta_2) \quad \forall \theta_1, \theta_2 \in \Theta \} \text{ mit } s$$

maximal und $h(\theta_1, \theta_2)$ als der HELLINGER-Distanz. Analog werden "c_x - zulässige" Entscheidungen bei gegebener Beobachtung x definiert.

Unter gewissen Voraussetzungen gelten Optimalitätseigenschaften, z. B. untere Schranken für s .

D. MORGENSTERN : Kritik an der Verwendung der Theorie stationärer Prozesse für EKG-Auswertung

Im stationären stochastischen Prozeß $X(t)$ mit $E(X(t))=0$ ist die Kovarianzfunktion $c(\tau) = E(X(t) \cdot X(t+\tau)) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t+\tau) x(t) dt$; die letzte Gleichheit gilt nur bei stetigen Spektralfunktionen. Bei rein periodischen Prozessen $X(t) = \sum_V (a_V \cos v\omega t + b_V \sin v\omega t)$ erhält man $c(\tau) = \sum_V (a_V^2 + b_V^2) / 2 \cdot \cos v\omega \tau$, d.h. die Kovarianzfunktion enthält nicht mehr die volle Information (ω, a_V, b_V) über den Prozeß. Außerdem ist die Spektralfunktion nicht stetig, sodaß obige Mittelung nicht gilt.

Für die Auswertung wird daher vorgeschlagen, jede nur nahezu gleiche Periode separat zu analysieren. Für die so erhaltenen Koeffizienten lassen sich Mittelwerte und empirische Varianzen angeben.

B. NGUYEN : BAYES' entscheidungstheoretische Lösung für den multiplen Vergleich

WALLER (1967) hat die BAYES' Lösung für den multiplen Vergleich als Entscheidungsproblem formuliert. Im Fall linearer Verlustfunktionen ist die gefundene Lösung konservativer als der TUKEY-Test (für kleine F-Werte) und weniger konservativ als der FISHER - Signifikanz-Grenzdifferenz-Test (für große F-Werte). Die Lösung von WALLER läßt sich auf quadratische Verlustfunktionen übertragen.

H. NOWAK : Über zweiseitige Intervall-Hypothesen

Anstelle der üblichen zweiseitigen Hypothesen, z. B.

$H_0: \mu - \mu_0 = 0$, $H_1: \mu - \mu_0 \neq 0$ sollen "Intervall-Hypothesen" betrachtet werden, z. B. $H_0: |\mu - \mu_0| \leq \Delta$, $H_1: |\mu - \mu_0| > \Delta$. Dabei ist Δ ein vorzugebender Genauigkeitsparameter. Derartige Intervall-Hypothesen entsprechen vielen Problemen der Medizinischen Statistik besser; jedoch werden in der Literatur keine Tests hierfür angegeben.

Einen Lösungsansatz - hier für den 1-Stichproben-Fall - bietet die Betrachtung der gemeinsamen Verteilung der beobachteten Zufallsvariable X und des zufälligen Parameters M . Dabei sind die bedingte Verteilung $X|M$ und die Randverteilung von M gegeben. Das Analogon zum GAUß-Test wird für den Sonderfall einer Gleichverteilung des Parameters M angegeben. Ist die Randverteilung eine Ein-Punkt-Verteilung und die bedingte Verteilung aus einer einparametrischen Exponentialfamilie, lassen sich zusätzliche Optimalitätseigenschaften angeben.

R. REPGES : Nichtparametrische Tests für p - faktorielle
Versuchspläne

Nach den Arbeiten von SEN (1970) für den 2^p - und von REPGES (1974) für den 3^p - Faktor - Versuch wird der allgemeine Fall behandelt, daß der Faktor i auf n_i -Stufen ($i = 1, \dots, p$) vorliegt. Das Modell sei $\underline{x} = \underline{A}\underline{T} + \underline{\epsilon}$ mit einer bezüglich Null symmetrischen und bezüglich Permutation der Argumente invarianten Verteilungsfunktion $F(\underline{\epsilon})$. Ist dann \underline{A} orthonormal, so ist $\underline{t} = \underline{A}' \underline{x}$ in jeder Komponente symmetrisch um $\underline{1}$. Für eine gegebene Teilmenge der Komponenten ist jede zweidimensionale Verteilung symmetrisch in den Komponenten und jede höherdimensionale Verteilung symmetrisch bezüglich Null. Für diese Teilmenge gelten die von SEN beschriebenen Vorzeichen - und Rangtests.

