

## **Simultantestprozedur für globale Nullhypothesen bei beliebiger Abhängigkeitsstruktur der Einzeltests**

Lindner, Klaus

*Publication date:*  
2001

*Document Version*  
Verlags-PDF (auch: Version of Record)

[Link to publication](#)

*Citation for pulished version (APA):*

Lindner, K. (2001). *Simultantestprozedur für globale Nullhypothesen bei beliebiger Abhängigkeitsstruktur der Einzeltests*. (Final; Band 11, Nr. 3). Universität Lüneburg.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

University  
of Applied Sciences

Fachhochschule  
Nordostniedersachsen



Lüneburg  
Buxtehude  
Suderburg

Arbeitsberichte aus der Informatik



## Simultantestprozedur für globale Nullhypothesen bei beliebiger Abhängigkeitsstruktur der Einzeltests

K. Lindner

11. Jahrgang, Heft 3, September 2001, ISSN 0939-8821

Technical Reports and Working Papers

Hrsg: Hinrich E. G. Bonin

Volgershall 1, D-21339 Lüneburg

Phone: xx49.4131.677175 Fax: xx49.4131.677140

## Simultantestprozedur für globale Nullhypothesen - bei beliebiger Abhängigkeitsstruktur der Einzeltests.

**Zusammenfassung:** Viele Probleme der Angewandten Statistik zeichnen sich dadurch aus, daß die zu überprüfende „globale“ Nullhypothese als Schnittmenge von  $n$  Einzelhypothesen aufgefaßt werden kann.

Aus unterschiedlichen Gründen (z.B. um differenzierterer Aussagen willen) möchte man statt der globalen Hypothese die  $n$  Einzelhypothesen testen. Entscheidend ist dabei, das Gesamtsignifikanzniveau trotz beliebiger Abhängigkeiten der Einzeltests unter Kontrolle zu behalten.

In der Literatur sind dazu verschiedene Methoden vorgeschlagen worden. Eine gute Übersicht über die klassischen Methoden findet sich bei MILLER [1981]. In der neueren Literatur sind eine Reihe von Varianten der BONFERRONI-Methode vorgeschlagen worden, z.B. bei KRAUTH/LIENERT [1973], RÜGER [1978], LINDNER [1979], HOMMEL [1983].

Eine übergreifende Zusammenfassung und Verallgemeinerung der letztgenannten Methoden bildet den Gegenstand der vorliegenden Arbeit.

Eine Anwendung dieser neuen Methode zur Auswertung von Kontingenztafeln wird diskutiert.

### 1. Einführung

Bei vielen Problemen der anwendungsorientierten Statistik steht man häufig vor dem folgenden Problem:

Eine zu testende - globale - Nullhypothese setzt sich aus einer Reihe von Einzelhypothesen zusammen.

Typisch für diese Situationen ist es meistens, daß sowohl ein Test der globalen Hypothese als auch Tests für die Einzelhypothesen von praktischem Interesse sind.

Dabei möchte man die globale Nullhypothese testen, um zu prüfen, ob das Datenmaterial überhaupt irgendwelche interessanten Teilaussagen vermuten läßt, die dann im einzelnen noch aufzusuchen wären; während man die Einzelhypothesen deshalb überprüfen will, weil die Einzeltests zu wesentlich differenzierteren Aussagen führen als der Globaltest.

Ein typisches Beispiel für die beschriebene Situation ist die Auswertung von zweidimensionalen Kontingenztafeln, den sogenannten Kreuztabellen. Möchte man etwa – wie in einer Untersuchung von WEIHE et al. [1987] – feststellen, ob Zusammenhänge zwischen der Höhe des Anfangsjahresgehaltes von Hochschulabsolventen und dem jeweiligen Unternehmenstyp bestehen, so befragt man Absolventen nach den Ausprägungen dieser beiden Merkmale, trägt dann die Ergebnisse in einer Kreuztabelle ein und untersucht, ob Abhängigkeiten zwischen bestimmten Merkmalsausprägungen bestehen.

Das statistische Standardinstrument für diese Situation ist der  $\chi^2$  - Test der Kontingenztafel auf Unabhängigkeit zu einem vorgegebenen Signifikanzniveau.

Hierzu sind aber sofort zwei wichtige Anmerkungen zu machen.

Erstens kann man diesen globalen  $\chi^2$  - Test häufig deshalb überhaupt nicht durchführen, weil die Voraussetzungen dafür nicht erfüllt sind. Es müssen nämlich, da die  $\chi^2$  - Verteilungsaussage asymptotischen Charakter hat, die unter der Nullhypothese der Unabhängigkeit erwarteten Häufigkeiten in jeder Zelle hinreichend groß sein, das heißt im wesentlichen, der Stichprobenumfang muß hinreichend groß sein. Ist dies nicht der Fall, muß auf die Durchführung des globalen  $\chi^2$  - Tests verzichtet werden. Es erhebt sich die Frage, wie man dann bei der Auswertung fortfahren soll, worauf wir gleich zurückkommen werden.

Zweitens würde der globale  $\chi^2$  - Test, wenn der Stichprobenumfang seine Durchführung überhaupt gestattet, im Falle der Ablehnung der Nullhypothese nur eine sehr allgemeine Aussage liefern, nämlich die, daß irgendwelche Abhängigkeiten zwischen den beiden Merkmalen vorhanden sind, nicht aber, bei welchen Ausprägungskombinationen der beiden Merkmale diese Abhängigkeiten auftreten. Auf unser obiges Beispiel bezogen, würde man dadurch eventuell zeigen können, daß es zwar Abhängigkeiten zwischen Unternehmenstyp und Jahresanfangsgehältern gibt, nicht aber, von welcher Art diese Abhängigkeiten sind. Genau letzteres will man aber eigentlich erfahren.

Beide soeben diskutierten Probleme bei der Auswertung von Kontingenztafeln sind für den Anwender von entscheidender Bedeutung.

Ähnliche Situationen treten in nahezu allen Bereichen der Angewandten Statistik auf, z.B. bei der Analyse von Zeitreihen, bei Konfigurationsfrequenzanalysen oder bei multifaktoriellen Varianzanalysen.

Eine Methode, mit den oben skizzierten Problemen qualifiziert umzugehen, soll hier am Beispiel der Kontingenztafeln demonstriert werden.

Nehmen wir an, es liege eine 5 x 6 - Kontingenztafel vor; das Merkmal A (Unternehmenstyp) habe 5 Ausprägungen (z.B. sehr kleine, kleine, mittlere, große, sehr große Unternehmen), das Merkmal B (Jahresanfangsgehalt) sei in 6 Ausprägungen aufgegliedert.

Die Kontingenztafel besteht dann aus 30 Zellen, und die globale Unabhängigkeitshypothese besagt, daß in allen Zellen Unabhängigkeit vorliegt (anschaulich gesprochen, daß in keiner der Zellen die Besetzungszahl zu über- oder unterproportionaler Realisierung tendiert).

In beiden oben angesprochenen Fällen (sowohl bei Nicht-Durchführbarkeit des globalen  $\chi^2$  - Tests wegen zu kleiner Besetzungszahlen, als auch bei der unbedingten Forderung nach Aussagen über die Unabhängigkeit pro Zelle), bietet sich nun folgende Vorgehensweise an, wobei anzumerken ist, daß der Grundgedanke dieser Methode nicht etwa auf Kontingenztafeln beschränkt ist, sondern sich auf alle statistischen Probleme übertragen läßt, bei denen eine Globalhypothese sich aus einer Reihe von Einzelhypothesen zusammensetzt.

Bei der 5 x 6 - Kontingenztafel kann man die globale Nullhypothese der Unabhängigkeit auffassen als das gemeinsame Auftreten von 30 Nullhypothesen der Unabhängigkeit pro Zelle.

Für jede einzelne Zelle aber kann man die Hypothese der Unabhängigkeit immer testen - auch im Falle eines kleinen Stichprobenumfanges. Hierfür steht nämlich in jedem Falle der exakte FISHER-Test zur Verfügung. Aus der 5 x 6 Felder -Tafel wird unter Berücksichtigung der jeweils interessierenden Zelle und unter geeigneter Zusammenfassung der restlichen Zellen eine 2 x 2 - Felder - Tafel gemacht, für die mit Hilfe der hypergeometrischen Verteilung dann der exakte Test von FISHER durchgeführt werden kann, der im übrigen für den 2 x 2 -Felder Test in seiner randomisierten Fassung (TOCHER [1950]) sogar gleichmäßig bester unverfälschter (uniformly most powerful unbiased) Test ist, also für den Unabhängigkeitstest pro Zelle das schärfste statistische Instrument überhaupt darstellt, WITTING [1966].

Man führt also 30 exakte FISHER - Tests durch und erhält dadurch pro Zelle eine Aussage über die Unabhängigkeit bzw. deren Verletzung.

Auf den ersten Blick scheinen nun möglicherweise mit der Durchführung dieser 30 Einzeltests beide der oben angesprochenen Probleme gelöst zu sein: Erstens ist dieses Vorgehen auch bei kleinem Stichprobenumfang möglich, und zweitens erhält man differenziertere Aussagen als beim globalen  $\chi^2$  - Test.

Leider sind aber die Probleme damit noch keineswegs befriedigend gelöst, denn es erhebt sich die Frage nach dem einzuhaltenden Signifikanzniveau. Dieses wollen wir in der vorliegenden Arbeit immer mit  $\gamma$  bezeichnen. Nehmen wir an, jeden der Einzeltests führen wir zum Niveau  $\alpha \leq \gamma$  durch, dann erhebt sich die Frage, wie groß die Wahrscheinlichkeit ist, daß mindestens einer der Einzeltests fälschlicherweise zur Ablehnung der (Einzel)-Nullhypothese der Unabhängigkeit führt.

Diese Wahrscheinlichkeit sollte einerseits natürlich kleiner als  $\gamma$  sein, andererseits steigt sie aber mit wachsender Anzahl der durchzuführenden Einzeltests stark an.

