Z. Wahrscheinlichkeitstheorie verw. Geb. 4, 260—264 (1965)

Die Irrtumswahrscheinlichkeit des χ^2-Testes im Grenzfall der Poissonverteilung

Von

K. Egg, H. Rüst und B. L. van der Waerden

Summary. If the expectations of two Poisson distributions are compared by means of the χ^2-test, the probability of an error of the first kind is never much larger than the asymptotic value 0.01 or 0.02 or 0.05 on the three most usual levels.

Der χ^2-Test ist sicher dann gerechtfertigt, wenn die Erwartungswerte np_i der Anzahlen Objekte in allen betrachteten Klassen sehr groß sind. Der Beweis wurde in einer grundlegenden Arbeit von Neyman und Pearson unter sehr allgemeinen Voraussetzungen geliefert [2].

Es fragt sich nun, wie weit man mit den Erwartungswerten heruntergehen kann, ohne daß die Irrtumswahrscheinlichkeit des Testes zu groß wird.

Häufig findet man in der Literatur Bemerkungen von der Art, die Erwartungswerte np_i sollten mindestens 4 oder 5 betragen. Neuere Autoren aber, die die Frage sorgfältiger geprüft haben, kommen zu dem Schluß, daß man in vielen Fällen beträchtlich kleinere Erwartungswerte zulassen kann. Cochran [1] verwendet folgendes Kriterium: Wird der χ^2-Test auf dem 5%-Niveau angewandt und liegt die wahre Irrtumswahrscheinlichkeit zwischen 4% und 6%, oder beim 1%-Niveau zwischen 0,7% und 1,5%, so ist die Abweichung als unbedeutend zu betrachten. Er kommt zu dem Schluß, daß auf dem 5%-Niveau bei der Prüfung der Normalität einer Verteilung ein einziger Erwartungswert bis $^1/_2$ heruntergehen darf, sofern die übrigen genügend groß sind. Dasselbe gilt auf dem 1%-Niveau, wenn die Zahl der Freiheitsgrade mehr als 6 beträgt. Auf dem 5%-Niveau dürfen auch zwei Erwartungswerte bis 1 heruntergehen. Bei einer Vergrößerung der Zahl der Freiheitsgrade und bei kleinen Erwartungswerten fallen die Irrtumswahrscheinlichkeiten im allgemeinen kleiner aus als die angenommenen Niveau-Werte; man bleibt also bei der Anwendung des Testes auf der sicheren Seite.

Bei einer kleinen Zahl von Freiheitsgraden liegt die Sache unter Umständen weniger günstig. Wird eine einzige beobachtete Häufigkeit mit einer angenommenen Wahrscheinlichkeit verglichen, so hat man zwei Klassen und einen Freiheitsgrad. Wendet man nun den χ^2-Test auf dem 5%-Niveau an und verlangt, daß die wahre Irrtumswahrscheinlichkeit unter 6,5% bleibt, so genügt es nach einer neuen Untersuchung von E. Studer [3] nicht, zu fordern, daß die Erwartungswerte np und $n - np$ in beiden Klassen mindestens 5 betragen, sondern man muß verlangen, daß sie mindestens 8 betragen. Verlangt man (wie Cochran) eine wahre Irrtumswahrscheinlichkeit von höchstens 6%, so muß man mit den Erwartungswerten sogar bis 16 oder noch höher hinaufgehen.

Besonders wichtig für viele Anwendungen ist der Fall, daß zwei Wahrscheinlichkeiten mittels der sogenannten 2 mal 2 Tafel miteinander verglichen werden sollen. Man habe in $N = n_1 + n_2$ unabhängigen Versuchen zwei Häufigkeiten x/n_1 und y/n_2 beobachtet. Die Nullhypothese besagt, daß in allen N Versuchen die Wahrscheinlichkeit p des beobachteten Ereignisses immer den gleichen Wert hat. Wenn nun

$$\chi^2 = \frac{(x n_1 - y n_2)^2 N}{n_1 n_2 (x+y)(N-x-y)} \tag{1}$$

größer als A_1, A_2 oder A_5 ausfällt, so wird die Nullhypothese auf dem 1%, 2% oder 5% Niveau verworfen. Die Schranken A sind so gewählt, daß die Irrtumswahrscheinlichkeit für feste p, wenn n_1 und n_2 gegen ∞ gehen, asymptotisch 0,01 bzw. 0,02 bzw. 0,05 beträgt. Einen einfachen Beweis dieser asymptotischen Aussage, nach einer Idee von M.-P. GEPPERT, findet man bei VAN DER WAERDEN [*4*].

Für kleine Werte von n_1 und n_2 haben GILDEMEISTER und VAN DER WAERDEN [*5*] die Irrtumswahrscheinlichkeit des Testes als Funktion p berechnet und graphisch dargestellt. Es stellte sich heraus, daß in den meisten Fällen die Irrtumswahrscheinlichkeit kleiner oder jedenfalls nicht viel größer ist als der asymptotische Wert, besonders dann, wenn der Faktor N in (1) durch $(N - 1)$ ersetzt wird, wozu es auch theoretisch gute Gründe gibt [*4*].

