

Hintergrundfaktoren bei qualitativen Variablen – ein Computerprogramm

Von N. Wermuth, B. K. Yun und H. Gönner

Zusammenfassung

Mögliche Einflußfaktoren auf den Zusammenhang zweier Variablen sind häufig bekannt. Ob ein Hintergrundfaktor jedoch einen wesentlichen Einfluß ausübt, muß jeweils anhand von Beobachtungen entschieden werden. Bei qualitativen Variablen bietet die Theorie der logarithmisch-linearen Modelle eine Entscheidungshilfe. Wir beschreiben ein entsprechendes Computerprogramm.

Summary

Potential confounding variables for the interrelation of two variables are frequently well-known. But, for each set of data it is necessary to decide whether a back-ground factor exerts a substantial influence or not. In the case of qualitative variables such a decision can be guided by the theory of log-linear models. We describe an appropriate computer program.

Problemstellung

Bei empirischen Untersuchungen sollten stets die Auswirkungen von Hintergrundfaktoren bedacht werden. Dies ist nötig, um Fehlinterpretationen zu vermeiden. Es kann zum Beispiel der Zusammenhang von zwei Variablen oder Merkmalen verfälscht wiedergegeben werden, wenn der Einfluß eines dritten Merkmals nicht erkannt wird. Bei medizinischen Studien werden mögliche Hintergrundfaktoren wie Alter, Geschlecht, Krankheiten in der Anamnese eines Patienten häufig routinemäßig miterfaßt. Wie stark jedoch solche Merkmale einen untersuchten Zusammenhang beeinflussen oder gar verändern, ist in jedem einzelnen Fall zu prüfen. Besonders wichtig ist eine Überprüfung im allgemeinen, wenn die Vergleichbarkeit verschiedener Kollektive nicht mit Hilfe von geplanten Experimenten sichergestellt werden kann, und im besonderen, wenn das ursächliche Wirken einer Variablen erforscht werden soll (COCHRAN 1965, KOLLER 1964), also die Auswirkungen von Hintergrundfaktoren mit denen des vermuteten ursächlichen Faktors verwechselt werden könnten.

Bei qualitativen oder klassifizierten quantitativen Merkmalen kann die Frage nach der Wirkung eines Hintergrundfaktors wie folgt gestellt werden: Genügt es, einen Zusammenhang so darzustellen, wie er sich im Gesamtkollektiv vorliegender Beobachtungen widerspiegelt oder muß die Aussage über die Art des Zusammenhangs modifiziert werden, weil in Teilkollektiven, die durch den Hintergrundfaktor festgelegt sind, andere Abhängigkeiten bestehen als im Gesamtkollektiv? Wegen der Wichtigkeit dieser Frage erstellten wir ein Computerprogramm, mit dessen Hilfe eine Antwort darauf gegeben werden kann. Das Programm berechnet Tests für eine Dreifaktor- und eine Zweifaktor-Interaktion in einer dreidimensionalen Kontingenztafel im Rahmen logarithmisch-linearer Modelle (BIRCH 1963, GOODMAN 1970,

BISHOP et al. 1975, WERMUTH 1976). Die Diskussion darüber, wie Interaktionen in Kontingenztafeln zu definieren und zu prüfen sind, ist noch nicht abgeschlossen (BARTLETT 1936, ROY und KASTENBAUM 1956, DARROCH 1962, 1974). Alternativverfahren zu den von uns verwendeten Tests im Sonderfall von zweiklassigen Merkmalen, genauer für $2 \times 2 \times K$ Kontingenztafeln, wurden unlängst verglichen (FLEISS 1973, MCKINLAY 1975). Die Plausibilität unseres Vorgehens wird an verschiedenen Daten dargestellt.

Methodik

Wir nehmen an, daß insgesamt n Beobachtungen für zwei hauptsächlich untersuchte Merkmale (Variable 1 und 2) und für einen Hintergrundfaktor (Variable 3) vorliegen. Dabei kann der Hintergrundfaktor entweder ein einzelnes Merkmal oder aber eine Kombination aus mehreren Einzelmerkmalen sein. Die Variable 1 habe I Ausprägungen oder Klassen; $i=1, \dots, I$ kennzeichne die einzelnen Ausprägungen. Weiterhin bezeichne $j=1, \dots, J$ und $k=1, \dots, K$ die Ausprägungen der Variablen 2 bzw. der Variablen 3. Wird jede einzelne Beobachtung einer der Ausprägungskombinationen (i, j, k) zugeteilt, so erhält man eine dreidimensionale $(I \times J \times K)$ Kontingenztafel, in der n_{ijk} die beobachtete und m_{ijk} die unter bestimmten Modellannahmen - erwartete Fallzahl in der Zelle (i, j, k) angeben. Folgen die Zellenbesetzungen einer Multinomialverteilung, so ist die Wahrscheinlichkeit (p) für $n = \sum_{ijk} n_{ijk}$ und für jede Ausprägungskombination durch (1) gegeben:

$$p = \frac{n! n^{-n}}{\prod_{i,j,k} n_{ijk}!} \prod_{i,j,k} (m_{ijk})^{n_{ijk}}. \quad (1)$$