H. RIEDWYL : Signifikanztests, ein Problem

In der Regel werden vor Durchführung statistischer Tests Entscheidungen über die Art der Hypothesen und die Art der Tests getroffen, die sicher einen Einfluß auf die gesamte Irrtumswahrscheinlichkeit α haben. Es wird daher vorgeschlagen, Entscheidungsbäume aufzustellen die eine Übersicht über alle (mehrstufigen) Entscheidungen liefern und für die die Kontrolle des α gewährleistet ist. In diesem Sinn liefern bei gleichem α Simultane Vergleiche nach GABRIEL weniger "Signifikanzen" als nach SCHEFFÉ, da bei GABRIEL der Entscheidungsbaum vollständig ist und bei SCHEFFÉ nur gewisse, subjektiv ausgewählte Entscheidungen getestet werden. Am Beispiel des Schätzens und Testens des Abstands zwischen zwei Parabeln $E(Y_i) = \alpha_i + \beta_i x + \gamma_i x^2$, $i=1,2$ mit geschätzten $\hat{\alpha}_i$, $\hat{\beta}_i$, $\hat{\gamma}_i$ wird ein Entscheidungsbaum aufgestellt.

C.L. RÜMKE : Fragen bei einer Kovarianzanalyse

In der Praxis tritt häufig der Fall auf, daß ganze Versuchspläne wiederholt durchgeführt werden. Die einzelnen Versuchspläne sollen z. B. kovarianzanalytisch ausgewertet werden. Dabei kann es geschehen, daß der Einfluß der Kovariablen nur bei einem Teil der Versuche sig-

nifikant ist. Folgende Fragen stellen sich :

1. Sollen alle Versuchspläne nach dem gleichen Modell ausgewertet werden (z. B. Kovarianzanalyse), obwohl die Voraussetzungen nur teilweise erfüllt sind?
2. Wenn die Kovariable z. B. das Mäusegewicht ist und der Einfluß des Gewichts bei allen Versuchen als gleich angesehen werden sollte, darf dann ein gemeinsamer Regressionskoeffizient angegeben werden? Welche Freiheitsgrade sind in diesem Fall zu wählen?

L. SACHS : Numerischer Vergleich von 11 Konkurrenten des klassischen Vierfelder- χ^2 -Tests bei kleinem Stichprobenumfang

Anhand des dem Vierfelderhomogenitätstest zugrundeliegenden Vergleichs zweier Binomialwahrscheinlichkeiten nach GILDEMEISTER und VAN DER WAERDEN (1943) wird für den klassischen Vierfelder- χ^2 -Test mit Stetigkeitskorrektur nach YATES und 11 konkurrierende Prüfgrößen untersucht, wie gut die Approximation der vorgegebenen Irrtumswahrscheinlichkeit durch die per Programm für viele Werte von n_1 , n_2 , p , α berechnete wahre Irrtumswahrscheinlichkeit ist. Hierzu werden Abweichungsgrade und Gütekriterien definiert. Insbesondere zwei von GART (1966) vorgeschlagene modifizierte Likelihoodverhältnisse erwiesen sich als den anderen Prüfgrößen deutlich überlegen.

W. SCHAAPFMA : An introduction to the univariate classification problem

On the basis of the outcome of the univariate random variables

$$x_0, x_1^{(1)}, \dots, x_{n_1}^{(1)}, x_1^{(2)}, \dots, x_{n_2}^{(2)}$$

the individual corresponding to x_0 has to be identified with one

of the samples $x_1^{(i)}, \dots, x_{n_i}^{(i)}$, i.i.d. $N(\mu_i, \sigma^2)$, $i = 1, 2$.

The parameters μ_1, μ_2, σ^2 are unknown. A UMRI - (uniformly minimum risk invariant -) rule is constructed and related to the maximum likelihood rule. The probabilities of misclassification are given in an explicit form as a function of $(\frac{\mu_1 - \mu_2}{\sigma})^2$. For these errors a confidence interval can also be constructed. Finally another approach is discussed where the a-posteriori-probability is to be estimated.

W. STUCKY : Probleme beim Testen von Übergangswahrscheinlichkeiten in 2-stufigen biologischen Modellen

Bei einem zweistufigen Modell treten auf jeder Stufe Verluste auf, so daß die Anzahlen von ursprünglich a_0 auf a_1 und a_2 zurückgehen. Die unbekanntes Übergangswahrscheinlichkeiten sind p_{01} und p_{12} . Wird ein entsprechender Versuch bei 2 Behandlungen jeweils mit mehreren Ausgangsgruppen wiederholt, so wird dieser Versuch meist mit Hilfe des t-Tests oder des χ^2 -Tests für die erhaltenen Anzahlen ausgewertet. Diese Verfahren sind in der Regel als ungeeignet zu verwerfen. Als mögliche Alternative kommt die Prüfung der relativen Häufigkeiten mit nichtparametrischen Verfahren in Frage. Dabei neu auftretende Probleme werden diskutiert.