Anmerkung: Die bei diesen Tests auftretenden Abhängigkeiten sind überaus kompliziert und gerade deshalb typisch für die diskutierte Vorgehensweise: In der nachfolgenden allgemeinen Betrachtung werden generell beliebige, insbesondere auch im einzelnen nicht bekannte Abhängigkeiten zwischen den Einzeltests zugelassen.

Einen ersten Anhaltspunkt, den Zusammenhang zwischen  $\alpha$  und  $\gamma$  betreffend, erhält man, wenn man die (hier sicher nicht gegebene) Unabhängigkeit der n Einzeltests unterstellt.

Dann nämlich wäre (vgl. etwa LINDNER [1979]) bei vorgegebenem Einzelsignifikanzniveau  $\alpha$  das Gesamtsignifikanzniveau

$$\gamma = 1 - (1 - \alpha)^n .$$

In unserem Beispiel würde sich mit  $\alpha = 0.05$  und  $n = 30$  ein Gesamtsignifikanzniveau von  $\gamma = 0.785$  ergeben, was sicherlich als unerträglich hoch anzusehen wäre.

Es würde immerhin bedeuten, daß bei Zutreffen der globalen Nullhypothese mit knapp 80 % - iger Wahrscheinlichkeit mindestens ein Fehlurteil zustande käme.

Um diesen nicht tolerierbaren Anstieg des Gesamtsignifikanzniveaus zu verhindern, müßte man das Signifikanzniveau der Einzeltests verringern; im Spezialfall der Unabhängigkeit der Einzeltests müßte bei vorgegebenem Gesamtsignifikanzniveau  $\gamma$  das Einzelsignifikanzniveau gemäß

$$(1) \quad \alpha^* = 1 - \sqrt[n]{1 - \gamma}$$

gewählt werden.

Allgemein nennt man diesen Vorgang der Reduzierung des Einzelsignifikanzniveaus zur Einhaltung des Gesamtsignifikanzniveaus eine  $\alpha$  - Justierung .

Eine Justierung gemäß (1) kommt aber - wie gesagt - nur bei unabhängigen Tests in Betracht.

Wir wollen hier aber beliebige Abhängigkeiten zulassen, was natürlich zu noch strengerer  $\alpha$  - Justierung führen muß.

## 2. Bekannte Methoden zur Einhaltung des Gesamtsignifikanzniveaus

### 2.1 Ein einfaches Beispiel

Um die nachfolgend dargestellten Methoden einfach illustrieren zu können, betrachten wir zunächst eine besonders einfach strukturierte Fragestellung.

Ein Marketingexperte befaßt sich mit dem folgenden Problem:

Ein neues Getränk soll auf den Markt gebracht werden. Entschieden werden soll nun noch die Farbe der Dose (Verpackung) - zur Diskussion stehen die Farben blau, grün, weiß, braun, gelb und rot, die als Farbe1, Farbe2, usw. numeriert sind.

$N = 24$  Testpersonen stehen zur Verfügung. Jede Testperson soll den Geschmack des Getränks beurteilen. Dazu wird ihr dasselbe Getränk in jeweils 6 unterschiedlich gefärbten Dosen zur Geschmacksbeurteilung dargeboten. Sie soll angeben, welche der 6 Proben sie geschmacklich favorisiert.

Der Marketingexperte soll anhand der Testergebnisse die - etwas naive, hier nur zu Demonstrationszwecken aufgeworfene - Frage beantworten, ob der Geschmack unabhängig von der Dosenfarbe beurteilt wird (globale Fragestellung), wobei er natürlich gleichzeitig an der weitergehenden Fragestellung interessiert ist, welche der Farben denn nun favorisiert wird (differenziertere Fragestellung).

Wenn das Geschmacksurteil von der Farbe der Dose nicht abhängig ist, verteilen sich die favorisierten Proben zufällig auf die 6 Farben, und zwar jeweils mit den Wahrscheinlichkeiten

$$p_j = \frac{1}{6} \quad , \quad j = 1, 2, \dots, 6 \quad .$$

Mit  $X_1, X_2, \dots, X_6$  seien die Zufallsvariablen bezeichnet, die angeben, wieviel Testpersonen jeweils die Farbe 1, Farbe 2, ..., Farbe 6 favorisieren.

Die Realisierungen der  $X_1, X_2, \dots, X_6$  unterliegen dann bekanntlich gemeinsam einer Multinomialverteilung

$$M(N; p_1, p_2, \dots, p_6) .$$

Die (globale) Nullhypothese  $H_0$  der Unabhängigkeit des Geschmacksurteils von der Farbe der Dose lautet dann ausführlich :

$$H_0: \begin{cases} p_1 = p_1^{(0)} = \frac{1}{6} \\ p_2 = p_2^{(0)} = \frac{1}{6} \\ \vdots \\ p_6 = p_6^{(0)} = \frac{1}{6} \end{cases}$$

Ist der Stichprobenumfang groß genug, kann ein Globaltest für  $H_0$  durchgeführt werden, und zwar mit Hilfe des üblichen  $\chi^2$  - Anpassungstests.

Die in der einschlägigen Literatur (z.B. COCHRAN [1954]) angegebene Bedingung zur Anwendung dieses Tests lautet, daß in mindestens 80% der Ausprägungen (hier also in mindestens 5 von den 6 Ausprägungen) die erwarteten Häufigkeiten  $N * p_i^{(0)} \geq 5$  sein sollen.

Wenn aber z.B.  $N = 24$  ist, kann diese Bedingung nie erfüllt werden, womit der klassische Globaltest für dieses Problem ausfällt.

Die Zerlegung des Problems in 6 Einzeltests liegt auf der Hand.

Die einzelnen Variablen  $X_1, X_2, \dots, X_6$  sind offensichtlich jeweils binomialverteilt gemäß

$$Bi(N; p_j) \quad \text{mit } j = 1, 2, \dots, 6.$$

Mit

$$H_0^{(j)} : p_j^{(0)} = \frac{1}{6} \quad \text{für } j = 1, 2, \dots, 6$$

ist  $H_0$  offenbar äquivalent mit dem Zutreffen aller sechs Einzelnullhypothesen  $H_0^{(j)}$  :

$$H_0 = H_0^{(1)} \wedge H_0^{(2)} \wedge \dots \wedge H_0^{(6)} ,$$

was kurz als

$$H_0 = \bigcap_{j=1}^6 H_0^{(j)}$$

geschrieben werden kann, da die Hypothesen bekanntlich auch als Teilmengen eines Gesamt-parameterraumes aufgefaßt werden können.

Die sechs Einzeltests sind nun die üblichen Binomialtests, mit denen nach Vorgabe des Einzelsignifikanzniveaus  $\alpha$  anhand der Realisierungen der 6 Variablen  $X_1, X_2, \dots, X_6$

über die Einzel-Nullhypothesen  $H_0^{(j)}$  entschieden wird.

Das einzuhaltende Gesamtsignifikanzniveau  $\gamma$  geben wir uns vor.

Die entscheidende Frage lautet:

Wie muß das Signifikanzniveau  $\alpha$  für die Einzeltests gewählt werden, und wie hat nach Durchführung der Einzeltests die Entscheidungsvorschrift über die globale  $H_0$  zu lauten, damit insgesamt  $\gamma$  eingehalten wird.

Getestet werden soll also die globale Nullhypothese

$$H_0 = \bigcap_{j=1}^n H_0^{(j)},$$

indem die n Einzel-Nullhypothesen  $H_0^{(j)}$ ,  $j = 1, 2, \dots, n$ , in n Einzeltests geprüft werden.

Zur Wahl der Entscheidungsvorschrift über die globale  $H_0$  in Abhängigkeit der Ausgänge der n Einzeltests und - gekoppelt damit - zur Wahl des Einzelsignifikanzniveaus  $\alpha$  stehen bereits einige Verfahren zur Verfügung. Diese sollen in den nächsten 3 Teilabschnitten vorgestellt werden.

## 2.2 Die BONFERRONI - Methode

Das älteste und bisher in der Anwendung am häufigsten anzutreffende Verfahren ist die BONFERRONI - Methode, (vgl. dazu etwa MILLER [1981] oder den sehr anwendungsorientierten Aufsatz von KRAUTH/LIENERT [1973]).

Bei dieser Methode wählt man das justierte Einzelsignifikanzniveau gemäß

$$(2) \quad \alpha = \frac{\gamma}{n},$$

wobei  $\gamma$  wieder das Gesamtsignifikanzniveau und n die Anzahl der durchzuführenden Einzeltests ist;

und die Entscheidungsvorschrift für die globale  $H_0$  lautet:

$$(3) \quad \begin{cases} \text{Lehne die Globalhypothese } H_0 \text{ genau dann ab,} \\ \text{wenn wenigstens einer der } n \text{ Einzeltests zur Ablehnung der} \\ \text{entsprechenden Einzel - Nullhypothese führt.} \end{cases}$$



Man kann leicht einsehen, weshalb dieses Verfahren immer die Einhaltung von  $\gamma$  garantiert.

Es sei  $P$  eine beliebige aus  $H_0$  stammende Wahrscheinlichkeitsverteilung über dem Stichprobenraum.

(Im Beispiel aus 2.1 gibt es sowieso nur eine Verteilung, nämlich die genannte Multinomialverteilung).

Wegen (2) gilt dann für die Ablehnungsbereiche  $A_j$  ( $j = 1, 2, \dots, n$ ) der Einzeltests

$$(4) \quad P(A_j) \leq \alpha = \frac{\gamma}{n} .$$

Der Gesamtablehnungsbereich  $A$  der globalen  $H_0$  besteht wegen (3) offenbar aus der Vereinigungsmenge der Einzelablehnungsbereiche:

$$A = \bigcup_{j=1}^n A_j$$

Nach der leicht mit vollständiger Induktion beweisbaren BONFERRONI - Ungleichung (daher der Name des Verfahrens) gilt dann

$$P(A) = P\left(\bigcup_{j=1}^n A_j\right) \leq \sum_{j=1}^n P(A_j)$$

und damit wegen (4)

$$P(A) \leq n * \alpha = \gamma .$$

Also:

$$(5) \quad \left\{ \begin{array}{l} \text{Wird nach } \alpha - \text{Justierung gemäß (2) die globale } H_0 \text{ aufgrund} \\ \text{der Entscheidungsvorschrift (3) abgelehnt, so wird dabei das} \\ \text{vorgegebene Gesamtsignifikanzniveau } \gamma \text{ eingehalten.} \end{array} \right.$$

Wird nach  $\alpha$  - Justierung gemäß (2) die globale  $H_0$  aufgrund der Entscheidungsvorschrift (3) abgelehnt, so wird dabei einerseits das vorgeschriebene Gesamtsignifikanzniveau  $\gamma$  eingehalten, und andererseits erhält der Anwender durch die Einzeltestentscheidungen differenzierte Informationen darüber, an welchen Stellen denn die globale Hypothese nicht zutrifft.