Der Logarithmus der Likelihoodfunktion wird damit

$$\ln L = \ln \frac{n!}{\prod_{i,j,k} n_{ijk}!} - n \ln n + \sum_{i,j,k} n_{ijk} \ln m_{ijk} \quad (2)$$

Nach BIRCH (1963) lässt sich die erwartete Fallzahl m_{ijk} als Summe von Haupt- und Wechselwirkungsparametern darstellen:

$$\ln m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)}, \quad (3)$$

wobei sich die Effektparameter ähnlich wie in der Varianzanalyse zu Null summieren: Für die Hauptwirkungsparameter gilt:

$$\sum_i u_{1(i)} = \sum_j u_{2(j)} = \sum_k u_{3(k)} = 0;$$

für die Parameter, die die Zweifaktor-Interaktionen erfassen:

$$\begin{aligned} \sum_j u_{12(ij)} &= \sum_k u_{13(ik)} = 0 & \forall i \\ \sum_i u_{12(ij)} &= \sum_k u_{23(jk)} = 0 & \forall j \\ \sum_i u_{13(ik)} &= \sum_j u_{23(jk)} = 0 & \forall k; \end{aligned}$$

und für die Parameter der Dreifaktor-Interaktion:

$$\begin{aligned} \sum_i u_{123(ijk)} &= 0 & \forall jk; \\ \sum_j u_{123(ijk)} &= 0 & \forall ik; \\ \sum_k u_{123(ijk)} &= 0 & \forall ij. \end{aligned}$$

In diesem sogenannten gesättigten Modell (GOODMAN 1970) ist der Maximum-Likelihood Schätzer \hat{m}_{ijk} für m_{ijk} die beobachtete Fallzahl n_{ijk} .

Wenn dagegen keine Dreifach-Interaktion vorliegt, so sind $u_{123(ijk)} = 0 \forall ijk$ und der von den Parametern abhängige Teil der Likelihoodfunktion (2) vereinfacht sich zu

$$\ln L \propto n u + \sum_i n_{i..} u_{1(i)} + \sum_j n_{.j.} u_{2(j)} + \sum_k n_{..k} u_{3(k)} + \sum_{ij} n_{ij.} u_{12(ij)} + \sum_{ik} n_{i.k} u_{13(ik)} + \sum_{jk} n_{.jk} u_{23(jk)}, \quad (4)$$

mit z.B. $n_{ij.} = \sum_k n_{ijk}$ als den Fallzahlen in der zweidimensionalen Randtafel der Variablen 1, 2. Aus (4) ist abzulesen, daß bei fehlender Dreifaktor-Interaktion die zweidimensionalen (und deren) Randtafeln ausreichen, um alle Effektparameter zu schätzen. Die Bezeichnung dieses Modells mit 12/13/23 soll daran erinnern. Außerdem stimmen für Modell 12/13/23 alle drei beobachteten zweidimensionalen Randtafeln mit den Randtafeln der Maximum-Likelihood Schätzer überein (BIRCH 1963):

$$\hat{m}_{ij.} = n_{ij.}; \quad \hat{m}_{i.k} = n_{i.k}; \quad \hat{m}_{.jk} = n_{.jk}. \quad (5)$$

Ein iterativer Anpassungsalgorithmus ist nötig, um die Schätzer \hat{m}_{ijk} , die der Bedingung (5) genügen, zu bestimmen (BISHOP 1969). Man beginnt mit $m_{ijk}^{(0)} = 1 \forall ijk$ und berechnet

$$I. \quad \hat{m}_{ijk}^{(1)} = \frac{n_{ij.}}{\hat{m}_{ij.}^{(0)}} \hat{m}_{ijk}^{(0)}$$

$$II. \quad \hat{m}_{ijk}^{(2)} = \frac{n_{i.k}}{\hat{m}_{i.k}^{(1)}} \hat{m}_{ijk}^{(1)} \quad (6)$$

$$III. \quad \hat{m}_{ijk}^{(3)} = \frac{n_{.jk}}{\hat{m}_{.jk}^{(2)}} \hat{m}_{ijk}^{(2)}$$

Ein Anpassungszyklus besteht aus Schritt I bis III. Vom zweiten Zyklus an wird $\hat{m}_{ijk}^{(0)}$ durch $\hat{m}_{ijk}^{(3)}$ aus dem vorhergehenden Zyklus ersetzt. Die Iterationen enden, sobald die Abweichungen zwischen $\hat{m}_{ijk}^{(2)}$ und $\hat{m}_{ijk}^{(3)}$ klein genug sind. Unter der Bedingung, daß alle $m_{ijk} > 0$, ist nachgewiesen (BROWN 1959), daß dieser Algorithmus zu den durch (5) definierten Maximum-Likelihood Schätzern führt.