Bei den praktischen Anwendungen kommt es häufig vor, daß n_1 und n_2 groß sind, aber x und y nicht. Um diese Fälle theoretisch zu behandeln, kann man von der Annahme ausgehen, daß n_1 und n_2 gegen ∞ streben, während pn_1 und pn_2, die Erwartungswerte von x und y, fest bleiben. Nennt man diese Erwartungswerte v und w, so hat man für x und y im Limes for $n_1 \to \infty$ und $n_2 \to \infty$ Poissonverteilungen anzunehmen. Die Wahrscheinlichkeit eines Wertepaares (x, y) ist dann

$$P(x, y) = \frac{v^x}{x!} e^{-v} \cdot \frac{w^y}{y!} e^{-w}. \tag{2}$$

Im folgenden soll der Fall $n_1 = n_2$, also $v = w$ numerisch untersucht werden. Alle zu berechnenden Wahrscheinlichkeiten sind Funktionen von w.

Die Formel für χ^2 ergibt für $n_1 = n_2$ im Grenzfall $N \to \infty$:

$$\chi^2 = \frac{(x-y)^2}{x+y}. \tag{3}$$

Die Formel (2) für $P(x, y)$ vereinfacht sich für $v = w$ zu

$$P(x, y) = \frac{w^{x+y}}{x!\,y!} e^{-2w}. \tag{4}$$

Um die Irrtumswahrscheinlichkeit des Testes zu erhalten, hat man $P(x, y)$ zu summieren über alle ganzzahligen x und y, die die Bedingung $\chi^2 > A$, oder, was dasselbe ist,

$$(x-y)^2 > (x+y)A \tag{5}$$

erfüllen. Die gesuchte Irrtumswahrscheinlichkeit ist also

$$P = \sum_{(5)} P(x, y) \tag{6}$$

summiert über alle x und y, die (5) erfüllen.

Man kann zu (5) noch die Nebenbedingung

$$x < y \tag{7}$$

hinzufügen und nachher die Summe verdoppeln:

$$P = 2 \sum_{(5),(7)} \frac{w^{x+y}}{x!\,y!} e^{-2w}. \tag{8}$$

Ferner kann man $x + y = z$ setzen und

$$P = 2\,e^{-2w} \sum_{1}^{\infty} c_z \frac{w^z}{z!} \tag{9}$$

schreiben, wobei c_z die ganze Zahl

$$c_z = \sum_{(11),(12)} \binom{z}{x} \tag{10}$$

ist. Die Summation erstreckt sich über diejenigen x, die die Bedingungen

$$(z - 2x)^2 > zA \tag{11}$$

$$z - 2x > 0 \tag{12}$$

erfüllen.

Man kann (9) ein wenig anders schreiben, indem man statt der c_z die Koeffizienten

$$d_z = 2^{-z} c_z \tag{13}$$

einführt. Die d_z sind die bedingten Wahrscheinlichkeiten, daß x und y die Ungleichung (5) erfüllen, falls $x + y = z$ einen bestimmten Wert hat. Statt (9) hat man dann

$$P = 2\,e^{-2w} \sum_{z=1}^{\infty} d_z \frac{(2w)^z}{z!}. \tag{14}$$

Die d_z haben gegenüber den c_z den rechnerischen Vorteil, daß sie alle dieselbe Größenordnung haben. Man hat nämlich für große z asymptotisch

$$2\,d_z \sim \alpha\,(\alpha = 0{,}01 \text{ oder } 0{,}02 \text{ oder } 0{,}05). \tag{15}$$

Die Reihe in (14) konvergiert ungefähr so gut wie die Exponentialreihe für e^{2w}. Das größte Glied dieser Exponentialreihe ist das Glied mit $z = [2w]$. Weiter als bis etwa $z = [2w + 4\sqrt{2w}]$ braucht man in der Reihe nicht zu gehen.

Die folgende Tabelle gibt $10^3 P$ als Funktion von w für die drei Fälle

$$A_1 = 6{,}64;\ A_2 = 5{,}41;\ A_5 = 3{,}84\,,$$

die der Reihe nach dem 1%-, 2%- und 5%-Niveau entsprechen. Die asymptotischen Werte von $10^3 P$ sind in den drei Fällen

$$10;\ 20;\ 50\,.$$

Die Tabelle wurde auf der elektronischen Rechenmaschine des Rechenzentrums der Universität Zürich berechnet.