Wir wollen uns diese Prozedur nun an unserem kleinen Beispiel aus 2.1 ansehen.

Ein Gesamtsignifikanzniveau von  $\gamma = 0.10$  geben wir uns vor.

$n = 6$  zweiseitige Binomialtests für die jeweiligen Nullhypothesen

$$H_0^{(j)} : p_j = \frac{1}{6} , \quad j = 1, 2, \dots, 6$$

sollen durchgeführt werden.

Gemäß (2) haben wir in den Einzeltests mit einem justierten  $\alpha = \frac{\gamma}{6} = 0,01\overline{66}$  zu arbeiten, welches sich wegen der Zweiseitigkeit (die Auftretenshäufigkeiten können ja zu klein oder zu groß sein) je zur Hälfte - also  $\frac{\alpha}{2} = 0,008\overline{33}$  - auf beide Seiten verteilt.

Das heißt also:  $H_0^{(j)}$  wird im Einzeltest dann abgelehnt, wenn für die aufgetretene Realisierung  $x_j$  der Variablen  $X_j$  entweder

$$P(X_j \leq x_j) \leq 0,008\overline{33} \quad \text{oder}$$

$$P(X_j \geq x_j) \leq 0,008\overline{33}$$

gilt.

Die Wahrscheinlichkeiten für  $P(X_j \leq x_j)$  bzw. für  $P(X_j \geq x_j)$  ergeben sich bei

$N = 24$  nach:

$$P(X_j = x_j) = \binom{24}{x_j} * \left(\frac{1}{6}\right)^{x_j} * \left(\frac{5}{6}\right)^{(24-x_j)} \quad \text{für } j=1,2,\dots,6 \quad \text{und } x_j = 0,1,2,\dots,24$$

Man kann sie für die hier interessierenden Werte der folgenden Tabelle entnehmen.

Tabelle 1: Wahrscheinlichkeiten der  $Bi\left(N = 24; \frac{1}{6}\right)$ -Verteilung:

Auftretens- häufigkeit $x_j$	$P(X_j = x_j)$	$P(X_j \leq x_j)$	$P(X_j \geq x_j)$
0	0,0126	0,0126	1
1	0,0604	0,0730	.
2	0,1389	0,2118	.
3	0,2037	0,4155	.
4	0,2139	.	.
5	0,1711	.	.
6	0,1084	.	0,1995
7	0,0557	.	0,0912
8	0,0237	.	0,0354
9	0,0084	.	0,0118
10	0,0025	.	0,0033
11	0,0006	.	0,0008
$\vdots$	$\vdots$	$\vdots$	$\vdots$
24	$\approx 0$	1	$\approx 0$

Nehmen wir nun an, bei den  $N = 24$  Testpersonen hätten sich die folgenden Auftretenshäufigkeiten  $x_j$  der favorisierten Farben ergeben :

Datensatz (A):

Ausprägung (Farbe) j	Farbe 1 (blau)	Farbe 2 (grün)	Farbe 3 (weiß)	Farbe 4 (braun)	Farbe 5 (gelb)	Farbe 6 (rot)
Auftretenshäufigkeit $x_j$	1	2	4	1	6	10
Restwahrscheinlichkeit nach Tabelle 1	0,0730	0,2118	.	0,0730	0,1995	0,0033

Die unter der Nullhypothese zu erwartenden Häufigkeiten wären  $4 = 24/6$  . Wie wahrscheinlich bzw. unwahrscheinlich die jeweiligen Abweichungen der tatsächlich beobachteten Häufigkeiten  $x_j$  von dieser erwarteten Häufigkeit sind, geben die aus Tabelle 1 abzulesenden Restwahrscheinlichkeiten  $P(X_j \leq x_j)$  bzw.  $P(X_j \geq x_j)$  an, und zwar je nachdem, ob die jeweilige Auftretenshäufigkeit  $x_j$  größer oder kleiner als die unter der Nullhypothese zu erwartende Häufigkeit 4 ist.

Führt man nun die 6 Einzeltests für den Datensatz (A) durch, so stellt man fest, daß nur die Ausprägung  $x_6 = 10$  der Variablen  $X_6$  mit einer Restwahrscheinlichkeit von  $P(X_6 \geq 10) = 0.0033$

unter den BONFERRONI - justierten Wert  $\frac{\alpha}{2} = 0,00833$  fällt.

Die Ausprägungen für die nächst extremeren Werte  $x_1 = x_5 = 1$  haben jeweils die Restwahrscheinlichkeiten  $0.01258 > 0,00833$ .

Bei den 6 Einzeltests wird also die Nullhypothese  $H_0^{(6)}$  abgelehnt, während alle anderen

$H_0^{(j)}$  ( $j = 1,2,3,4,5$ ) beibehalten werden.

Das heißt zusammengefaßt:

Beim Datensatz (A) wird die Einzelhypothese  $H_0^{(6)}$  abgelehnt, und deshalb wird nach der BONFERRONI-Entscheidungsvorschrift (3) auch die globale Nullhypothese  $H_0$  auf dem 10%-Niveau ( $\gamma = 0.1$ ) abgelehnt.

Der Marketingexperte hat nun aufgrund dieses Experiments auf dem 10%-Niveau die Aussage abgesichert, daß sich die Geschmacksbeurteilung nicht unabhängig von der Verpackungsfarbe ergibt.

Neben dieser - auf dem Niveau  $\gamma$  abgesicherten - Aussage erhält der Marketingexperte die zusätzliche Information, daß es wahrscheinlich die Farbe 6 (rot) ist, welche die Geschmacksbeurteilung positiv beeinflusst.

Anmerkung: Dieses Beispiel ist natürlich nur seiner Einfachheit halber gewählt worden, um das Prinzip ohne viel Rechnung deutlich werden zu lassen; die praktischen Anwendungsmöglichkeiten dieses Prinzips liegen, wie oben bereits erwähnt, in fast allen Bereichen der Angewandten Statistik. Aus Demonstrationsgründen bleiben wir weiter bei diesem einfachen Beispiel.

### 2.3 Die RÜGER-Methode

Nehmen wir nun einmal an, der Marketingexperte hätte nicht den Datensatz (A) erhalten, sondern der Versuch hätten die folgenden Daten ergeben:

Datensatz (B):

Ausprägung (Farbe) j	Farbe 1 (blau)	Farbe 2 (grün)	Farbe 3 (weiß)	Farbe 4 (braun)	Farbe 5 (gelb)	Farbe 6 (rot)
Auftretenshäufigkeit $x_j$	1	2	2	1	9	9
Restwahrscheinlichkeit nach Tabelle 1	0,0730	0,2118	0,2118	0,0730	0,0118	0,0118

Geht man auch hier wieder nach der BONFERRONI - Methode vor, so stellt man sofort fest, daß keine der Restwahrscheinlichkeiten unter den BONFERRONI - justierten Wert

$$\frac{\alpha}{2} = 0,00833$$

fällt.

Nach der BONFERRONI - Methode würde also keine der Einzelhypothesen und damit natürlich auch nicht die globale Nullhypothese  $H_0$  verworfen werden.

Der besseren Vergleichbarkeit wegen stellen wir einmal die beiden Datensätze direkt nebeneinander:

Tabelle 2 (Datensätze (A und B)):

Ausprägung (Farbe) j	Farbe 1 (blau)	Farbe 2 (grün)	Farbe 3 (weiß)	Farbe 4 (braun)	Farbe 5 (gelb)	Farbe 6 (rot)
Auftretenshäufigkeit $x_j$ bei Datensatz (A)	1	2	4	1	6	10
Auftretenshäufigkeit $x_j$ bei Datensatz (B)	1	2	2	1	9	9

Rein von der Intuition her würde man sagen, daß die Abweichungen von der in allen 6 Zellen jeweils erwarteten Häufigkeit 4 bei Datensatz (B) insgesamt stärker ist als bei (A).

Woran liegt es nun, daß die BONFERRONI - Methode bei Datensatz (B) die Nullhypothese nicht ablehnt - wohl aber bei Datensatz (A) ?

Es liegt an der extremen  $\alpha$  - Justierung und dem damit verbundenen Powerverlust im Einzeltest.

Anmerkung: Bei der Entscheidungsvorschrift (3) darf aber die  $\alpha$  - Justierung nicht schwächer ausfallen. Dies kann man leicht einsehen; wenn nämlich die Abhängigkeitsstruktur der Einzeltests so beschaffen ist, daß die jeweiligen Ablehnungsbereiche paarweise disjunkt sind, tritt in der BONFERRONI - Ungleichung das Gleichheitszeichen ein, und damit wird das vorgegebene Gesamtsignifikanzniveau  $\gamma$  bei der BONFERRONI - Methode voll ausgeschöpft.

Um nun den starken Powerverlust im Einzeltest bei der BONFERRONI - Methode zu vermeiden, hat RÜGER [1978] folgende simultane Testentscheidungsmethode vorgeschlagen.

Das Gesamtsignifikanzniveau  $\gamma$  sei wieder vorgegeben.