Eine fehlende Dreifaktor-Interaktion impliziert, daß die sogenannten Kreuzproduktverhältnisse einander gleich sein müssen (ROY und KASTENBAUM 1956):

$$\frac{m_{IJK} m_{ijk}}{m_{iJK} m_{Ijk}} = \frac{m_{IJK} m_{ijk}}{m_{iJK} m_{Ijk}} \quad (7)$$

$$(1 \leq i \leq I-1; \quad 1 \leq j \leq J-1; \quad 1 \leq k \leq K-1).$$

In diesem Sinne kann bei Modell 12/13/23 von einem gleichartigen Zusammenhang in allen Teilkollektiven gesprochen werden. Die inhaltliche Interpretation eines Kreuzprodukt-Verhältnisses als relative Chance (odds-ratio) wird später am Beispiel beschrieben.

Zur Überprüfung, ob bestimmte Modellannahmen in einer vorliegenden Kontingenztafel erfüllt sind, verwenden wir Likelihoodquotiententests (abgekürzt $LQ-\chi^2$). So wird Modell 12/13/23, bzw. die Hypothese $H_{01}: u_{123(ijk)} = 0 \forall ijk$, mit der Teststatistik

$$\chi^2_{[12/13/23]} = -2 \ln \prod_{i,j,k} \left[\frac{\hat{m}_{ijk}}{n_{ijk}} \right]^{n_{ijk}} \quad (8)$$

überprüft. Dabei ist \hat{m}_{ijk} durch (5) oder (6) bestimmt und für große n ist (8) annähernd Chi-Quadrat-verteilt mit $(I-1)(J-1)(K-1)$ Freiheitsgraden.

Wird H_{01} abgelehnt, so bestehen ungleichartige Abhängigkeiten in den Teilkollektiven und unsere Ausgangsfrage ist beantwortet: es liegt ein wesentlicher Einfluß des Hintergrundfaktors vor. In diesem Fall wäre es irreführend, nur den Zusammenhang im Gesamtkollektiv zu präsentieren und zu interpretieren. Stattdessen muß die Art des Zusammenhangs für die Variablen 1 und 2 in jedem der K Teilkollektive des Hintergrundfaktors getrennt beschrieben werden.

Wird dagegen H_{01} nicht abgelehnt, so sind zur Beurteilung des Hintergrundfaktors zwei weitere Hypothesen zu prüfen, diejenigen, daß die partielle (H_{02}) und daß die marginale (H_{03}) Abhängigkeit der Variablen 1 und 2 fehlen. Stimmen die Testergebnisse für H_{02} und H_{03} überein, derart, daß entweder beide ein signifikantes oder beide ein nicht-signifikantes Resultat ergeben, so kann man folgern, daß der Einfluß des Hintergrundfaktors auf den Zusammenhang zwischen Variable 1 und 2 unerheblich ist. Es genügt in

diesem Fall, den Zusammenhang im Gesamtkollektiv darzustellen. Wird dagegen H_{03} abgelehnt, aber H_{02} kann nicht verworfen werden¹⁾, so bedeutet dies, daß im Gesamtkollektiv ein Zusammenhang durch den Hintergrundfaktor nur vorgetäuscht wird, daß die beiden Variablen aber tatsächlich in den Teilkollektiven voneinander unabhängig sind.

Zu prüfen ist H_{02} : $u_{12}(ij) = 0 \forall ij$, unter der Voraussetzung $u_{123}(ijk) = 0 \forall ijk$, mit der Teststatistik

$$\chi^2 [\text{part. } 12] = -2 \ln \prod_{i,j,k} \left[\frac{\hat{m}_{ijk}}{\hat{m}_{i.jk}} \right]^{n_{ijk}}, \quad (9)$$

die annähernd einer Chi-Quadrat-Verteilung mit $(I-1)(J-1)$ Freiheitsgraden folgt. Dabei ist $\hat{m}_{i.jk}$ wie zuvor definiert und \hat{m}_{ijk} ist der Maximum-Likelihood Schätzer für dasjenige Modell, in dem die unbedingte Unabhängigkeit der Variablen 1 und 2 postuliert wird (Modell 13/23):

$$\hat{m}_{ijk} = \frac{n_{i.k} \cdot n_{.jk}}{n_{..k}} \quad (10)$$

Eine einfachere Berechnung von (9) ist möglich, wenn man berücksichtigt, daß für die Likelihood-Quotientenprüfgrößen die folgende Additivitätseigenschaft gilt:

$$\chi^2 [\text{part. } 12] = \chi^2 [13/23] - \chi^2 [12/13/23] \quad (11)$$

Dabei ist

$$\chi^2 [13/23] = -2 \ln \prod_{i,j,k} \left[\frac{\hat{m}_{ijk}}{n_{i.jk}} \right]^{n_{ijk}} \quad (12)$$

die Prüfgröße für Modell 13/23, bzw. für H :

$$u_{12}(ij) = u_{123}(ijk) = 0 \text{ (BIRCH 1963).}$$

Marginale Unabhängigkeit liegt vor, wenn die beiden Variablen im Gesamtkollektiv unabhängig sind, wenn also die Hypothese

$$H_{03}: m_{ij.} = \frac{m_{i..} \cdot m_{.j.}}{m_{...}} \text{ gilt, die mit der Teststatistik}$$

$$\chi^2 [1/2] = -2 \ln \prod_{i,j} \left[\frac{n_{i..} \cdot n_{.j.} / n_{...}}{n_{ij.}} \right]^{n_{ij.}}$$

bei $(I-1)(J-1)$ Freiheitsgraden geprüft wird.

