

Biometrieübung 11

Kontingenztafeln

Aufgabe

1. 2x2-Kontingenztafeln

100 weibliche Patienten sind mit einer konventionellen Therapie behandelt worden. 85 Patientinnen wurden geheilt, 15 sind gestorben. Von 81 Patientinnen, die mit einer neuen Therapie behandelt wurden, konnten 77 geheilt entlassen werden. 4 sind gestorben.

- Stellen Sie aus den genannten Häufigkeiten eine 2x2-Kontingenztafel auf.
- Wie groß sind die erwarteten Häufigkeiten, wenn diese proportional zu den Randsummen sein sollen ?
- Ist der Heilungsprozentsatz stochastisch unabhängig von der angewandten Therapie (Nullhypothese) ?

Bei männlichen Patienten wurden folgende Zahlen ermittelt:

Konventionelle Therapie: 85 geheilt, 20 gestorben

neue Therapie: 70 geheilt, 6 gestorben

- Wiederholen Sie die Lösung der Aufgaben a bis c für die männlichen Patienten.
- Berechnen Sie für die weiblichen und für die männlichen Patienten jeweils den Kontingenzkoeffizienten. Ist die Abhängigkeit des Heilungserfolges von der Therapie bei den weiblichen oder bei den männlichen Patienten größer ?
- Unter welcher Bedingung erreicht der Kontingenzkoeffizient seinen größten Wert ?

2. kx2-Kontingenztafeln

Bei 300 Personen wurden Geschlecht und Haarfarbe notiert. Beim Geschlecht wurden weiblich und männlich unterschieden, bei der Haarfarbe schwarz, braun, blond und rot.

Folgende Häufigkeiten wurden gefunden:

weiblich/schwarz	55
männlich/schwarz	32
weiblich/braun	65
männlich/braun	43
weiblich/blond	64
männlich/blond	16
weiblich/rot	16
männlich/rot	9

- a. Stellen Sie aus den genannten Häufigkeiten eine $k \times 2$ - oder $2 \times k$ -Kontingenztafel auf.
 - b. Wie groß sind die erwarteten Häufigkeiten, wenn diese proportional zu den Randsummen sein sollen ?
 - c. Ist die Haarfarbe stochastisch unabhängig von Geschlecht (Nullhypothese) ?
 - d. Welche Haarfarbe(n) ist (sind) vom Geschlecht abhängig ?
-

Letzte Änderung: 27.02.1999

Kontakt: [Wolfgang Stümer](#)



Biometrieübung 11

Kontingenztafeln

Lösung

1. 2x2-Kontingenztafeln

a. 2x2-Kontingenztafel

weiblich Patienten	geheilt	gestorben	Insgesamt
konventionelle Therapie	85	15	100
neue Therapie	77	4	81
Insgesamt	162	19	181

b. erwartete Häufigkeiten

weiblich Patienten	geheilt	gestorben
konventionelle Therapie	89,5	10,5
neue Therapie	72,5	8,5

c. Test

Nullhypothese: Der Heilungsprozentsatz ist stochastisch unabhängig von der angewandten Therapie.

Alternativhypothese: Der Heilungsprozentsatz ist abhängig von der angewandten Therapie.

$$\hat{\chi}^2 = 4,82$$

$$\alpha = 0,05$$

$$FG = 1$$

$$\chi_{FG,\alpha}^2 = 3,84$$

Da $\hat{\chi}^2 > \chi_{FG,\alpha}^2$ ist, wird die Nullhypothese abgelehnt. Der Heilung bei den weiblichen Patienten ist von der angewandten Therapie abhängig.

d. männliche Patienten

2x2-Kontingenztafel

männliche Patienten	geheilt	gestorben	Insgesamt
---------------------	---------	-----------	-----------

konventionelle Therapie	85	20	105
neue Therapie	70	6	76
Insgesamt	155	26	181

erwartete Häufigkeiten

männliche Patienten	geheilt	gestorben
konventionelle Therapie	89,9	15,1
neue Therapie	65,1	10,9

Test

Nullhypothese: Der Heilungsprozentsatz ist stochastisch unabhängig von der angewandten Therapie.