Man wähle vor (!) der Datenerhebung einen festen Wert

$$k \in \{1, 2, \dots, n\}$$

und justiere dann das Signifikanzniveau der Einzeltests folgendermaßen:

$$(6) \quad \alpha = \frac{k * \gamma}{n} .$$

Die Entscheidungsvorschrift für die globale  $H_0$  soll dann lauten:

$$(7) \quad \left\{ \begin{array}{l} \text{Lehne die Globalhypothese } H_0 \text{ genau dann ab,} \\ \text{wenn wenigstens } k \text{ der } n \text{ Einzeltests zur Ablehnung der} \\ \text{entsprechenden Einzel - Nullhypothesen führen.} \end{array} \right.$$

RÜGER [1978] hat - mit sehr großem beweistechnischem Aufwand - nachgewiesen, daß die  $\alpha$  - Justierung (6) - gekoppelt mit der Entscheidungsvorschrift (7) - einerseits das globale Signifikanzniveau  $\gamma$  einhält und andererseits in dem Sinne scharf ist, daß es Abhängigkeitsstrukturen der Einzeltests gibt, bei denen das Gesamtsignifikanzniveau voll ausgeschöpft wird.

Einfachere Beweise der RÜGERschen Aussagen findet man bei LINDNER [1979] und bei MORGENSTERN [1980].

Wenden wir nun die RÜGER - Methode auch wieder mit  $\gamma = 0,1$  auf unser einfaches Beispiel (in den zwei Varianten der Datensätze (A) und (B)) an.

Wir nehmen an, wir hätten (vor der Datenerhebung)  $k = 2$  gewählt.

Das Signifikanzniveau der Einzeltests wäre dann nach (6) als

$$\alpha = \frac{2 * 0,1}{6} = 0,0\overline{33}$$

zu berechnen.

Wegen der Zweiseitigkeit der Binomialtests muß im Einzeltest die Restwahrscheinlichkeit den Wert

$$\frac{\alpha}{2} = 0,0166$$

unterschreiten, um eine Ablehnung der Einzelnullhypothese  $H_0^{(j)}$  zu erreichen.

Um nach RÜGERS Entscheidungsvorschrift (7) auf dem vorgegebenen Niveau  $\gamma$  die globale

Nullhypothese  $H_0$  ablehnen zu können, müssen jetzt wegen der Wahl von  $k = 2$ , mindestens zwei der Einzeltests zur Ablehnung führen.

Wie man feststellt, ist dies bei dem Datensatz (B) der Fall, die Farben "gelb" und "rot" wurden jeweils 9- statt der erwarteten 4-mal favorisiert.

Da aber

$$P(X_j \geq 9) = 0,0118 \leq 0,0166 = \frac{\alpha}{2} \quad \text{für } j = 5, 6$$

gilt, ist für beide Farben jeweils die Nullhypothese

$$H_0^{(j)}: p_j = \frac{1}{6} \quad (j = 5, 6)$$

abzulehnen.

Bei Datensatz (B) ist also die globale Nullhypothese  $H_0$  wegen (7) abzulehnen.

Anders sieht es beim Datensatz (A) aus. Dort ist nur die 10-malige Favorisierung von "rot" so unwahrscheinlich, daß

$$P(X_6 \geq 10) = 0,0033 \leq 0,0166 = \frac{\alpha}{2}$$

gilt, alle anderen Restwahrscheinlichkeiten bleiben über dieser Schranke.

Daher kann im Datensatz (A) gemäß (7) die globale  $H_0$  nicht abgelehnt werden - eben weil  $k = 2$  gewählt worden ist.

Vor- und Nachteile der RÜGER - Methode liegen damit auf der Hand:

Der Vorteil gegenüber der BONFERRONI - Methode besteht natürlich in der nicht so extremen  $\alpha$  - Justierung, was den Powerverlust im Einzeltest weniger gravierend macht.

Der Nachteil liegt in der relativ willkürlichen Festlegung des Wertes  $k$ ; wobei noch einmal zu betonen ist, daß  $k$  vor der Datenerhebung festgelegt werden muß. Falls man  $k$  erst anhand der Daten festlegt, erhöht sich natürlich das Gesamtsignifikanzniveau weit über die vorgegebene Schranke  $\gamma$ .

RÜGERs Methode bietet sich also vorwiegend dann an, wenn man bereits vor der Datenerhebung eine grobe Vorstellung davon hat, an wieviel Stellen die globale Nullhypothese verletzt sein könnte.

Dementsprechend kann man dann  $k$  wählen. In vielen praktischen Fällen wird man sich mit  $k = 2$  begnügen, weil dort der Powergewinn im Einzeltest gegenüber der BONFERRONI - Methode schon sehr groß ist, das Risiko, wegen weniger als  $k$  Einzelablehnungen  $H_0$  beibehalten zu müssen, aber unter allen  $k > 1$  noch am kleinsten ist.

Praktisch kann aber auch  $k = 2$  schon zu groß sein, wie sich am Beispiel der Beurteilung des Datensatzes (A) gezeigt hat.

Erwünscht wäre unter praktischen Gesichtspunkten häufig ein "Mittel" zwischen der BONFERRONI-Methode und der RÜGER - Methode mit  $k = 2$ .

Die sehr allgemeine, im dritten Abschnitt der vorliegenden Arbeit bewiesene Ungleichung (21) wird uns ein solches Vorgehen ermöglichen. Zunächst wollen wir der Vollständigkeit halber aber noch die HOMMEL – Methode vorstellen.

#### 2.4 Die HOMMEL - Methode

Der Willkürlichkeit der Vorgabe von  $k$  bei der RÜGER - Methode, die deren Hauptnachteil darstellt, wird durch eine von HOMMEL [1983] vorgestellte Methode begegnet.

Grob gesagt, wird bei der HOMMEL-Methode abgeschätzt, wie stark sich das Gesamtsignifikanzniveau gegenüber dem vorgegebenen  $\gamma$  vergrößert, wenn man bei der RÜGER-Methode den Wert von  $k$  erst nach der Datenerhebung so wählt, daß man eben für  $k$  Einzeltests auch  $k$  Ablehnungen erhält.

Dementsprechend stärker werden die Einzelsignifikanzniveaus

$$(8) \quad \alpha_k = \frac{k * \gamma}{C(n) * n}$$

mit einer nur von  $n$  abhängigen Konstanten  $C(n)$  dann so justiert, daß insgesamt das vorgegebene Gesamtniveau  $\gamma$  wieder eingehalten wird.

Damit ist der Hauptnachteil der RÜGER - Methode, die mehr oder weniger willkürliche Vorgabe von  $k$ , behoben - allerdings um den Preis einer stärkeren  $\alpha$  - Justierung.

Entscheidend bei dieser Methode ist die Berechnung des Wertes  $C(n)$ .

Nach HOMMEL [1983] ist  $C(n)$  folgendermaßen zu wählen:

$$(9) \quad C(n) = \sum_{k=1}^n \frac{1}{k}$$

Nach HOMMEL ist nun folgendermaßen vorzugehen:

$$(10) \quad \left\{ \begin{array}{l} \text{Für jeden der } n \text{ Einzeltests berechne man nach (8) und (9) } n \\ \text{Einzelsignifikanzschranken } \alpha_1 \leq \alpha_k \leq \dots \leq \alpha_n \text{ und lehne die} \\ \text{Globalhypothese } H_0 \text{ genau dann ab, wenn ein } k \in \{1, 2, \dots, n\} \text{ so} \\ \text{existiert, daß wenigstens } k \text{ der Einzeltests die jeweilige Einzel-} \\ \text{nullhypothese } H_0^{(j)} \text{ auf dem Niveau } \alpha_k \text{ ablehnen.} \end{array} \right.$$

HOMMEL [1983] hat bewiesen, daß bei diesem Vorgehen einerseits  $\gamma$  eingehalten wird, und daß diese Justierung insofern scharf ist, als es Abhängigkeitsstrukturen der Einzeltests gibt, bei denen  $\gamma$  voll ausgeschöpft wird.

Wenden wir nun auch dieses Verfahren auf unser kleines Beispiel aus 2.1 an.

Dazu berechnen wir zunächst nach (8) und (9) die Einzelsignifikanzschranken  $\alpha_k$  für  $k = 1, 2, \dots, n$ .

Der Wert von  $C(n)$  ergibt sich zu

$$C(n) = \sum_{k=1}^6 \frac{1}{k} = 2,45.$$

Damit ergeben sich die Werte für

$$\alpha_k = \frac{k * 0,1}{2,45 * 6} \quad k = 1, 2, \dots, 6.$$

Diese Werte müssen wegen der Zweiseitigkeit des Binomialtests wieder halbiert, und die so

berechneten Werte  $\frac{\alpha_k}{2}$  müssen in aufsteigender Reihenfolge mit den der Größe nach geordneten

Restwahrscheinlichkeiten  $p^{(k)}$  ( $k=1, 2, \dots, n$ ) der Daten verglichen werden.

Die Entscheidungsvorschrift (10) läuft dann darauf hinaus, daß die Globalhypothese  $H_0$  genau dann abgelehnt wird, wenn für ein  $k \in \{1, 2, \dots, n\}$

$$p^{(k)} \leq \frac{\alpha_k}{2}$$

gilt.

Die so berechneten Werte  $\frac{\alpha_k}{2}$  der Einzeltests und die in aufsteigender Reihenfolge geordneten

Restwahrscheinlichkeiten  $p^{(k)}$  der Datensätze (A) und (B) sind in der folgenden Tabelle zusammengefaßt.



Tabelle 3:

k	$\frac{\alpha_k}{2}$	in aufsteigender Reihenfolge geordnete Restwahrscheinlichkeiten $p^{(k)}$ für:	
		Datensatz (A)	Datensatz (B)
1	0,0034	0,0033 (*)	0,0118
2	0,0068	0,0730	0,0118
3	0,0102	0,0730	0,0730
4	0,0136	0,1995	0,0730
5	0,0170	0,2118	0,2118
6	0,0204	.	0,2118

Nach der HOMMEL-Methode wird also beim Datensatz (A) die globale  $H_0$  abgelehnt, wie schon bei der BONFERRONI-Methode. Beim Datensatz (B) jedoch bewirkt die (verglichen mit RÜGER  $k=2$ ) strengere Justierung, die aber notwendig ist, will man die willkürliche Vorgabe von  $k=2$  vermeiden, daß eine Ablehnung der globalen  $H_0$  nicht erreicht werden kann.