Das Flußdiagramm der wichtigsten Subroutine (Abb. 1) zeigt, daß nur die Hypothese der gleichartigen Abhängigkeit in allen Teilkollektiven (H_{01}) und die Hypothese der Unabhängigkeit im Gesamtkollektiv (H_{03}) geprüft werden, falls H_{01} abgelehnt wird; daß aber - falls H_{01} nicht abgelehnt wird - zusätzlich die Hypothese der partiellen Unabhängigkeit (H_{02}) überprüft wird.

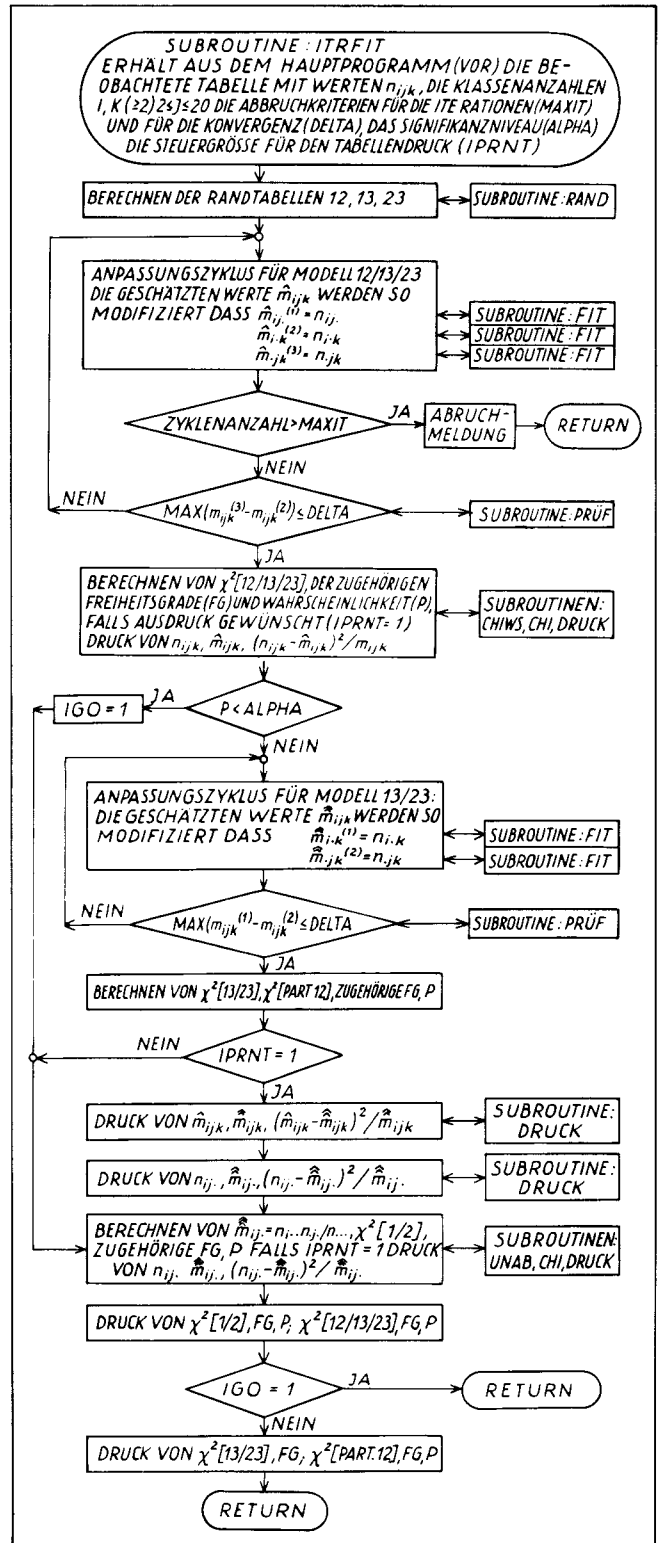
Ergebnisse

Wir stellen das geschilderte Prüfverfahren an drei Beispielen mit jeweils andersgeartetem Einfluß des

¹⁾ Die Möglichkeit, daß eine partielle Assoziation nachgewiesen wird (H_{02} wird abgelehnt) und daß gleichzeitig H_{03} nicht abgelehnt werden kann, ist zwar theoretisch nicht auszuschließen (BIRCH 1963), aber praktisch fast nie zu erwarten.