Tabelle. *Irrtumswahrscheinlichkeit mal 10^3*

w	$10^3 P(w)$		
	$A_1=6{,}64$	$A_2=5{,}41$	$A_5=3{,}84$
0	0,0	0,0	0,0
0,5	0,0	0,0	2,1
1	0,1	0,4	14,0
1,5	0,4	2,1	29,9
2	1,3	5,1	41,3
2,5	2,4	8,7	46,2
3	3,7	12,1	47,0
3,5	4,8	14,7	46,0
4	5,5	16,3	44,7
4,5	6,1	17,1	43,8
5	6,5	17,3	43,5
5,5	6,8	17,1	43,6
6	7,0	16,9	44,2
6,5	7,3	16,7	45,1
7	7,5	16,7	46,2
7,5	7,6	16,9	47,3
8	7,8	17,1	48,5
8,5	7,9	17,4	49,5
9	8,0	17,8	50,5
9,5	8,1	18,1	51,3
10	8,2	18,3	51,9
10,5	8,3	18,6	52,4
11	8,4	18,8	52,7
11,5	8,6	19,0	53,0
12	8,7	19,1	53,0
12,5	8,8	19,3	53,0
13	8,8	19,5	52,9
13,5	8,9	19,6	52,7
14	9,0	19,8	52,5
14,5	9,1	20,0	52,2
15	9,1	20,0	51,9
15,5	9,2	20,1	51,5
16	9,2	20,2	51,2
16,5	9,3	20,3	50,8
17	9,3	20,3	50,5
17,5	9,4	20,3	50,2
18	9,4	20,3	50,0
18,5	9,5	20,3	49,8
19	9,5	20,2	49,6
19,5	9,6	20,2	49,5
20	9,6	20,1	49,4

Man sieht aus der Tabelle, daß in dem untersuchten Bereich von $w = 0$ bis $w = 20$ die Irrtumswahrscheinlichkeiten sich nie weit über die asymptotischen Werte erheben. Der Maximum von $10^3 P(w)$ in diesem Bereich ist in den drei untersuchten Fällen

$$9{,}6;\ 20{,}3;\ 53{,}0\,.$$

Es ist also nicht nötig, bei der Anwendung des χ^2-Testes zu verlangen, daß der gemeinsame Erwartungswert w der beiden Klassen genügend groß sein soll. Bei kleinen Werten von w sind die Irrtumswahrscheinlichkeiten sogar erheblich kleiner als die asymtotischen Werte; man bleibt also bei kleinem w auf der sicheren Seite.

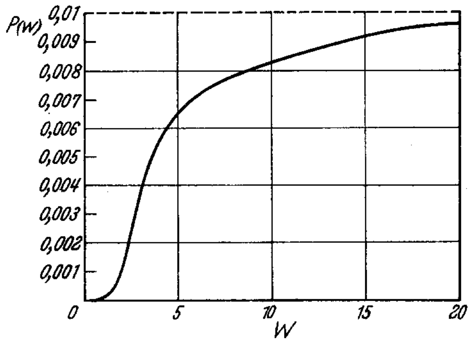

Fig. 1

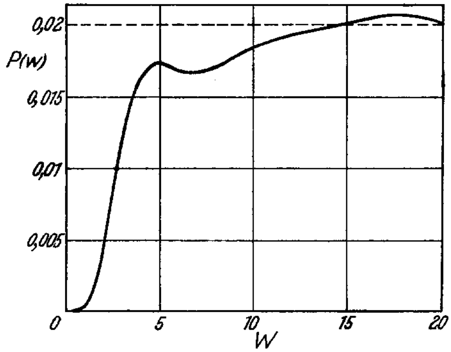

Fig. 2

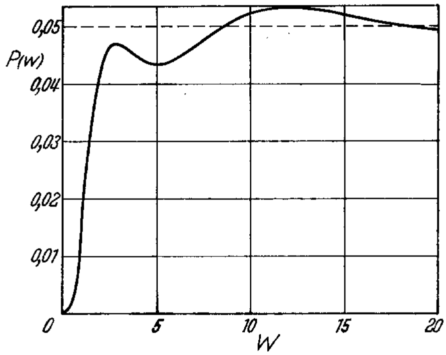

Fig. 3

Fig. 1
Irrtumswahrscheinlichkeit $P(w)$ auf dem 1%-Niveau

Fig. 2
Irrtumswahrscheinlichkeit $P(w)$ auf dem 2%-Niveau

Fig. 3
Irrtumswahrscheinlichkeit $P(w)$ auf dem 5%-Niveau

Literatur

[1] Cochran, W. G.: The χ^2-test of goodness of fit. Ann. math. Statistics **23**, 315—345 (1952).
[2] Neyman, J., and E. S. Pearson: On the use and interpretation of test criteria. Biometrica **20** A, 175—240 und 263—294 (1928).
[3] Studer, E.: Diss. Zürich. Wird demnächst veröffentlicht.
[4] Van der Waerden, B. L.: Mathematische Statistik. § 9 C. Berlin-Göttingen-Heidelberg: Springer 1957.
[5] Ber. Verh. sächs. Akad. Wiss. Leipzig **95**, 145—150 (1943).

Mathematisches Institut
der Universität Zürich
Zürich (Schweiz)

(Eingegangen am 24. September 1964)