Alternativhypothese: Der Heilungsprozentsatz ist abhängig von der angewandten Therapie.

$$\hat{\chi}^2 = 4,46$$

$$\alpha = 0,05$$

$$FG = 1$$

$$\chi_{FG, \alpha}^2 = 3,84$$

Da $\hat{\chi}^2 > \chi_{FG, \alpha}^2$ ist, wird die Nullhypothese abgelehnt. Der Heilung bei den männlichen Patienten ist von der angewandten Therapie abhängig.

e. Kontingenzkoeffizienten

	Kontingenzkoeffizient (CC)	korrigierter Kontingenzkoeffizient (CC _{korr})
weibliche Patienten	0,1611	0,2278
männliche Patienten	0,1550	0,2193

Die Abhängigkeit des Heilungserfolges von der Therapie ist bei den weiblichen Patienten größer (0,1611 > 0,1550). Aber der Unterschied bei den Kontingenzkoeffizienten der weiblichen und männlichen Patienten ist so gering, daß man ihn vernachlässigen kann.

f. maximaler Kontingenzkoeffizient

Der maximale Kontingenzkoeffizient für die Vierfeldertafel beträgt 0,7071; er tritt stets bei vollkommener Kontingenz auf, also wenn die entgegengesetzten Komplementärereignisse unbesetzt bleiben. In unserem

Beispiel müßten bei der konventionellen Therapie alle Patienten sterben und bei der neuen Therapie alle Patienten geheilt werden (oder umgekehrt).

2. kx2-Kontingenztafeln

a. 4x2-Kontingenztafel

Haarfarbe	weiblich	männlich	Insgesamt
schwarz	55	32	87
braun	65	43	108
blond	64	16	80
rot	16	9	25
Insgesamt	200	100	300

b. erwartete Häufigkeiten

Haarfarbe	weiblich	männlich
schwarz	58,0	29,0
braun	72,0	36,0
blond	53,3	26,7
rot	16,7	8,3

c. Test

Nullhypothese: Die Haarfarbe ist stochastisch unabhängig vom Geschlecht .

Alternativhypothese: Die Haarfarbe ist abhängig vom Geschlecht .

$$\hat{\chi}^2 = 8,99$$

$$\alpha = 0,05$$

$$FG = k-1 = 3$$

$$\chi_{FG,\alpha}^2 = 7,81$$

Da $\hat{\chi}^2 > \chi_{FG,\alpha}^2$ ist, wird die Nullhypothese abgelehnt. Die Haarfarbe ist abhängig vom Geschlecht .

d. Welche Haarfarben sind abhängig?

2x2-Kontingenztafel (schwarz)

Haarfarbe	weiblich	männlich	Insgesamt
schwarz	55	32	87
braun/blond/rot	145	68	213
Insgesamt	200	100	300

erwartete Häufigkeiten (schwarz)

Haarfarbe	weiblich	männlich
schwarz	58,0	29,0
braun/blond/rot	142,0	71,0

Test (schwarz)

Nullhypothese: Die Haarfarbe schwarz ist stochastisch unabhängig vom Geschlecht .

Alternativhypothese: Die Haarfarbe schwarz ist abhängig vom Geschlecht .

$$\hat{\chi}^2 = 0,66$$

$$\alpha = 0,05$$

$$FG = 1$$

$$\chi_{FG,\alpha}^2 = 3,84$$

Da $\hat{\chi}^2 < \chi_{FG,\alpha}^2$ ist, wird die Nullhypothese **nicht** abgelehnt. Die Haarfarbe schwarz ist stochastisch unabhängig vom Geschlecht .

2x2-Kontingenztafel (braun)

Haarfarbe	weiblich	männlich	Insgesamt
braun	65	43	108
schwarz/blond/rot	135	57	192
Insgesamt	200	100	300

erwartete Häufigkeiten (braun)

Haarfarbe	weiblich	männlich
braun	72,0	36,0
schwarz/blond/rot	128,0	64,0

Test (braun)

Nullhypothese: Die Haarfarbe braun ist stochastisch unabhängig vom Geschlecht .