Wiederum drängt sich der Wunsch auf, einen Mittelweg zwischen BONFERRONI-Methode und RÜGER-Methode (mit  $k=2$ ) zu gehen.

Im Abschnitt 3 wird uns das gelingen.

Bemerkung: Der Vollständigkeit halber sei nochmals angemerkt werden, daß die Abschätzungen des Gesamtsignifikanzniveaus  $\gamma$  weder bei der BONFERRONI-, noch bei der RÜGER-, noch bei der HOMMEL - Methode zu grob sind. RÜGER [1978] und HOMMEL [1983] haben nachgewiesen, daß die jeweiligen Abschätzungen des Gesamtsignifikanzniveaus - und damit auch die jeweiligen  $\alpha$  - Justierungen – in dem Sinne scharf sind als jeweils Abhängigkeitsstrukturen existieren, bei denen durch die vorgenommenen  $\alpha$  - Justierungen und die entsprechenden Entscheidungsvorschriften das Gesamtsignifikanzniveau  $\gamma$  voll ausgeschöpft wird.

### 3. Eine allgemeine Simultantestmethode

#### 3.1 Grundsätzliche Bemerkungen

Die drei diskutierten Simultantestmethoden (BONFERRONI, RÜGER, HOMMEL) , bzw. die ihnen jeweils zugrundeliegenden wahrscheinlichkeitstheoretischen Abschätzungen, sind Spezialfälle eines sehr viel allgemeineren Zusammenhanges, den wir jetzt beweisen werden.

Diese allgemeinere Aussage erlaubt dann beispielsweise auch, eine Simultantestprozedur zu entwickeln, die einem 'Mittel' zwischen der BONFERRONI-Methode und der RÜGER-Methode (mit  $k=2$ ) entspricht, die zur Verfügung zu haben, uns schon in Abschnitt 2.3 wünschenswert erschien.

Um den Beweis für diese Methode sauber führen zu können, müssen wir mathematisch etwas weiter ausholen.

Gegeben sei ein meßbarer Raum, der sogenannte Stichprobenraum  $(X, \mathbf{B})$ . Die Stichproben  $x$  sind die Elemente von  $X$ .

$\mathbf{B}$  ist die Borelsche  $\sigma$  - Algebra über  $X$ .

Auf  $(X, \mathbf{B})$  existiert eine Wahrscheinlichkeitsverteilung  $P$ , die jedoch nicht bekannt ist. Aufgrund der Kenntnis des speziell vorliegenden Problems weiß man lediglich, daß  $P$  einer bestimmten, durch den Parameter  $\vartheta$  parametrisierten Klasse

$$\Pi = \{P_{\vartheta} \mid \vartheta \in \Theta\}$$

von in Betracht kommenden Wahrscheinlichkeitsverteilungen auf  $(X, \mathbf{B})$  angehört.

Die Menge  $\Theta$  aller möglichen Parameter  $\vartheta$  werde Parameterraum genannt.

Im Parameterraum  $\Theta$  sind nun  $n$  spezielle Teilmengen

$$H_0^{(j)} \subseteq \Theta \quad \text{als Nullhypothesen ausgezeichnet, } (j = 1, 2, \dots, n).$$

Ferner sei ein beliebig aber monoton gestaffeltes System von Signifikanzschranken vorgegeben:

$$0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n < 1.$$

Dementsprechend liege für alle  $j = 1, 2, \dots, n$  jeweils ein monotones System von Ablehnungsbereichen

$$A_j^1 \subseteq A_j^2 \subseteq \dots \subseteq A_j^n$$

vor mit

$$\sup_{\vartheta \in H_0^{(j)}} P_{\vartheta}(A_j^k) \leq \alpha_k \quad \text{für alle } j = 1, \dots, n \text{ und } k = 1, \dots, n.$$

Zu jeder Einzelhypothese  $H_0^{(j)}$  (bzw. zu jedem Einzeltest der  $H_0^{(j)}$ ,  $j=1, \dots, n$ ) gibt es also ein System von  $n$  monoton gestaffelten Ablehnungsbereichen, wobei wir  $A_j^k$  den Ablehnungsbereich der Stufe  $k$  für  $H_0^{(j)}$  nennen.

Ferner wird vorausgesetzt, daß die Globalhypothese nicht leer ist:

$$H_o := \bigcap_{j=1}^n H_0^{(j)} \neq \emptyset.$$

Wir wählen nun ein fixes aber beliebiges  $\vartheta \in \Theta$

und beziehen nun alle folgenden Wahrscheinlichkeitsaussagen auf dieses  $\vartheta$  :

$$P( \dots ) = P_{\vartheta}( \dots ).$$

Den Ablehnungsbereich der Globalhypothese definieren wir folgendermaßen:

$$A := \{x \mid \exists k, \text{ soda\ss } x \text{ in wenigstens } k \text{ der } A_1^k, \dots, A_n^k \text{ liegt} \}.$$

Das heißt, die Entscheidung über die Globalhypothese  $H_0$  wird nach folgender Vorschrift getroffen:

$$(11) \quad \left\{ \begin{array}{l} \text{Lehne die globale Nullhypothese } H_0 \text{ genau dann ab, wenn ein} \\ k \in \{1, 2, \dots, n\} \text{ so existiert, da\ss die Stichprobe } x \text{ in wenigstens } k \\ \text{der Ablehnungsbereiche der Stufe } k \text{ fällt, wenn also wenigstens} \\ k \text{ der Einzeltests die jeweilige Einzel – Nullhypothese } H_0^{(j)} \text{ auf} \\ \text{dem Niveau } \alpha_k \text{ ablehnen.} \end{array} \right.$$

Diese Entscheidungsvorschrift lautet so ähnlich wie die Entscheidungsvorschrift (10) nach HOMMEL [1983]. Der Unterschied liegt darin, daß die Signifikanzschranken  $\alpha_k$  jetzt außer der Monotonie-Bedingung keiner der HOMMELschen Bedingung (8) vergleichbaren Gesetzmäßigkeit mehr unterliegen müssen.

Gesucht ist nun ganz allgemein (d.h. unter Zulassung beliebiger Abhängigkeiten der

Ablehnungsbereiche  $A_j^k$ ) eine möglichst scharfe Abschätzung der Wahrscheinlichkeit von A nach

oben, welche nur noch von der Größe der vorgegebenen Werte  $\alpha_1, \alpha_2, \dots, \alpha_n$  (und natürlich von n) abhängig ist:

$$P(A) \leq S(\alpha_1, \alpha_2, \dots, \alpha_n).$$

Das heißt: Gesucht ist eine obere Schranke  $S(\alpha_1, \alpha_2, \dots, \alpha_n)$  für  $P(A)$ , die nur von den numerischen Werten der Signifikanzschranken  $\alpha_1, \alpha_2, \dots, \alpha_n$  abhängt, nicht aber von der speziellen Lage der Einzelablehnungsbereiche  $A_j^k$ . Diese obere Schranke  $S(\alpha_1, \alpha_2, \dots, \alpha_n)$  sollte natürlich möglichst klein sein. Am besten wäre es, wenn  $S(\alpha_1, \alpha_2, \dots, \alpha_n)$  sogar eine obere Grenze (also eine kleinst-mögliche obere Schranke) für  $P(A)$  darstellen würde; dies wäre z.B. der Fall, wenn bei bestimmten Abhängigkeitsstrukturen der Einzelablehnungsbereiche  $A_j^k$  das Gleichheitszeichen  $P(A) = S(\alpha_1, \alpha_2, \dots, \alpha_n)$  angenommen wird.

Für spezielle Vorgaben der numerischen Werte von  $\alpha_1, \alpha_2, \dots, \alpha_n$  haben RÜGER [1978], LINDNER [1979] und HOMMEL [1983] jeweils eine solche obere Grenze gefunden. Ein wesentliches Ergebnis der vorliegenden Arbeit besteht nun darin, eine solche obere Grenze  $S(\alpha_1, \alpha_2, \dots, \alpha_n)$  zu finden - und zwar für beliebig vorgegebene Werte  $\alpha_1, \alpha_2, \dots, \alpha_n$ , die nur der Monotoniebedingung

$$0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n < 1$$

genügen müssen, und sonst keinen weiteren Beschränkungen unterliegen.

Die Ergebnisse von RÜGER [1978], LINDNER [1979] und HOMMEL [1983] werden sich als Spezialfälle dieses allgemeineren Ergebnisses herausstellen.

### 3.2 Die Abschätzung für $P(A)$ durch $S(\alpha_1, \alpha_2, \dots, \alpha_n)$ .

Um die Wahrscheinlichkeit von A zu bestimmen, benutzen wir den in der Wahrscheinlichkeitstheorie üblichen Kunstgriff der Aufspaltung von A in paarweise disjunkte Teilmengen.

Zu diesem Zweck definieren wir für  $k = n, n-1, n-2, \dots, 3, 2, 1$  die Mengen

$$M_n := \{x \mid x \text{ liegt in allen } n \text{ Bereichen } A_1^n, \dots, A_n^n\}$$

.

.

$$M_k := \{x \mid x \text{ liegt in mindestens } k \text{ der } A_1^k, \dots, A_n^k\} - \bigcup_{v=k+1}^n M_v$$

Aus der Definition der Mengen  $M_k$  folgt zunächst unmittelbar deren paarweise Disjunktheit.

Ferner gilt

$$M_k := \{x \mid x \text{ liegt in genau } k \text{ der } A_1^k, \dots, A_n^k\}$$

Beweis : Sei  $x \in M_k$ .

Widerspruchsannahme: x liegt in mindestens (k+1) der  $A_1^k, \dots, A_n^k$ . Dann aber liegt wegen

$$A_j^k \subseteq A_j^{k+1}$$

x auch in mindestens (k+1) der  $A_1^{k+1}, \dots, A_n^{k+1}$ ,

woraus sofort folgt:

$$x \in \bigcup_{v=k+1}^n M_v \quad (\text{Widerspruch zur paarweisen Disjunktheit der } M_k).$$

Ferner läßt sich zeigen:

$$(12) \quad A = \bigcup_{k=1}^n M_k.$$

Beweis zu (12): Es sei zunächst  $x \in A$ . Dann existiert nach Definition von  $A$  ein  $k$  so, daß  $x$  in mindestens  $k$  der  $A_1^k, \dots, A_n^k$  liegt.

