

1/1984

BEMERKUNGEN ZU KONTINGENZTAFELN

nach A. C. Kimber, University of Surrey
Originaltitel in 'Teaching Statistics' Vol. 5 (1983),
Nr. 3 : Some Remarks on Contingency Tables
Übersetzung: B. Wollring

Zweidimensionale Kontingenztafeln werden in fast allen ein-
führenden Statistik-Kursen diskutiert. Zu Recht, denn diese
einfachen Tafeln treten auf natürliche Weise bei vielen Fra-
gestellungen auf, besonders in den Sozialwissenschaften,
insbesondere dann, wenn die Gegenstände der Untersuchung
sich leicht in bestimmte Kategorien einteilen lassen, z. B.
männlich-weiblich, gegen-für-indifferent bezüglich eines
politischen Standpunktes etc. Ferner kann der bekannte χ^2 -
Test bei ihnen leicht durchgeführt werden und wird von Stu-
denten intuitiv akzeptiert.

Eines der Hauptziele, beim Unterrichten von Statistik, ist
es, den Studenten dazu zu befähigen, Statistik in alltäglichen
Situationen zu verwenden. Fast immer werden die Stu-
denten vor verwirrenden und zweideutigen Graphen oder Schau-
bildern gewarnt, z. B. angepaßten Kurven oder Piktogrammen.
Die Fußfallen beim Analysieren von Kontingenztafeln werden
leider weit seltener herausgestellt, und das, obwohl Vier-
feldertafeln recht häufig in Zeitungen und nicht wissen-
schaftlichen Zeitschriften auftreten, speziell im Zusammen-
hang mit "statistischen gefundenen Ergebnissen". Oft geht
es um den Zusammenhang von mehr als zwei Variablen, und die
Ergebnisse werden in Form mehrerer zweidimensionaler Kon-
tingenztafeln dargestellt. Dummerweise kann das getrennte
Untersuchen jeder einzelnen Tafel zu Fehlergebnissen füh-
ren.

Ein Beispiel für diesen Effekt gibt ein Artikel in der
Zeitschrift *The Guardian* vom 23.06.1982 über die Probleme
von Frauen am Arbeitsplatz. Die Ergebnisse basieren auf ei-
ner Erhebung, bei denen einer großen Anzahl arbeitender
Frauen bestimmte Fragen gestellt und die Häufigkeiten der
Antworten in diversen zweidimensionalen Kontingenztafeln

dargestellt wurden. (Angegeben waren Prozent-Anteile, aus dem Stichprobenumfang haben wir die absoluten Anteile rekonstruiert.) Tabelle 1 a zeigt zum Beispiel die Häufigkeiten für die Antworten Besser-Gleich-Schlechter auf die Frage "Glauben Sie, daß es die Arbeitsatmosphäre besser oder schlechter macht, wenn Männer und Frauen am Arbeitsplatz zusammen arbeiten?", wie sie von einer Stichprobe von Frauen gegeben wurden, die nach Altersgruppen unterschieden waren. Ausgerüstet mit unserem Wissen zum χ^2 -Test können wir testen, ob Antwort und Alter gekoppelt sind oder nicht. Die erwarteten Antworthäufigkeiten unter der Annahme, daß keine Koppelung besteht, zeigt Tabelle 1 b.

Antwort	(a) Erfragte Häufigkeiten			(b) Geschätzte Häufigkeiten		
	Alter 15-24	25-34	über 35	Alter 15-24	25-34	über 35
Besser	115	73	178	107.74	68.77	189.49
Gleich	13	11	19	12.66	8.08	22.26
Schlechter	13	6	51	20.60	13.15	36.25

Tabelle 1: Erfragte (a) und geschätzte (b) Antworthäufigkeiten

Der Wert von χ^2 ist 15.68, was bei 4 Freiheitsgraden hochsignifikant ist. Es scheint daher, daß Alter und Einstellung gekoppelt sind.

Die Frauen wurden bei der Untersuchung jedoch nicht nur in Altersgruppen, sondern auch nach anderen Merkmalen eingeteilt, zum Beispiel alleinstehend-verheiratet-geschieden oder danach, ob sie vorwiegend mit Männern oder Frauen zusammenarbeiten. Dies läßt unsere naive Interpretation der Tabelle 1 a von zweifelhaftem Wert erscheinen, denn diese Variablen wurden bei der Auswertung ignoriert, obwohl sie vielleicht wichtige Variablen sind. Um darzustellen, was das "Heraussummieren" einer "wichtigen Variablen" bewirken kann, geben wir nun zwei Beispiele an, bei denen die vollständigen Daten verfügbar sind.