H. THÖNI : "Korrelation" unkorrelierter Beobachtungen

In einem Experiment wird das Wachstum w_1, \dots, w_n von n Pflanzen in einer Reihe gemessen. Korreliert man die benachbarten Paare $x_i = w_i, y_i = w_{i+1}, i = 1, \dots, n-1$, so erhält man einen von Null signifikant verschiedenen Korrelationskoeffizienten. Damit kann ein Biologe eine nachbarliche Beeinflussung des Wachstums nachweisen! Eine theoretische Überprüfung zeigt jedoch, daß bei unabhängigen w_i die Punkte (x_i, y_i) nicht mehr unabhängig sind. Eine Approximation des Erwartungswertes des Korrelationskoeffizienten liefert einen von Null verschiedenen Wert und erklärt den beobachteten Korrelationskoeffizienten. (Anmerkung: Für $n = 3$ würde man stets $|r| = 1$ beobachten.)

N. VICTOR: Fehlerraten von Zuordnungsregeln bei unvollständiger Information

Die auf SKARABIS (1970) zurückgehenden Definitionen verschiedener Fehlerraten werden auf den Fall der Verwendung unzutreffender Modelle ausgedehnt, d. h. anstelle der Wahrscheinlichkeitsräume $(X, F_X, P(\theta_1))$ werden $(X, F_X, P(\theta_1^i))$ ($i=1, \dots, n$ Gruppen) zugrunde gelegt mit $\mathcal{W}^1 \subset \mathcal{W}$ für die zugehörigen Parameterräume. Die Verwendbarkeit verschiedener Schätzmethoden für diesen Fall wird kurz diskutiert. Es wird vorgeschlagen, für praktische Probleme das Modell aus geeigneten Modellfamilien $(X, F_X, P(\theta_1^j))$ ($j=1, \dots, k$), mit $\theta_1^j \in \mathcal{W}^j \quad \forall 1, j$ und $\mathcal{W}^1 \subset \dots \subset \mathcal{W}^k$ so auszuwählen, daß die zugehörige "plug-in-rule" minimale zu erwartende Fehlerraten liefert.

E. WALTER: Eine Modifikation des Vorzeichentests und des Median-tests

Um die Effizienz des Vorzeichentests für die Prüfung der Symmetrie bezüglich Null und des Mediantests für die Prüfung zweier unabhängiger Stichproben zu erhöhen, wird vorgeschlagen, die mittleren Beobachtungen nicht zu berücksichtigen (beim Vorzeichentest, die r dem Betrage nach kleinsten, beim Mediantest jeweils $r/2$ Beobachtungen, die gerade kleiner bzw. größer als der gemeinsame Median sind). Auf die äußeren Beobachtungen ist dann der übliche Test anzuwenden. Die Anzahl der nicht zu berücksichtigenden Beobachtungen wurde für die üblichen Signifikanzgrenzen so bestimmt, daß die lokale Schärfe bei Normalverteilung maximiert wird. Die asymptotische Effizienz bezüglich des t-Tests wird dadurch in beiden Fällen von 0,64 auf 0,81 erhöht. Für kleine n kommt hierzu noch die Erhöhung durch die bessere Ausnutzung der Irrtumswahrscheinlichkeit. Auch die dem modifizierten Vorzeichentest entsprechenden Konfidenzbereiche sind im Mittel kleiner.

Da die optimale Anzahl der nicht zu berücksichtigenden Elemente von α abhängt, können Stichproben konstruiert werden, die z. B. bei $\alpha = 0,01$ signifikant, bei $\alpha = 0,05$ aber nicht signifikant sind.

S. WELLEK: Zur statistischen Analyse medizinischer Längsschnittdaten mittels MARKOFF - Ketten

N zufällige Beobachtungseinheiten werden intervallweise zwischen einem individuell variablen Erstbeobachtungsalter $x_0^{(i)}$ ($i=1, \dots, N$) und einem für alle Objekte gleichen Zeitpunkt w der letzten Datenerhebung nach den folgenden beiden Merkmalen beschrieben:

- (1) Überleben oder Absorption durch eine von a möglichen Abgangsursachen
- (2) Zugehörigkeit zur 1-ten, 2-ten, ... ,g-ten Ausprägung eines beliebigen g-stufigen Klassifikationsmerkmals ("Typenmerkmal")

Diese Beobachtungen werden betrachtet als bezüglich des Prozeßanfangs mehr oder minder unvollständige Realisierungen einer zeitlich inhomogenen MARKOFF-Kette 1. Ordnung mit g transienten und $a \cdot g$ absorbierenden Zuständen. Für die Parameter der Kette werden bedingterwartungstreue ML-Schätzer angegeben. Außerdem werden die zweiten zentralen Momente dieser Schätzfunktionen untersucht.