Also liegt  $x$  (nach Def. der  $M_1, \dots, M_n$ ) in  $M_k$  oder in  $\bigcup_{v=k+1}^n M_v$ .

Umgekehrt folgt trivialerweise aus  $x \in M_k$  sofort  $x \in A$ .

Es gilt also:

$$P(A) = \sum_{k=1}^n P(M_k) .$$

Wir wählen nun ein festes  $k$  mit  $0 < k \leq n$ .

Dann gibt es genau  $n$  Ablehnungsbereiche der Stufe  $k$ :  $A_1^k, A_2^k, \dots, A_n^k$ .

Mit diesem festen  $k$  definieren wir nun für alle  $i = 1, 2, \dots, n$  die folgenden Bereiche:

$$B_i^k := \{x \mid x \text{ liegt in genau } i \text{ der } A_1^k, \dots, A_n^k\} .$$

Per Definition sind die  $B_1^k, B_2^k, \dots, B_n^k$  paarweise disjunkt.

Aufgrund der Definition der Einzelablehnungsbereiche  $A_j^k$  gilt

$$n * \alpha_k \geq \sum_{j=1}^n P(A_j^k) \quad \text{für alle } k=1, \dots, n .$$

Nach FELLER [1968] gilt außerdem

$$P(A_1^k) + P(A_2^k) + \dots + P(A_n^k) = P(B_1^k) + 2 * P(B_2^k) + \dots + n * P(B_n^k) ,$$

zusammengefaßt also:

$$(13) \quad n * \alpha_k \geq \sum_{i=1}^n i * P(B_i^k)$$

Wir benötigen nun noch eine Beziehung zwischen den Mengen  $M_v$  ( $v = 1, \dots, k$ ) und  $B_i^k$  ( $i = 1, \dots, n$ ).

Und zwar gilt:

$$(14) \quad v \leq k \Rightarrow M_v \subseteq B_v^k \cup B_{v+1}^k \cup \dots \cup B_n^k .$$

Beweis zu (14): Es sei  $x \in M_v$ . Dann liegt  $x$  in mindestens  $v$  der  $A_1^v, \dots, A_n^v$ .

Wegen  $v \leq k$  liegt  $x$  dann auch in mindestens  $v$  der  $A_1^k, \dots, A_n^k$ .

Folglich liegt  $x$

“in genau  $v$  der  $A_1^k, \dots, A_n^k$  “ oder “in genau  $v+1$  der  $A_1^k, \dots, A_n^k$  “ oder .....

..... oder “in genau  $n$  der  $A_1^k, \dots, A_n^k$  “ ,

das heißt:  $x \in B_v^k \cup \dots \cup B_n^k$ .

Aus der Aussage (14) folgt unmittelbar:

$$(15) \quad i < v \leq k \Rightarrow B_i^k \cap M_v = \emptyset .$$

Die Beziehung (13) schreiben wir nun um, indem wir für alle  $i = 1, 2, \dots, n$  die Mengen  $B_i^k$  mit der Menge

$\{M_1 \cup M_2 \cup \dots \cup M_k\}$  schneiden.

Diese entstehenden Schnittmengen sind dann natürlich Teilmengen der  $B_i^k$  ( $i = 1, \dots, n$ );

und weil die Mengen  $M_1, M_2, \dots, M_k$  paarweise disjunkt sind, kann die Additivitätseigenschaft der Wahrscheinlichkeit ausgenutzt werden, was unmittelbar zu der folgenden Aussage führt:

$$(16) \quad n * \alpha_k \geq \sum_{v=1}^k \sum_{i=1}^n i * P(B_i^k \cap M_v)$$

Wegen (15) kann man gleichwertig dafür schreiben:

$$n * \alpha_k \geq \sum_{v=1}^k \sum_{i=v}^n i * P(B_i^k \cap M_v) .$$

Nun kann man weiter nach unten abschätzen:

$$(17) \quad n * \alpha_k \geq \sum_{v=1}^k \sum_{i=v}^n v * P(B_i^k \cap M_v) = \sum_{v=1}^k v * \sum_{i=v}^n P(B_i^k \cap M_v) .$$

Die innere Summe in der rechten Seite von (17) ist aber wegen (14) nichts anderes als  $P(M_v)$  und wir erhalten damit:

$$(18) \quad n * \alpha_k \geq \sum_{v=1}^k v * P(M_v) .$$

Da wir  $k$  zwar fest aber beliebig gewählt hatten, gilt die Aussage (18) für alle  $k = 1, 2, \dots, n$ .

Das heißt, (18) stellt ein System von  $n$  Ungleichungen dar, welche man sich durch den Index  $k$  durchnummeriert denken kann.

Wir multiplizieren die ersten  $(n - 1)$  dieser  $n$  Ungleichungen (also diejenigen, die aus (18) für  $k = 1, 2, \dots, n-1$  entstehen) jeweils mit dem Faktor

$$\frac{1}{k(k+1)}$$

und die letzte Ungleichung mit dem Faktor

$$\frac{1}{n} .$$

Dann addieren wir alle  $n$  Ungleichungen und erhalten:

$$\sum_{k=1}^{n-1} \sum_{v=1}^k \frac{v * P(M_v)}{k(k+1)} + \sum_{v=1}^n \frac{v * P(M_v)}{n} \leq \sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)} + \alpha_n$$

Durch Umsortieren der Summanden in der Doppelsumme ergibt sich daraus

$$\sum_{v=1}^{n-1} \sum_{k=v}^{n-1} \frac{v * P(M_v)}{k(k+1)} + \sum_{v=1}^n \frac{v * P(M_v)}{n} \leq \sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)} + \alpha_n$$

bzw.

$$\sum_{v=1}^{n-1} \sum_{k=v}^{n-1} \frac{v * P(M_v)}{k(k+1)} + \sum_{v=1}^{n-1} \frac{v * P(M_v)}{n} + P(M_n) \leq \sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)} + \alpha_n ,$$

was sich auch folgendermaßen schreiben läßt:

$$(19) \quad \sum_{v=1}^{n-1} v * P(M_v) \left[ \sum_{k=v}^{n-1} \frac{1}{k(k+1)} + \frac{1}{n} \right] + P(M_n) \leq \sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)} + \alpha_n .$$

Behauptung: Für alle  $n = 2, 3, \dots$  und alle  $v = 1, 2, \dots, (n-1)$  gilt:

$$(20) \quad \sum_{k=v}^{n-1} \frac{1}{k(k+1)} + \frac{1}{n} = \frac{1}{v}$$

Beweis zu (20): Mit Hilfe von vollständiger Induktion läßt sich direkt zeigen, daß für alle natürlichen Zahlen  $n$  gilt:

$$\sum_{k=1}^n \frac{1}{k(k+1)} = 1 - \frac{1}{n+1} .$$

Daraus folgt sofort:

$$\sum_{k=1}^{n-1} \frac{1}{k(k+1)} + \frac{1}{n} = 1 \quad \text{für alle } n=2, 3, \dots ,$$

womit sich sofort für alle  $n=2, 3, \dots$  und alle  $v=1, 2, \dots, (n-1)$  ergibt:

$$\sum_{k=v}^{n-1} \frac{1}{k(k+1)} + \frac{1}{n} = 1 - \sum_{k=1}^{v-1} \frac{1}{k(k+1)} = 1 - \left(1 - \frac{1}{v}\right) = \frac{1}{v} .$$

Unter Verwendung der nun bewiesenen Behauptung (20) erhält man aus (19) unmittelbar die folgende Aussage:

$$(21) \quad P(A) = \sum_{k=1}^n P(M_k) \leq \sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)} + \alpha_n .$$

Als Ergebnis können wir also festhalten:

Bei Anwendung der Entscheidungsvorschrift (11) wird das Gesamtsignifikanzniveau  $\gamma$  eingehalten,

wenn die Werte  $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$  so gewählt werden, daß

$$(22) \quad \sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)} + \alpha_n = \gamma$$

gilt.

Im Zuge der obigen Herleitung ist nun mehrfach  $P(A)$  nach oben abgeschätzt worden, so daß sich natürlich die Frage erhebt, wie grob die Abschätzung (21) ist.

Dies hängt im Einzelfall natürlich stark von der Struktur der Abhängigkeitsverhältnisse zwischen den Einzelablehnungsbereichen ab. Wichtig ist aber die Frage, ob die Abschätzung (21) nicht generell zu grob ist.

Dies ist nicht der Fall, denn man kann zeigen, daß die Einzelablehnungsbereiche so strukturiert sein können, daß in (21) das Gleichheitszeichen gilt.

Dies soll im folgenden geschehen.

Dazu zeigen wir zunächst:

$$(23) \quad \left\{ \begin{array}{l} \text{Das Gleichheitszeichen wird in (21) immer dann angenommen, wenn für} \\ \text{alle } k = 1, 2, \dots, n \text{ gilt: } P(M_k) = \frac{n(\alpha_k - \alpha_{k-1})}{k} \quad (\text{mit } \alpha_0 = 0). \end{array} \right.$$

Beweis zu (23): Nehmen wir an, die Wahrscheinlichkeiten  $P(M_k)$  haben die in (23) angegebenen Werte.

Dann ergibt sich:

$$P(A) = \sum_{k=1}^n P(M_k) = n * \left[ \left(1 - \frac{1}{2}\right) \alpha_1 + \left(\frac{1}{2} - \frac{1}{3}\right) \alpha_2 + \dots + \left(\frac{1}{k} - \frac{1}{k+1}\right) \alpha_k + \dots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \alpha_{n-1} \right] + \alpha_n$$

Wegen  $\frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)}$  ist aber der Ausdruck auf der rechten Seite der vorstehenden

Gleichung nichts anderes als  $\sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)} + \alpha_n$ , womit die Aussage (23) bewiesen ist.

Damit reduziert sich die Frage also darauf, ob es Konstellationen der Einzelablehnungsbereiche  $A_j^k$

so gibt, daß für alle  $k = 1, \dots, n$  gilt:  $P(M_k) = \frac{n(\alpha_k - \alpha_{k-1})}{k}$ .