Hintergrundfaktors vor. Im ersten Beispiel gibt es einen wesentlichen Einfluß des Hintergrundfaktors, weil in den Teilkollektiven verschiedenartige, gegenläufige Assoziationen vorliegen. Im zweiten Beispiel ist der Einfluß des Hintergrundfaktors unwesentlich, da die beiden Variablen in allen Teilkollektiven gleichartig assoziiert und marginal wie auch partielle stark voneinander abhängig sind. Im dritten und letzten Beispiel liegt wieder ein

Abb. 1: Flußdiagramm der wichtigsten Subroutine



deutlicher Einfluß des Hintergrundfaktors vor, weil die beiden interessierenden Variablen im Gesamtkollektiv (also marginal) als abhängig erscheinen, während in den Teilkollektiven (also partiell) tatsächlich keine Abhängigkeit nachgewiesen werden kann.

Der erste Datensatz ist G. LIENERT (1969) entnommen und gibt das Ergebnis eines Experimentes mit 65 Personen wieder. Wir untersuchen den Zusammenhang zwischen Affektivitätsbeeinflussung (Variable 1) und Denkstörung (Variable 2) nach LSD-Einnahme. Bewußtseinstrübung (Variable 3) wird als potentieller Hintergrundfaktor angesehen. Jedes Merkmal ist zweiklassig. Abb. 1 zeigt zunächst den Zusammenhang für die Variablen 1 und 2 im Gesamtkollektiv.

Die beiden Symptome Affinitätsbeeinflussung und Denkstörung treten scheinbar voneinander unabhängig auf. Das zeigt sich im Testergebnis für H_{03} , die Unabhängigkeitshypothese im Gesamtkollektiv. Es zeigt sich auch darin, daß das beobachtete Kreuzproduktverhältnis annähernd gleich eins ist: die Chancen für eine Änderung in der Affektivität sind 23 zu 11 bei Personen mit Denkstörung und 19 zu 12 bei Personen ohne Denkstörung. Das Kreuzproduktverhältnis

$$\frac{23 \cdot 12}{11 \cdot 9} = \frac{23}{19} \cdot \frac{12}{12} = 1,3$$

gibt diese sogenannte relative Chance für die Affektivitätsbeeinflussung wieder.

Daß die beiden Symptome jedoch nur scheinbar unabhängig sind, wird anhand des Testergebnisses für H_{01} eindeutig nachgewiesen: Modell 12/13/13 bzw. die Annahme einer fehlenden Drei-Faktor-Interaktion ist nicht mit den Daten vereinbar. Bei der Betrachtung der durch den Hintergrundfaktor definierten Teilkollektive zeigt sich die Bewußtseinstrübung als deutlich dominierendes Symptom: Im Kollektiv der 37 Personen mit getrübttem Bewußtsein werden entweder die beiden anderen Symptome gleichzeitig oder keines der beiden Symptome beobachtet. Im Kollektiv der 28 Personen mit ungetrübttem Bewußtsein dagegen treten die beiden Symptome fast nur isoliert auf. Diese gegenläufigen Assoziationen in den Teilkollektiven bewirken die scheinbare Unabhängigkeit der Symptome 1 und 2 im Gesamtkollektiv.

Die Daten des zweiten Beispiels entstammen der von der Deutschen Forschungsgemeinschaft getragenen, prospektiven Studie "Schwangerschaftsverlauf und Kindesentwicklung". In Abb. 3 wird der zwischen Schwangerschaftsausgang (Variable 1) und Schwangerschaftsdauer (Variable 2) bestehende, medizinisch leicht erklärbare Zusammenhang gezeigt: je kürzer die Gestationszeit, desto geringer sind die Überlebenschancen des Neugeborenen.

Man erwartet intuitiv, daß dieser Zusammenhang in allen Kliniken (Variable 3) gleichartig sein muß. Daher eignen sich Beobachtungen für diese drei Merkmale dazu, die Leistungsfähigkeit unseres Verfahrens weiter zu veranschaulichen.