Alternativhypothese: Die Haarfarbe braun ist abhängig vom Geschlecht .

$$\hat{\chi}^2 = 3,19$$

$$\alpha = 0,05$$

$$FG = 1$$

$$\chi_{FG,\alpha}^2 = 3,84$$

Da $\hat{\chi}^2 < \chi_{FG,\alpha}^2$ ist, wird die Nullhypothese **nicht** abgelehnt. Die Haarfarbe braun ist stochastisch unabhängig vom Geschlecht .

2x2-Kontingenztafel (blond)

Haarfarbe	weiblich	männlich	Insgesamt
blond	64	16	80
schwarz/braun/rot	136	84	220
Insgesamt	200	100	300

erwartete Häufigkeiten (blond)

Haarfarbe	weiblich	männlich
blond	53,3	26,7
schwarz/braun/rot	146,7	73,3

Test (blond)

Nullhypothese: Die Haarfarbe blond ist stochastisch unabhängig vom Geschlecht .

Alternativhypothese: Die Haarfarbe blond ist abhängig vom Geschlecht .

$$\hat{\chi}^2 = 8,73$$

$$\alpha = 0,05$$

$$FG = 1$$

$$\chi_{FG,\alpha}^2 = 3,84$$

Da $\hat{\chi}^2 > \chi_{FG,\alpha}^2$ ist, wird die Nullhypothese abgelehnt. Die Haarfarbe blond ist abhängig vom Geschlecht .

2x2-Kontingenztafel (rot)

Haarfarbe	weiblich	männlich	Insgesamt
rot	16	9	25
schwarz/braun/blond	184	91	275
Insgesamt	200	100	300

erwartete Häufigkeiten (rot)

Haarfarbe	weiblich	männlich
rot	16,7	8,3
schwarz/braun/blond	183,3	91,7

Test (rot)

Nullhypothese: Die Haarfarbe rot ist stochastisch unabhängig vom Geschlecht .

Alternativhypothese: Die Haarfarbe rot ist abhängig vom Geschlecht .

$$\hat{\chi}^2 = 0,09$$

$$\alpha = 0,05$$

$$FG = 1$$

$$\chi_{FG,\alpha}^2 = 3,84$$

Da $\hat{\chi}^2 < \chi_{FG,\alpha}^2$ ist, wird die Nullhypothese **nicht** abgelehnt. Die Haarfarbe rot ist stochastisch unabhängig vom Geschlecht .

Nur die Haarfarbe blond ist abhängig vom Geschlecht. Bei den Farben schwarz, braun und rot besteht keine Abhängigkeit vom Geschlecht.

Letzte Änderung: 27.08.1999

Kontakt: [Wolfgang Stümer](#)



Biometrieübung 11

Kontingenztafeln

Formeln

Inhalt

[Auswertung von Vierfeldertafeln \(2x2 Kontingenztafeln\)](#)

[Tabelle der Signifikanzschranken der Chi-Quadrat-Verteilung](#)

Auswertung von Vierfeldertafeln (2x2 Kontingenztafeln)

- Analyse kategorialer Daten:
- Neben Alternativmerkmalen mit nur zwei Ausprägungen (z.B. einwandfrei – nicht einwandfrei) unterscheidet man nominal skalierte Merkmale, wobei Namen oder Bezeichnungen für bestimmte Kategorien existieren (z.B. ledig, verheiratet, geschieden) sowie ordinal skalierte Merkmale, bei denen die Kategorien aus Namen oder Bezeichnungen für Ränge, d.h. für Intensitätsstufen des Merkmals bestehen (z.B. die Benotungsskala in der Schule).
- Werden die Ausprägungen zweier kategorialer Merkmale gemeinsam betrachtet, so liegt im einfachsten Fall eine Vierfeldertafel vor. Ziel der Analyse einer Vier- oder Mehrfeldertafel ist die Erfassung der Beziehungsstruktur der in der Tafel verknüpften kategorialen Merkmale. Wird die stochastische Unabhängigkeit beider Merkmale geprüft, so liegt eine Kontingenztafel vor.