Dies aber ist der Fall. Bei LINDNER [1979] ist nämlich gezeigt worden, daß allgemein durch  $n$  Ereignisse der Wahrscheinlichkeit  $\beta$  ein Ereignis der Wahrscheinlichkeit  $n\beta/k$   $k$ -fach überdeckt werden kann.

Damit können, wie man sich elementargeometrisch leicht überlegt, die Einzelablehnungsbereiche  $A_j^k$  für alle  $k = 1, \dots, n$  und  $j = 1, \dots, n$  so liegen, daß (23) erfüllt ist.



### 3.3 BONFERRONI-, RÜGER- und HOMMEL-Abschätzungen als Spezialfälle

Wie man leicht verifiziert, sind BONFERRONI-, RÜGER- und HOMMEL- Methode jeweils Spezialfälle des in Abschnitt 3.2 dargestellten allgemeinen Prinzips.

Die BONFERRONI – Methode erhält man aus der in 3.2 vorgeschlagenen Entscheidungsvorschrift

indem man speziell  $\alpha_1 = \alpha_2 = \dots = \alpha_n = \frac{\gamma}{n}$  wählt.

Man erhält dann gemäß (21) die folgende Abschätzung:

$$P(A) \leq \sum_{v=1}^{n-1} \frac{n * \alpha_v}{v(v+1)} + \alpha_n = \sum_{v=1}^{n-1} \frac{n * \frac{\gamma}{n}}{v(v+1)} + \frac{\gamma}{n} = \gamma \left[ \sum_{v=1}^{n-1} \frac{1}{v(v+1)} + \frac{1}{n} \right] = \gamma ,$$

und  $H_0$  wird genau dann abgelehnt, wenn mindestens eine der Einzelhypothesen  $H_0^{(j)}$  abgelehnt wird.

Die RÜGER – Methode ergibt sich, indem man für ein fixes k speziell

$\alpha_1 = \alpha_2 = \dots = \alpha_{k-1} = 0$  und  $\alpha_k = \alpha_{k+1} = \dots = \alpha_n = \frac{k * \gamma}{n}$  wählt.

Man erhält dann gemäß (21) die folgende Abschätzung:

$$P(A) \leq \sum_{v=1}^{n-1} \frac{n * \alpha_v}{v(v+1)} + \alpha_n = \sum_{v=k}^{n-1} \frac{n * \frac{k * \gamma}{n}}{v(v+1)} + \frac{k * \gamma}{n} = k * \gamma \left[ \sum_{v=k}^{n-1} \frac{1}{v(v+1)} + \frac{1}{n} \right] = \gamma ,$$

und  $H_0$  wird genau dann abgelehnt, wenn mindestens k der Einzelhypothesen  $H_0^{(j)}$  abgelehnt werden.

Die HOMMEL – Methode ergibt sich, indem man für  $v = 1, 2, \dots, n$  speziell

$\alpha_v = \frac{v * \gamma}{C(n) * n}$  mit  $C(n) = \sum_{v=1}^n \frac{1}{v}$  wählt.

Man erhält dann gemäß (21) die folgende Abschätzung:

$$P(A) \leq \sum_{v=1}^{n-1} \frac{n * \alpha_v}{v(v+1)} + \alpha_n = \sum_{v=1}^{n-1} \frac{n * \frac{v * \gamma}{C(n) * n}}{v(v+1)} + \frac{\gamma}{C(n)} = \frac{\gamma}{C(n)} \left[ \sum_{v=1}^{n-1} \frac{1}{(v+1)} + 1 \right] = \gamma ,$$

und  $H_0$  wird genau dann abgelehnt, wenn ein k so existiert, daß wenigstens k der Einzeltests die jeweilige Einzelnullhypothese  $H_0^{(j)}$  auf dem Niveau  $\alpha_k$  ablehnen.

### 3.4 Vorschläge zur Wahl der $\alpha_k$

Wie kann man nun die Werte von Signifikanzschranken  $\alpha_1, \alpha_2, \dots, \alpha_n$  der  $n$  Einzeltests wählen ?

(a) Da die Werte  $\alpha_1, \alpha_2, \dots, \alpha_n$  außer der Monotoniebedingung und der Relation (22) keinen weiteren Einschränkungen unterliegen, kann man zu ihrer Wahl sehr allgemein folgendermaßen vorgehen:

Man beginnt mit der Wahl von  $\alpha_1$ .

Für  $\alpha_1$  muß sicher

$$0 \leq \alpha_1 \leq \frac{\gamma}{n}$$

gelten, da mit  $\alpha_1 = \alpha_2 = \dots = \alpha_n = \frac{\gamma}{n}$  die rechte Seite von (21) genau gleich  $\gamma$  wird.

Die Frage lautet nun: Wie kann der Wert von  $\alpha_2$  (und danach der von  $\alpha_3$  usw.) gewählt werden ?

Sicher nicht kleiner als  $\alpha_1$  (wegen der Monotonie) aber natürlich auch nicht beliebig groß, da schließlich die Bedingung (22) erfüllt werden muß.

Um diese Frage zu beantworten, nehmen wir an, daß für ein beliebiges  $k_0 \in \{2, 3, \dots, n-1\}$  die Werte von  $\alpha_1, \alpha_2, \dots, \alpha_{k_0-1}$  bereits festgelegt worden sind.

In welchen Schranken kann dann  $\alpha_{k_0}$  gewählt werden ?

Wegen der Monotonie der  $\alpha_k$  muß sicher gelten

$$\alpha_{k_0-1} \leq \alpha_{k_0}.$$

Den maximal möglichen Wert  $\alpha_{\max}$  für  $\alpha_{k_0}$  erhält man, indem man

$$\alpha_{k_0} = \alpha_{k_0+1} = \dots = \alpha_n$$

setzt und (22) berücksichtigt, d.h.:

$$\begin{aligned} \gamma - \sum_{k=1}^{k_0-1} \frac{n * \alpha_k}{k(k+1)} &= \sum_{k=k_0}^{n-1} \frac{n * \alpha_{\max}}{k(k+1)} + \alpha_{\max} \\ &= n * \alpha_{\max} * \left[ \sum_{k=k_0}^{n-1} \frac{1}{k(k+1)} + \frac{1}{n} \right] \\ &= \frac{n * \alpha_{\max}}{k_0}, \end{aligned}$$

woraus sich die obere Grenze für  $\alpha_{k_0}$  ergibt:

$$\alpha_{k_0} \leq \alpha_{\max} = \frac{k_0 * \left[ \gamma - \sum_{k=1}^{k_0-1} \frac{n * \alpha_k}{k(k+1)} \right]}{n}$$

In den Grenzen  $\alpha_{k_0-1} \leq \alpha_{k_0} \leq \alpha_{\max}$  kann nun  $\alpha_{k_0}$  beliebig gewählt werden.

So kann fortgefahren werden bis  $\alpha_{n-1}$ .

Der Wert für  $\alpha_n$  ergibt sich dann gemäß (22) zu

$$\alpha_n = \gamma - \sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)}.$$

Die – sehr allgemeine – Vorschrift zur Wahl der Signifikanzschranken  $\alpha_1, \alpha_2, \dots, \alpha_n$  lautet also:

$$(24) \quad \left\{ \begin{array}{l} \text{Wähle } \alpha_1 \leq \gamma/n. \text{ Für } k=2,3, \dots, n-1 \text{ wähle dann (sukzessive) } \alpha_k \text{ so, daß} \\ \alpha_{k-1} \leq \alpha_k \leq \frac{k}{n} \left[ \gamma - \sum_{l=1}^{k-1} \frac{n * \alpha_l}{l(l+1)} \right] \text{ gilt. Wähle ferner } \alpha_n = \gamma - \sum_{k=1}^{n-1} \frac{n * \alpha_k}{k(k+1)}. \end{array} \right.$$

Innerhalb der in (24) genannten Grenzen kann man also die Signifikanzschranken  $\alpha_1, \alpha_2, \dots, \alpha_n$  frei wählen, wodurch eine außerordentliche Flexibilität gewährleistet wird. Bei jeder Wahl der  $\alpha_1, \alpha_2, \dots, \alpha_n$  gemäß (24) ist durch die im Abschnitt 3.2 bewiesene Relation (22) die Einhaltung des vorgegebenen Gesamtsignifikanzniveaus  $\gamma$  gesichert.

(b) Eine sehr spezielle Variante der Wahl von  $\alpha_1, \alpha_2, \dots, \alpha_n$  ergibt sich aus dem schon in Abschnitt 2.3 erwähnten Vorschlag, einen "Mittelweg" zwischen BONFERRONI- und RÜGER-Methode (mit  $k=2$ ) herzustellen.

Zu diesem Zweck wählt man :

$$(25) \quad \alpha_1 = 0.5 * \frac{\gamma}{n} \quad \text{und} \quad \alpha_2 = \dots = \alpha_n = 1.5 * \frac{\gamma}{n}.$$

(Man prüft leicht nach, daß bei dieser Wahl der  $\alpha_1, \dots, \alpha_n$  die Relation (22) eingehalten wird).

Der Vorteil gegenüber der RÜGER-Methode (mit  $k=2$ ) liegt darin, daß man die globale Nullhypothese  $H_0$  auch schon bei nur einem auf dem Niveau  $\alpha_1$  signifikanten Einzeltest ablehnen kann, während der Vorteil gegenüber der BONFERRONI-Methode darin besteht, daß man  $H_0$  auch dann ablehnen

kann, wenn zwar kein Einzeltest auf dem strengen BONFERRONI-Niveau  $\frac{\gamma}{n}$  signifikant wird, aber

zwei Einzeltests auf dem etwas schwächeren Niveau  $\alpha_2 = 1.5 * \frac{\gamma}{n}$  signifikant ausfallen. Der Vorteil

gegenüber der HOMMEL-Methode besteht natürlich darin, daß dort die Schranken  $\alpha_1$  und  $\alpha_2$  wesentlich restriktiver sind.