Die folgende Datenliste wurde oft veröffentlicht, wir haben sie bei EVERITT (1977) entnommen. Die Daten beschreiben das Überleben von Babys in Abhängigkeit davon, wieviel klinische Schwangerschaftsvorsorge die Mütter in Anspruch genommen haben (mehr-weniger), wobei jede eine von zwei bestimmten Kliniken (X-Y) aufgesucht hatte. Tabelle 2 zeigt die vollständigen Daten.

Klinik	Umfang der Vorsorge	Babys		Sterblichkeitsrate in %
		gestorben	überlebt	
X	weniger	3	176	1.7
	mehr	4	293	1.4
Y	weniger	17	197	7.9
	mehr	2	23	8.0

Tabelle 2: Vollständige Daten zum Überleben von Babys

Auch ohne eine eingehende Analyse, die in jedem Fall den Rahmen einer Einführung sprengt, ist es klar, daß es bei *einer gegebenen Klinik* wenig Anlaß zu der Annahme gibt, daß der Umfang der klinischen Schwangerschaftsvorsorge mit dem Überleben der Kinder zusammenhängt. Wenn wir aber die Ergebnisse der einzelnen Kliniken zu denen der Vierfeldertafel in Tabelle 3 zusammenfassen, so sieht die Sache ganz anders aus:

Vorsorge	Babys	
	gestorben	überlebt
weniger	20 (14.29)	373 (378.71)
mehr	6 (11.71)	316 (310.29)

Tabelle 3: Vereinfachte Daten zum Überleben von Babys

Die geschätzten Werte unter der Annahme, daß kein Zusammenhang zwischen Überleben und Vorsorgeumfang besteht, ist jeweils in Klammern angegeben. So entsteht ein χ^2 -Wert von 5.26, was bei einem Freiheitsgrad hochsignifikant ist. Würden wir also nur die Vierfeldertafel 3 betrachten, würden wir irrtümlich zu dem Schluß kommen, daß Überleben und Vor-

sorgeumfang stark voneinander abhängen.

Das dritte Beispiel ist Teil einer Studie zum Englisch-Australischen Test-Cricket von CROUCHER (1982). Von speziellem Interesse ist hier der Zusammenhang von der Anzahl der "lbw-dismissals" ("Aus-Würfen") unter den "batsmen" ("Aufschläger") von hohem bzw. niedrigem Rang und dem Ort des "dismissals". Die Anzahl der "lbw-dismissals" in den Spielen von 1946 bis 1980 gibt Tabelle 4 an. Zusammengefaßte Daten zeigt Tabelle 5.

Nationalität des batsman	Ort des Spiels	Position in der Schlagordnung		Anteil der batsmen mit hohem Rang
		1 - 5	6 - 11	
Engländer	England	50	35	59
	Australien	64	37	63
Australier	England	72	30	71
	Australien	36	40	47

Tabelle 4: Vollständige Daten der "lbw-dismissals"

Ort des Spiels	Position in der Schlagordnung	
	1 - 5	6 - 11
England	122	65
Australien	100	77

Tabelle 5: Vereinfachte Daten der "lbw-dismissals"

Die Resultate zeigen, daß es unter den australischen "batsmen" einen bemerkenswerten Unterschied gibt zwischen den hochrangigen "batsmen", die in England "lbw" ("aus") werden und denen, die es in Australien werden. Einen solch deutlichen Unterschied gibt es bei den englischen "batsmen" nicht. Das ist natürlich hinsichtlich der Kontroversen über Feststellbarkeit und neutrale Schiedsrichter in den Testspielen interessant. Lassen wir jedoch die Nationalität der "batsmen" beiseite und betrachten die so entstehenden Daten der Tabelle 5, so erhalten wir einen χ^2 -Wert von 2.96, was bei einem Freiheitsgrad eins nicht signifikant ist. Dies

legt - irrtümlicherweise - nahe, daß ein Zusammenhang von Austragungsort und Position in der Schlagordnung bei den "lbw"-Entscheidungen unwahrscheinlich ist.

Um zu dem Artikel zurückzukehren: Hier wird sicher ein sehr komplexer Gegenstand angesprochen, und es war gewiß eher die Absicht des Autors, die Dinge für den Leser zu vereinfachen und nicht ihn in die Irre zu führen. Der Student sollte jedoch auf die Gefahren einer solchen Vereinfachung aufmerksam gemacht werden, die manchmal weit davon entfernt, das Verständnis zu fördern, nur zur Verwirrung führt.

Literatur

CROUCHER, J. S.: Anglo-Australian Test Cricket Dismissals 1946 - 1980. - B.I.A.S 9(1982), S. 179 - 193

EVERITT, B. S.: The Analysis of Contingency Tables. - Chapman and Hall, London, 1977