H. WITTING: Zur Robustheit von Tests

Da bei der Anwendung eines statistischen Tests die Modellannahme, auf die er zugeschnitten ist, nie exakt erfüllt sein wird, müssen Tests weitgehend unempfindlich gegenüber Abweichungen von der Modellannahme sein. Für einfache Modelle, bei denen die Prüfgröße

von der Form $\hat{T}_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n T(x_i)$ ist (T monoton; $X_i, i \in N$;

stochastisch unabhängige gleiche Verteilung; Regulationsvoraussetzungen erfüllt), wird gezeigt, daß Abweichungen von der Modellannahme im Sinne der ϵ - contamination und δ - totale - Variation bereits für beliebig kleine $\epsilon, \delta > 0$ zu unendlichen Mittelwertverschiebungen der Prüfgröße führen können, falls T unbeschränkt ist. Ist T beschränkt, so läßt sich erreichen, daß der Test (auf

Kosten der Schärfe) das Niveau einhält und für hinreichend kleine ϵ, δ unverfälscht wird. Ein Vergleich über die PITMAN-Effizienz spricht auch hier zugunsten der robusten Tests, verg. HUBER (1965), RIEDER (1974).

F.X. WOHLZOGEN: Zur Wahl zweckmäßiger Testpläne bei sequentiellen Paarvergleichen

Für 2 Verfahren (A,B) sei die relative Überlegenheit von A gegenüber B im Paarvergleich zu prüfen. Der Test soll sequentiell als SPRT nach WALD durchgeführt werden. Ein Paar sei mit + bzw. - bewertet, wenn A dem Verfahren B über- bzw. unterlegen ist. Wenn die Effekte der Verfahren messbar sind, können theoretisch nur Paare mit der Bewertung + oder - auftreten. Der Anteil der + Paare sei \hat{p} . Im SPRT wird die Hypothese $p \geq p_0$ ($0 < p_0 < 1$) geprüft. Der "Indifferenzbereich" sei $p_0 - d < p < p_0 + d$. Die Wahrscheinlichkeiten für Fehlentscheidungen - Ablehnung der Hypothese, wenn $p \geq p_0 + d$, oder Annahme der Hypothese, wenn $p \leq p_0 - d$, - seien $\alpha = \beta$.

Sind die Effekte der zu vergleichenden Verfahren nicht messbar sondern nur qualitativ bewertbar (z. B. "gebessert", "nicht gebessert") so interessiert der Unterschied zwischen den Erfolgsraten unter A und B : $|\epsilon_A - \epsilon_B| = 2\delta$. Sei ϵ die mittlere Erfolgsrate, so besteht zwischen d und δ die Beziehung $d = \delta / 2(\epsilon - \epsilon^2 + \delta^2)$. Sind ϵ_A und ϵ_B bekannt, so können p , d und die zu erwartende Anzahl von "ties", d. h. (+,+) oder (-,-), berechnet werden. Unter diesen Voraussetzungen können zweckmäßige Testpläne gewählt werden, da der voraussichtlich für eine Entscheidung erforderliche Stichprobenumfang nur noch von der Wahl der Irrtumswahrscheinlichkeiten α, β abhängt.

R. ZENTGRAF: Einige Bemerkungen zu LANCASTERS Definition für das Fehlen von Interaktionen höherer Ordnung

Die additive Definition von LANCASTER für die Beschreibung von Interaktionen zwischen n Zufallsvariablen wird benutzt, um eine

Darstellung für die gemeinsame Verteilungsfunktion dieser n Zufallsvariablen für den Fall anzugeben, daß alle Interaktionen ab einer s -ten, $2 \leq s \leq n - 1$, Ordnung verschwinden. Diese Darstellung benutzt nur die 1 -, 2 -, ..., s -dimensionalen Randverteilungen. Es werden die Probleme bei der Anwendung dieser Darstellung auf qualitative Daten geschildert und für den Spezialfall dichotomer Variabler Beziehungen zwischen LANCASTERS Darstellung und solchen von BAHADUR und LAZARFELD erläutert.

H. Nowak (Aachen)

2
.
.
V