#### 4. Ein Beispiel

Abschließend soll noch an einem praktischen Beispiel die Anwendung der in Abschnitt 3 entwickelten Simultantestprozedur demonstriert werden.

Nehmen wir an, ein Wissenschaftler forsche – wie WEIHE et al. [1987] – nach Zusammenhängen zwischen der Höhe des Anfangsgehaltes von Hochschulabsolventen und den jeweiligen Unternehmenstypen, bei denen der Absolvent angestellt ist.

Die diesbezügliche Datenerhebung bei  $N = 85$  Absolventen habe zu folgenden Ergebnissen geführt:

Höhe des Anfangs- gehaltes		unter 70000 $A_1$	70000 –75000 $A_2$	75000 –80000 $A_3$	80000 –85000 $A_4$	85000 –90000 $A_5$	über 90000 $A_6$	Zeilen- summen
Betriebs- größe								
sehr klein	$B_1$	7	2	2	2	1	1	15
klein	$B_2$	3	5	4	4	4	1	21
mittel	$B_3$	2	5	5	4	4	1	21
groß	$B_4$	2	3	4	5	5	1	20
sehr groß	$B_5$	0	1	1	1	1	4	8
Spalten- summen		14	16	16	16	15	8	$N = 85$

Das vorgegebene (globale) Signifikanzniveau sei  $\gamma = 0,05$ . Ferner suche man nicht nach unterbesetzten sondern nur nach überbesetzten Zellen.

Ein globaler  $\chi^2$  - Test auf Unabhängigkeit in allen 30 Zellen dieser Kontingenztafel kann leider nicht durchgeführt werden, da die erwarteten Häufigkeiten pro Zelle viel zu klein sind, - ja sogar alle unter 5 liegen.

Daher wollen wir statt eines Globaltests in jeder einzelnen der 30 Zellen auf Unabhängigkeit testen. Dies soll mit Hilfe des exakten FISHER-Tests für  $2 \times 2$ -Tafeln geschehen, wobei aus inhaltlichen Gründen aber nur nach überbesetzten und nicht nach unterbesetzten Zellen gesucht werde.

Beispielsweise sieht die zu der Zelle  $A_1B_1$  gehörige  $2 \times 2$  -Tafel folgendermaßen aus.

Höhe des Anfangs- gehaltes		unter 70000 $A_1$	sonstig $A_2 - A_6$	Zeilensummen
Betriebsgröße				
sehr klein	$B_1$	7	8	15
sonstig	$B_2 - B_5$	7	63	70
Spaltensummen		14	71	$N = 85$

Die erwartete Häufigkeit in der untersuchten Zelle berechnet sich zu  $\frac{15}{85} * \frac{14}{85} * 85 = 2,47$ .

Die Restwahrscheinlichkeit für eine Zellbesetzungszahl von mindestens 7 berechnet sich mit Hilfe der hypergeometrischen Verteilung zu

$$P(X \geq 7) = \sum_{m=7}^{14} \frac{\binom{15}{m} * \binom{85-15}{14-m}}{\binom{85}{14}} = 0,00227$$

Entsprechend lassen sich die Restwahrscheinlichkeiten in den restlichen 29 Zellen berechnen, wobei sich herausstellt, daß überhaupt nur in zwei Zellen – nämlich  $A_1B_1$  und  $A_6B_5$  – Restwahrscheinlichkeiten von unter 5% auftreten.

Insbesondere ergibt sich für die Zelle  $A_6B_5$  eine Restwahrscheinlichkeit von:

$$P(X \geq 4) = \sum_{m=4}^8 \frac{\binom{8}{m} * \binom{85-8}{8-m}}{\binom{85}{8}} = 0,00206$$

Die Beurteilung dieser Ergebnisse mit Hilfe der BONFERRONI-Methode sieht folgendermaßen aus:

Nach (4) wird das Signifikanzniveau  $\alpha$  der Einzeltests berechnet gemäß:

$$\alpha = \frac{\gamma}{n} = \frac{0,05}{30} = 0,001\bar{6}$$

Es stellt sich heraus, daß keiner der 30 Einzeltests zu diesem Niveau signifikant ausfällt. Aufgrund der BONFERRONI-Methode muß die globale Nullhypothese der Unabhängigkeit in allen 30 Zellen also beibehalten werden. Es konnten mit der BONFERRONI-Methode keine Zusammenhänge zwischen Unternehmenstyp und Anfangsgehalt festgestellt werden.

Die Beurteilung der Ergebnisse mit Hilfe der RÜGER-Methode (mit  $k=2$ ) geschieht wie folgt:

Nach (6) wird das Signifikanzniveau  $\alpha$  der Einzeltests berechnet gemäß:

$$\alpha = \frac{2 * \gamma}{n} = \frac{2 * 0,05}{30} = 0,00\bar{3}.$$

Zwei der Zellbesetzungszahlen (nämlich in  $A_1B_1$  und  $A_6B_5$ ) stellen sich auf diesem Niveau als überzufällig heraus. Nach der RÜGERschen Entscheidungsvorschrift (7) kann also die globale Nullhypothese abgelehnt werden. Mit der RÜGERschen Methode (mit  $k=2$ ) konnten also in zwei Zellen die Abhängigkeiten zwischen Unternehmenstyp und Anfangsgehalt nachgewiesen werden. Dennoch muß hier aber auf einen gravierenden Nachteil der RÜGERschen Methode hingewiesen werden. Die Zahl  $k=2$  mußte vor der Datenerhebung festgelegt werden. Würde nur eine der Zellen überzufällig häufig besetzt sein, so hätte diese Tatsache mit der RÜGERschen Methode unter keinen Umständen entdeckt werden können.

Aufgrund der HOMMELschen Methode sieht die Beurteilung der Ergebnisse folgendermaßen aus:

Nach (8) und (9) werden die gestaffelten Signifikanzschranken  $0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$

wie folgt berechnet:

$$\alpha_k = \frac{k * \gamma}{C(n) * n} = \frac{k * 0,05}{3,995 * 30} = k * 0,00042$$

das heißt:

$$\alpha_1 = 0,00042$$

$$\alpha_2 = 0,00083$$

$$\alpha_3 = 0,00125$$

usw.

Es stellt sich heraus, daß kein k so existiert, daß k Einzeltests zum Niveau  $\alpha_k$  signifikant werden.

Die HOMMEL-Methode führt also ebenso wie die BONFERRONI-Methode dazu, daß keinerlei Zusammenhänge nachgewiesen werden können.

Analysieren wir nun noch abschließend die Ergebnisse mit der hier (in Abschnitt 3) neu entwickelten Simultantestmethode – aufgrund der Entscheidungsvorschrift (11), wobei die gestaffelten Signifikanzschranken  $0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$  berechnet werden gemäß (25):

$$\alpha_1 = \frac{0,5 * \gamma}{n} = \frac{0,5 * 0,05}{30} = 0,0008\bar{3} \quad \text{und}$$

$$\alpha_2 = \alpha_3 = \dots = \alpha_{30} = \frac{1,5 * \gamma}{30} = 0,0025.$$

Die Restwahrscheinlichkeiten von zwei Einzeltests (die erwähnten zwei Zellen betreffend) liegen unter der Signifikanzschranke  $\alpha_2 = 0,0025$ . Damit kann nach der Entscheidungsvorschrift (11) die globale Nullhypothese der Unabhängigkeit verworfen werden. Die Zellbesetzungszahlen in den beiden entsprechenden Zellen können als signifikante Überfrequenzen interpretiert werden.

Ohne den Nachteil der RÜGER-Methode, ein festes k vor der Datenerhebung festlegen zu müssen, kann mit der vorgeschlagenen Methode auf dem vorgegebenen globalen Signifikanzniveau also nachgewiesen werden, daß in den erwähnten zwei Zellen Abhängigkeiten zwischen Unternehmenstyp und Anfangsgehalt bestehen.

Sowohl mit der BONFERRONI-Methode als auch mit der HOMMEL-Methode konnte dieses Ergebnis nicht erreicht werden.

Abschließend sei nochmals darauf hingewiesen, daß die Wahl der  $\alpha_1, \alpha_2, \dots, \alpha_n$  gemäß (25) natürlich sehr speziell ist, - die unter (24) bereitgestellte Vorschrift zur Wahl der Signifikanzschranken  $\alpha_1, \alpha_2, \dots, \alpha_n$  bietet hingegen ein überaus breites Spektrum an Möglichkeiten zur Gestaltung von Simultantests.

## LITERATUR:

- COCHRAN, W.G., 1954: Some methods for strengthening the common  $X^2$  tests. Biometrics, 10, 417 - 451.
- FELLER, W., 1968: An Introduction to Probability Theory and Its Applications, Vol. 1, 3. Ed., New York.
- HOMMEL, G., 1983: Tests of Overall Hypothesis for Arbitrary Dependence Structures. Biom. J. , 25, 423 - 430.
- KRAUTH, J. und G.A. LIENERT, 1973: Nichtparametrischer Nachweis von Syndromen durch simultane Binomialtests. Biom. Zeitschrift, 15, 13 - 20.
- LINDNER, K., 1979: Das maximale Signifikanzniveau bei Simultantests. Proceedings in Operations Research, 9, 330 - 335.
- MILLER, R.G., 1981: Simultaneous statistical inference. 2. Ed., New York.
- MORGENSTERN, D., 1980: Berechnung des maximalen Signifikanzniveaus des Testes "Lehne  $H_0$  ab, wenn k unter n gegebenen Tests zur Ablehnung führen". Metrika, 27, 285 - 286.
- RÜGER, B., 1978: Das maximale Signifikanzniveau des Tests: "Lehne  $H_0$  ab, wenn k unter n gegebenen Tests zur Ablehnung führen ". Metrika, 25, 171 - 178.
- TOCHER, K.D., 1950: Extensions of the Neyman-Pearson Theory of tests to discontinuous variates. Biometrika, 37, 130 - 144.
- WEIHE, H.J., C.H. HENCKE, B. TRUNZ, 1987: Berufseintrittsbedingungen von Fachhochschulabsolventen. Frankfurt/Main.
- WITTING, H., 1966: Mathematische Statistik. Stuttgart.