Merkmalspaar II	Ereignisse (+)	Komplementär- ereignisse (-)	Insgesamt
Merkmalspaar I			
Erste Stichprobe	a	b	$a + b = n_1$
Zweite Stichprobe	c	d	$c + d = n_2$
Insgesamt	$a + c$	$b + d$	$n_1 + n_2 = n$

Die beiden Stichproben von Alternativdaten werden daraufhin untersucht, ob sie als Zufallsstichproben aus einer durch die Randsummen repräsentierten Grundgesamtheit aufgefaßt werden können. Ob Abweichungen der Verhältnisse $\mathbf{a/n_1}$ und $\mathbf{c/n_2}$ von dem Verhältnis $\mathbf{(a + c)/n}$ (Nullhypothese der Homogenität: $a/n_1 = c/n_2 = (a + c)/n$) als Zufallsabweichung auffaßbar sind.

Verteilen sich die Felderhäufigkeiten proportional zu den Randsummen?

Um dies zu entscheiden, bestimmen wir die unter dieser Annahme zu erwartenden Häufigkeiten, kurz

Erwartungshäufigkeiten E genannt.

Wir multiplizieren die Zeilensumme mit der Spaltensumme des Feldes und dividieren das Produkt durch den Umfang n der vereinten Stichproben.

z.B.

$$E_a = \frac{n_1 * (a + c)}{n}$$

Nullhypothese: Beide Stichproben entstammen einer gemeinsamen Grundgesamtheit mit der Erfolgswahrscheinlichkeit π .

Alternativhypothese: Beide Stichproben entstammen zwei verschiedenen Grundgesamtheiten mit den Erfolgswahrscheinlichkeiten π_1 und π_2 .

Erfolgswahrscheinlichkeit π ist das Ereignis dividiert durch die Zeilensumme.

Die Nullhypothese auf Gleichheit oder Homogenität beider Parameter (oder auf stochastische

Unabhängigkeit beider Merkmalsalternativen) wird anhand des χ^2 - Tests nicht abgelehnt oder abgelehnt.

$$\hat{\chi}^2 = \frac{(a - E_a)^2}{E_a} + \frac{(b - E_b)^2}{E_b} + \frac{(c - E_c)^2}{E_c} + \frac{(d - E_d)^2}{E_d}$$

durch Umformungen erhält man:

$$\hat{\chi}^2 = \frac{n(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}$$

Für kleines n ist n durch $(n-1)$ zu ersetzen.

$$\hat{\chi}_*^2 = \frac{(n - 1)(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}$$

Das Vierfelder- χ^2 besitzt nur einen Freiheitsgrad, da bei gegebenen Randsummen nur eine der 4 Häufigkeiten frei gewählt werden kann.

Die Nullhypothese wird abgelehnt, sobald das berechnete $\hat{\chi}^2$ größer ist als der Tabellenwert $\chi_{1,\alpha}^2$

χ^2 -Tests können lediglich die Existenz eines Zusammenhanges aufzeigen. Über die Stärke des Zusammenhanges sagen sie nichts aus. Ein Maßzahl für den Grad des Zusammenhanges ist der Kontingenzkoeffizient. Er ist als Maß für die Straffheit des Zusammenhanges der beiden Merkmale von

Vier- und Mehrfeldertafeln aus dem $\hat{\chi}^2$ - Wert nach der Formel zu berechnen:

$$CC = \sqrt{\frac{\hat{\chi}^2}{n + \hat{\chi}^2}} \quad CC_{\text{korr}} = \frac{CC}{CC_{\text{max}}}$$

Der maximale Kontingenzkoeffizient für die Vierfeldertafel beträgt 0,7071; er tritt stets bei vollkommener Kontingenz auf, also wenn die Felder b und c unbesetzt bleiben.

Vergleich mehrerer Stichproben von Alternativdaten (k*2 – Felder – Kontingenztafel)

Vorausgesetzt werden n unabhängige Beobachtungen.