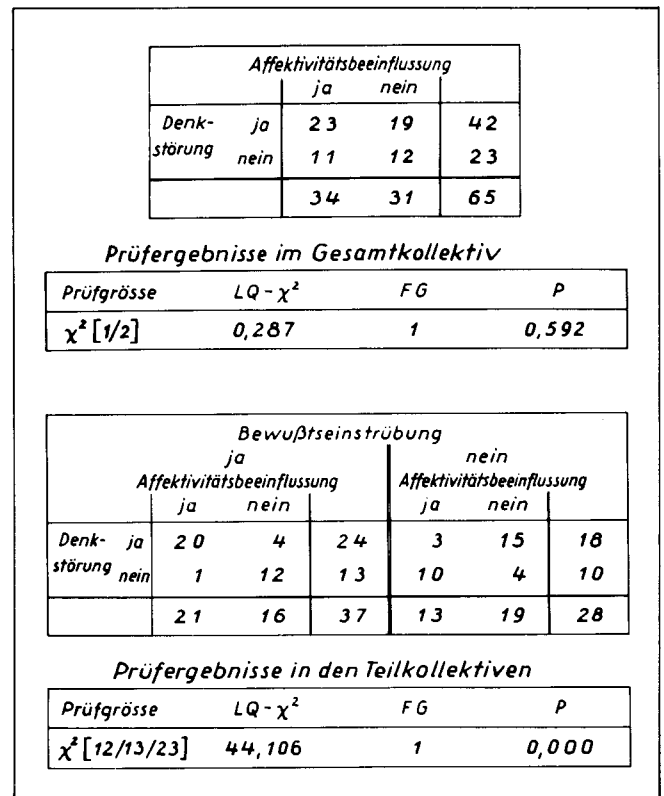
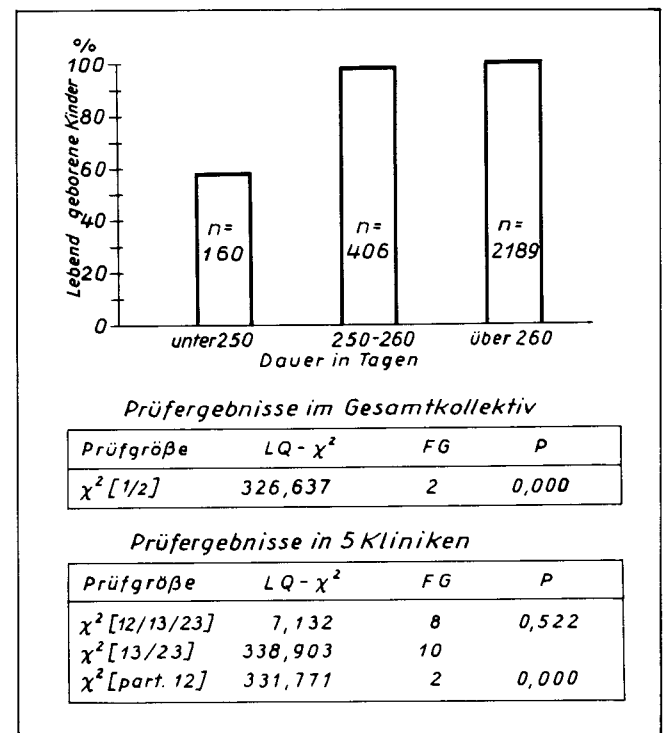


Abb. 2: Symptome nach LSD-Einnahme

Die Prüfergebnisse bestätigen die Erwartungen. Die Annahme der gleichartigen Abhängigkeit in den Kliniken kann nicht widerlegt werden und Schwangerschaftsausgang und Schwangerschaftsdauer sind im Gesamtkollektiv wie auch in den Teilkollektiven (den einzelnen Kliniken) extrem stark

Abb. 3: Schwangerschaftsausgang und Schwangerschaftsdauer für 2755 Neugeborene



assoziiert. Es besteht daher kein wesentlicher Einfluß des Hintergrundfaktors Klinik, und es genügt, die Art der Abhängigkeit im Gesamtkollektiv darzustellen.

Nur um zu verdeutlichen, in welchem Sinne die Variablen 1 und 2 in allen Kliniken gleichartig assoziiert sind, zeigen wir in Tabelle 1 die beobachteten und die unter der Annahme H_{01} geschätzten Werte für jede der fünf Kliniken. Gemäß Gleichung (7) müssen mehrere Kreuzproduktverhältnisse einander gleich sein. So geben

$$0,008 = \frac{1,16/204,84}{8,15/10,85} = \frac{2,27/665,73}{15,22/33,78} = \dots = \frac{4,20/218,80}{11,47/4,53}$$

und

$$0,31 = \frac{1,16/204,84}{0,69/37,31} = \frac{2,27/665,73}{1,51/137,49} = \dots = \frac{4,20/218,80}{2,33/37,67}$$

die geschätzten relativen Risiken für totgeborene oder perinatal verstorbene Kinder bei Tragzeiten über 260 Tagen und unter 250 Tagen und bei Tragzeiten über 260 Tagen und 251 bis 260 Tagen an. Es wird somit in allen fünf Kliniken geschätzt, daß bei Tragzeiten über 260 Tagen ein etwa um ein Drittel kleineres Risiko für einen schlechten Schwangerschaftsausgang besteht, als bei Tragzeiten zwischen

Tabelle 1: Schwangerschaftsausgang und Schwangerschaftsdauer in fünf Kliniken

Klinik	lebend geb. Kind	Schwangerschaftsdauer in Tagen					
		unter 250	251 bis 260	261 und mehr	unter 250	251 bis 260	261 und mehr
Bamberg	ja	11*	10,85**	37*	37,31**	205*	204,84**
	nein	8	8,15	1	0,69	1	1,16
Berlin-Charlottenburg	ja	33	33,78	139	137,49	665	665,73
	nein	16	15,22	0	1,51	3	2,27
Gießen	ja	13	11,40	59	61,01	280	279,59
	nein	6	7,60	3	0,99	1	1,41
Tübingen	ja	31	31,44	125	124,52	806	806,04
	nein	26	25,56	2	2,48	5	4,96
Kiel	ja	4	4,53	38	37,67	219	218,80
	nein	12	11,47	2	2,33	4	4,20