Stichprobe (oder 2. Merkmal)	Merkmal (+)	Merkmal (-)	Σ
1	x_1	$n_1 - x_1$	n_1
2	x_2	$n_2 - x_2$	n_2
·	·	·	·
j	x_j	$n_j - x_j$	n_j
·	·	·	·
k	x_k	$n_k - x_k$	n_k
Σ	x	n-x	n

Zur Entscheidung über Beibehalten oder Ablehnung der Nullhypothese der Homogenität (Gleichheit) k

binomialer Grundgesamtheiten dient der χ^2 - Test.

$$\hat{\chi}^2 = \frac{n^2}{x(n-x)} \left[\sum_{j=1}^k \frac{x_j^2}{n_j} - \frac{x^2}{n} \right] \quad \text{mit FG} = k - 1$$

n = Umfang der gesamten Stichproben, die "Ecksumme"

n_j = Umfang der einzelnen Stichproben **j**

x = Gesamtzahl der Stichprobenelemente mit dem Merkmal "+"

x_j = Häufigkeit des Merkmals "+" in der Stichprobe **j**



Biometrieübung 11

Kontingenztafeln

Tabelle der Signifikanzschranken der Chi-Quadrat-Verteilung

FG	Signifikanzniveau			
	0,05	0,025	0,01	0,001
1	3,84	5,02	6,63	10,83
2	5,99	7,38	9,21	13,82
3	7,81	9,35	11,34	16,27
4	9,49	11,14	13,28	18,47
5	11,07	12,83	15,09	20,51
6	12,59	14,45	16,81	22,46
7	14,07	16,01	18,48	24,32
8	15,51	17,53	20,09	26,12
9	16,92	19,02	21,67	27,88
10	18,31	20,48	23,21	29,59
11	19,68	21,92	24,73	31,26
12	21,03	23,34	26,22	32,91
13	22,36	24,74	27,69	34,53
14	23,68	26,12	29,14	36,12
15	25,00	27,49	30,58	37,70
16	26,30	28,85	32,00	39,25
17	27,59	30,19	33,41	40,79
18	28,87	31,53	34,81	42,31
19	30,14	32,85	36,19	43,82
20	31,41	34,17	37,57	45,31
21	32,67	35,48	38,93	46,80
22	33,92	36,78	40,29	48,27
23	35,17	38,08	41,64	49,73
24	36,42	39,36	42,98	51,18
25	37,65	40,65	44,31	52,62
26	38,89	41,92	45,64	54,05
27	40,11	43,19	46,96	55,48
28	41,34	44,46	48,28	56,89
29	42,56	45,72	49,59	58,30
30	43,77	46,98	50,89	59,70

31	44,99	48,23	52,19	61,10
32	46,19	49,48	53,49	62,49
33	47,40	50,73	54,78	63,87
34	48,60	51,97	56,06	65,25
35	49,80	53,20	57,34	66,62
36	51,00	54,44	58,62	67,98
37	52,19	55,67	59,89	69,35
38	53,38	56,90	61,16	70,70
39	54,57	58,12	62,43	72,06
40	55,76	59,34	63,69	73,40
41	56,94	60,56	64,95	74,74
42	58,12	61,78	66,21	76,08
43	59,30	62,99	67,46	77,42
44	60,48	64,20	68,71	78,75
45	61,66	65,41	69,96	80,08
46	62,83	66,62	71,20	81,40
47	64,00	67,82	72,44	82,72
48	65,17	69,02	73,68	84,04
49	66,34	70,22	74,92	85,35
50	67,50	71,42	76,15	86,66
51	68,67	72,62	77,39	87,97
52	69,83	73,81	78,62	89,27
53	70,99	75,00	79,84	90,57
54	72,15	76,19	81,07	91,87
55	73,31	77,38	82,29	93,17
56	74,47	78,57	83,51	94,46
57	75,62	79,75	84,73	95,75
58	76,78	80,94	85,95	97,04
59	77,93	82,12	87,17	98,32
60	79,08	83,30	88,38	99,61
61	80,23	84,48	89,59	100,89
62	81,38	85,65	90,80	102,17
63	82,53	86,83	92,01	103,44
64	83,68	88,00	93,22	104,72
65	84,82	89,18	94,42	105,99
66	85,96	90,35	95,63	107,26
67	87,11	91,52	96,83	108,53
68	88,25	92,69	98,03	109,79