* beobachtete Werte ** unter H_{01} geschätzte Werte

Tabelle 2: Augenhintergrunddiagnose und Leistungs- oder Funktionsstörungen des Auges in drei Altersgruppen

Alter in Jahren	Leistungs- oder Funktionsstörungen	Fundusurteil			
		normal	nicht organisch verändert	leichte organische Veränderung	erhebliche organische Veränderung
unter 35	ja	37	39	16	3
	nein	61	42	12	4
36 bis 50	ja	4	43	36	13
	nein	11	24	40	8
51 und mehr	ja	5	14	32	11
	nein	0	1	3	2

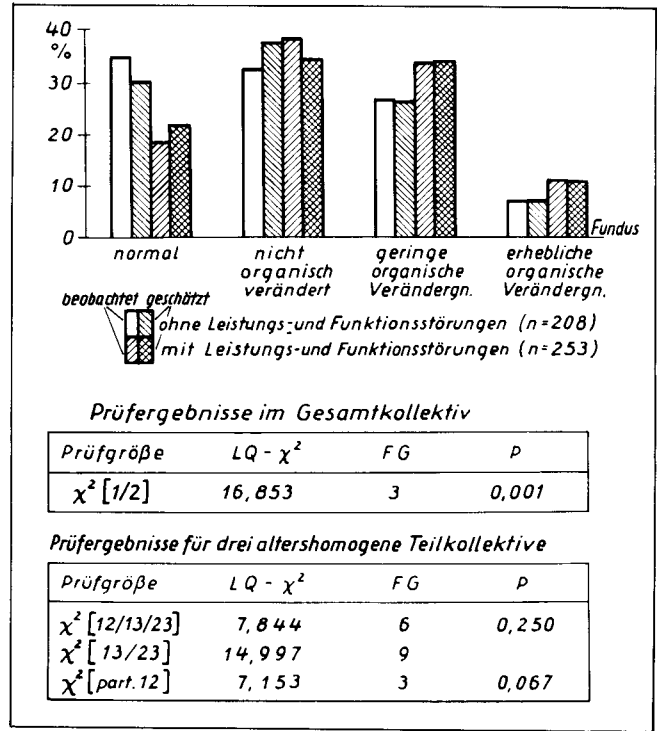


Abb. 4: Augenhintergrund und Augenstörungen bei 461 Hypertonikern

251 und 260 Tagen. Daß die so geschätzten Werte wenig von den Beobachtungen abweichen, ist sowohl aus Tabelle 1 als auch an dem Testergebnis für H_{01} in Abb. 3 abzulesen.

Die Daten des dritten Beispiels (Tab.2) wurden im Rahmen des Sonderforschungsbereiches (SFB) 36 im Teilprojekt Ophthalmologie (Leitung: A. Nover, P. Steinbach) erhoben. Für 461 Hypertoniker unterschiedlichen Alters interessierte die Abhängigkeit von Leistungs- oder Funktionsstörungen des Auges (Variable 1) und Augenhintergrunddiagnose (Variable 2). Die beobachteten Werte im Gesamtkollektiv in Abb. 4 lassen die beiden Merkmale als deutlich assoziiert erscheinen, normaler Augenhintergrund ohne Störungen des Auges treten deutlich häufiger auf als normaler Augenhintergrund mit Störungen (und $\chi^2 [1/2] = 16,85$ bei 3 Freiheitsgraden).

Als möglicher Hintergrundfaktor wurde das Alter vermutet: Leistungs- oder Funktionsstörungen des Auges häufen sich bei älteren Patienten und die Veränderungen im Augenhintergrund sind von der Hypertoniedauer, damit ebenfalls vom Alter der Patienten abhängig. Bei einer nur groben Aufgliederung der Patienten in drei Altersgruppen (unter 35, 36 bis 50 und über 50 Jahre) bestätigt sich diese Vermutung. Werden die drei nach Alter unterschiedenen Teilkollektive (mit je 214, 179 und 68 Patienten) untersucht, so

läßt sich keine wesentliche Abhängigkeit von Augenhintergrunddiagnose und Leistungs- oder Funktionsstörungen nachweisen ($\chi^2 [12/13/23] = 7,84$ bei 6 und $\chi^2 [\text{part. } 12] = 7,15$ bei 3 Freiheitsgraden). Die altersinhomogene Patientengruppe und die Altersabhängigkeit jeder der beiden Variablen täuschte die starke Abhängigkeit im Gesamtkollektiv der Patienten vor. In Abb. 4 sind außer den beobachteten noch geschätzte Prozentzahlen ($100 \times (\sum_k n_{i.k} \cdot n_{.jk} / n_{..k}) / n_{i..}$) angegeben. Diese bestätigen erneut, daß mit der Unabhängigkeitsannahme unter Berücksichtigung der altersinhomogenen Zusammensetzung der Patienten, die beobachteten Werte im Gesamtkollektiv gut reproduziert werden.