69	89,39	93,86	99,23	111,06
70	90,53	95,02	100,43	112,32
71	91,67	96,19	101,62	113,58
72	92,81	97,35	102,82	114,83
73	93,95	98,52	104,01	116,09
74	95,08	99,68	105,20	117,35
75	96,22	100,84	106,39	118,60
76	97,35	102,00	107,58	119,85
77	98,48	103,16	108,77	121,10
78	99,62	104,32	109,96	122,35
79	100,75	105,47	111,14	123,59
80	101,88	106,63	112,33	124,84
81	103,01	107,78	113,51	126,08
82	104,14	108,94	114,69	127,32
83	105,27	110,09	115,88	128,56
84	106,39	111,24	117,06	129,80
85	107,52	112,39	118,24	131,04
86	108,65	113,54	119,41	132,28
87	109,77	114,69	120,59	133,51
88	110,90	115,84	121,77	134,75
89	112,02	116,99	122,94	135,98
90	113,15	118,14	124,12	137,21
91	114,27	119,28	125,29	138,44
92	115,39	120,43	126,46	139,67
93	116,51	121,57	127,63	140,89
94	117,63	122,72	128,80	142,12
95	118,75	123,86	129,97	143,34
96	119,87	125,00	131,14	144,57
97	120,99	126,14	132,31	145,79
98	122,11	127,28	133,48	147,01
99	123,23	128,42	134,64	148,23
100	124,34	129,56	135,81	149,45
101	125,46	130,70	136,97	150,67
102	126,57	131,84	138,13	151,88
103	127,69	132,97	139,30	153,10
104	128,80	134,11	140,46	154,31
105	129,92	135,25	141,62	155,53
106	131,03	136,38	142,78	156,74

107	132,14	137,52	143,94	157,95
108	133,26	138,65	145,10	159,16
109	134,37	139,78	146,26	160,37
110	135,48	140,92	147,41	161,58
111	136,59	142,05	148,57	162,79
112	137,70	143,18	149,73	164,00
113	138,81	144,31	150,88	165,20
114	139,92	145,44	152,04	166,41
115	141,03	146,57	153,19	167,61
116	142,14	147,70	154,34	168,81
117	143,25	148,83	155,50	170,01
118	144,35	149,96	156,65	171,22
119	145,46	151,08	157,80	172,42
120	146,57	152,21	158,95	173,62
121	147,67	153,34	160,10	174,81
122	148,78	154,46	161,25	176,01
123	149,88	155,59	162,40	177,21
124	150,99	156,71	163,55	178,41
125	152,09	157,84	164,69	179,60
126	153,20	158,96	165,84	180,80
127	154,30	160,09	166,99	181,99
128	155,40	161,21	168,13	183,19
129	156,51	162,33	169,28	184,38
130	157,61	163,45	170,42	185,57
131	158,71	164,57	171,57	186,76
132	159,81	165,70	172,71	187,95
133	160,91	166,82	173,85	189,14
134	162,02	167,94	175,00	190,33
135	163,12	169,06	176,14	191,52
136	164,22	170,18	177,28	192,71
137	165,32	171,29	178,42	193,89
138	166,42	172,41	179,56	195,08
139	167,51	173,53	180,70	196,26
140	168,61	174,65	181,84	197,45
141	169,71	175,76	182,98	198,63
142	170,81	176,88	184,12	199,82
143	171,91	178,00	185,26	201,00
144	173,00	179,11	186,39	202,18

145	174,10	180,23	187,53	203,37
146	175,20	181,34	188,67	204,54
147	176,29	182,46	189,80	205,73
148	177,39	183,57	190,94	206,91
149	178,49	184,69	192,07	208,08
150	179,58	185,80	193,21	209,27

[zurück](#)

Letzte Änderung: 20.09.1999

Kontakt: [Wolfgang Stümer](#)