Diskussion

Wir haben beschrieben, wie bei qualitativen Daten die Wirkung von Hintergrundfaktoren mit Hilfe von logarithmisch-linearen Modellen geprüft werden kann. Das geschilderte Verfahren kann als Alternative zur Berechnung eines standardisierten Chi-Quadrat-Wertes (ARMITAGE 1966, KOLLER 1974) und als Alternative zum MANTEL-HAENSZEL (1959) Verfahren angesehen werden. Die beiden letztgenannten Methoden wurden ursprünglich vorgeschlagen, um mehrere Vierfeldertafeln zusammenzufassen: bei nur der Tendenz nach vorhandenen ähnlichen Assoziationen in den einzelnen Teiltafeln hofft man, durch Zusammenfassen einen Zusammenhangsnachweis erbringen zu können. Beide Verfahren setzen daher implizit gleichartige Assoziationen in den Teiltafeln voraus und entsprechen inhaltlich der Prüfung der partiellen Assoziation. Dies bedeutet, daß sie sich im allgemeinen nicht dazu eignen, die Auswirkungen von Hintergrundfaktoren zu prüfen.

Der wichtigste Vorzug des hier beschriebenen Verfahrens ist, daß nur dann die partielle Assoziation geprüft wird, falls die Annahme des gleichartigen Zusammenhangs in allen Teiltafeln mit den Beobachtungen zu vereinbaren ist. Die dabei verwendete Definition der Gleichartigkeit ist nicht auf Vierfeldertafeln beschränkt; sie basiert auf Kreuzproduktverhältnissen, die ihrerseits als relative Risiken interpretiert werden können. Als Nachteile sind zu nennen, daß die Berechnungen nur noch schlecht von Hand durchgeführt werden können und daß der Begriff des relativen Risikos nicht unmittelbar anschaulich ist (vergleiche auch BERKSON 1958).

Obwohl wir das Prüfen eines Hintergrundfaktors als statistisches Testverfahren dargestellt haben, ist das Verfahren auch bei der Suche nach wichtigen Hintergrundfaktoren einsetzbar. In diesem Fall sind die Prüfgrößen und die zugehörigen Quantilwerte als deskriptive Maße anzusehen, mit deren Hilfe die relative Bedeutung verschiedener potentieller Hintergrundfaktoren bewertet werden kann.

Wir danken Frau H. Nitsch für die technischen Hilfen, Herrn O. Pietschmann für das Zeichnen der Abbildungen und Frau H. Bianco für das maschinengeschriebene Manuskript.

Literaturverzeichnis

- ARMITAGE, P. (1966): The chi-square test for heterogeneity of proportions after adjustment for stratification. *J.Roy.Statist.Soc.* B28, 159-169
- BARTLETT, M.S. (1935): Contingency table interactions. *J.Roy.Statist.Soc.Suppl.* 2, 248-252
- BERKSON, J. (1958): Smoking and lung cancer: Some observations on two recent reports. *J.Amer.Statist.Assoc.* 53, 28-38
- BISHOP, Y.M.M., HOLLAND, P.W., FIENBERG, S.E. (1975): Discrete multivariate analysis. Theory and practice, M.I.T. Press, Cambridge
- BIRCH, M.W. (1963): Maximum likelihood in threeway contingency tables. *J.Roy.Statist.Soc.* B25, 220-233
- BROWN, D.T. (1959): A note on approximations to discrete probability distributions. *Information and Control* 2, 386-392
- COCHRAN, W.G. (1965): The planning of observational studies of human populations. *J.Roy.Statist.Soc.* A128, 234-265
- DARROCH, J.N. (1962): Interactions in multi-factor contingency tables. *J.Roy.Statist.Soc.* B24, 251-263
- DARROCH, J.N. (1974): Multiplicative and additive interaction in contingency tables. *Biometrika* 61, 207-214
- FLEISS, J. (1973): Statistical methods for rates and proportions. Wiley, New York
- GOODMAN, L.A. (1970): The multivariate analysis of qualitative data: interactions among multiple classifications. *J.Amer.Statist.Assoc.* 65, 225-256
- KOLLER, S. (1964): Systematik statistischer Schlußfehler. *Meth.Inform.Med.* 3, 113-117
- KOLLER, S. (1974): Schwangerschaftsverlauf und Kindesentwicklung. Unveröffentlicher Zwischenbericht. *Inst.f.Med.Statistik, Mainz*
- LIENERT, G.A. (1971): Die Konfigurationsfrequenzanalyse I. *Z.klin.Psych.Psychother.* 19, 99-115
- McKINLAY, S.M. (1975): The effect of bias on estimators of relative risk for pair-matched and stratified samples. *J.Amer.Statist.Assoc.* 70, 859-864

- ROY, S.N., KASTENBAUM, M.A. (1956): On the hypothesis of "no interaction" in a multi-way contingency table.
Ann.Math.Statist. 27, 749-757
- WERMUTH, N. (1976): Analogies between multiplicative models in contingency tables and covariance selection.
Biometrics 32, 95-108

Anschrift der Verfasser:

Dr. Nanny Wermuth
Byung Kuk Yun
Herbert Gönner
Institut für Med. Statistik
und Dokumentation
der Universität Mainz
Langenbeckstr. 1
6500 Mainz