On the Application of the χ^2 Method to Association and Contingency Tables, with Experimental Illustrations

Author(s): G. Udny Yule

Source: *Journal of the Royal Statistical Society*, Vol. 85, No. 1 (Jan., 1922), pp. 95-104
Published by: Wiley for the Royal Statistical Society
Stable URL: https://www.jstor.org/stable/2340522
Accessed: 20-04-2020 16:28 UTC

On the Application of the χ^2 Method to Association and Contingency Tables, with Experimental Illustrations.

By G. Udny Yule, C.B.E., M.A., F.R.S.

In a paper* published in the *Proceedings of the Royal Society of Medicine* in 1915, Dr. Greenwood and I directed attention to the discrepancy between the results given by the χ^2 method, as ordinarily applied to the fourfold table for estimating the probability that any given divergence from independence might have arisen by random sampling, and the results given by the more elementary test afforded by comparing $p_1 - p_2$ with its probable error, where in my notation for the cell-frequencies

$$p_1 = (\mathrm{AB})/(\mathrm{B}) \qquad p_2 = (\mathrm{A}\beta)/(\beta) \tag{1}$$

Invariably the probability of any arrangement, or of any less probable arrangement, having arisen by random sampling was greater when estimated by the χ^2 method than when the other test was applied. At the time of publication we were unable to clear up the source of the discrepancy, and as it was evident that, in judging data by the χ^2 test, we should certainly not attach importance to results which might be the mere effect of random sampling and were thus erring on the side of caution, we passed over the difficulty for the time. Later work carried us, however, towards the result given in the preceding note by Mr. Fisher. Dr. Greenwood and I were in correspondence on the general question, and on points arising out of it, for many months afterwards, and while I am alone responsible for the present note, having carried out the bulk of the experiments described, it was prompted entirely by our joint work.

Let δ denote the difference between the frequency (AB) and its independence-value (A)(B)/N. Then, as I showed in a former paper, the standard error of δ, ϵ_δ, is given in general by the equation†

$$\epsilon_\delta^2 = \frac{1}{\mathrm{N}^3}\left\{(\mathrm{A})(\alpha)(\mathrm{B})(\beta) + \mathrm{N}\delta[(\mathrm{A})-(\alpha)][(\mathrm{B})-(\beta)] - \mathrm{N}^2\delta^2\right\} \tag{2}$$

* "The Statistics of Anti-typhoid and Anti-cholera Inoculations, etc.," *Proc. Roy. Soc. Medicine*, vol viii, 1915, pp. 113–190.

† "On the Methods of Measuring Association between Two Attributes," *J.S.S.*, vol. lxxv, 1912, equation (39), p. 602.

If in this equation we make δ zero, so as to obtain the standard error of δ for sampling from a universe in which A and B are independent, we have the simplified form—

$$_0\epsilon_\delta^2 = \frac{1}{N^3}\{(A)\ (\alpha)\ (B)\ (\beta)\} \qquad (3)$$

But

$$\chi^2 = N\delta^2 \left\{ \frac{1}{(A)\ (B)} + \frac{1}{(A)\ (\beta)} + \frac{1}{(\alpha)\ (B)} + \frac{1}{(\alpha)\ (\beta)} \right\}$$
$$= \frac{N^3\delta^2}{(A)\ (\alpha)\ (B)\ (\beta)} \qquad (4)$$

Therefore we may write

$$_0\epsilon_\delta^2 = \frac{\delta^2}{\chi^2}$$

And therefore, if the number of observations is sufficiently large to enable us to regard the distribution of errors as normal, the frequency distribution of δ is given by

$$y = y_0 e^{-\frac{1}{2}\chi^2} \qquad (5)$$

Hence the probability of obtaining a value of δ exceeding, without respect to sign, the value corresponding to any assigned value of χ is

$$P = \sqrt{\frac{2}{\pi}}\int_\chi^\infty e^{-\frac{1}{2}\chi^2}d\chi \qquad (6)$$

But this is the expression for P corresponding to $n' = 2$, not $n' = 4$, in Professor Pearson's notation ; the value 4 is right when divergence is measured from a set of frequencies given *a priori*, but when divergence is measured from independence-values determined from the observations themselves it must be taken as 2. And the reason is fairly obvious. For a given number of observations N the number of algebraically independent cell-frequencies that can be given *a priori* is 3 (three cells of the table can be filled with any arbitrary figures, of which the total is not greater than N, and the fourth is then determined by N) ; but only *one* arbitrary value can be assigned to δ and the frequencies are then determined. The value of n' in each case exceeds by unity the number of *algebraically independent* data. Entering the tables with argument $n' = 4$ appears to represent an attempt to answer a different question, the value (A) (B)/N of the sample being taken as an approximation to the "independence value" in the universe. But in my view (*see* below) the approximation is biassed, and the results, consequently, more or less misleading.

This stage in the argument was reached not long after the publication of the paper of 1915, and Dr. Greenwood and I proceeded to check the result by a short series of experiments. Thirty fourfold tables of a hundred observations each were made by him by dice-throwing; ignoring 5's and 6's, throws of 1's, 2's, 3's and 4's were assigned to the compartments of the table so numbered. Another hundred tables of a hundred observations each were made by me by constructing a rough circular tray of millboard, with partitions dividing it into quadrants, mounting it on a vertical axis so that it could be rapidly spun, throwing a hundred counters into it while spinning, and then counting up the number in each quadrant. For each of these tables the value of χ^2 could then be computed: (*a*) from the *a priori* expectation of equal distribution, (*b*) from the independence-values given by the totals of rows and columns. The results were in very good consonance with theory. For the divergence from expectation the distribution was given by the values of P corresponding to $n' = 4$; for the divergence from independence, by the values of P corresponding to $n' = 2$. The two distributions were entirely different.

But this soon *suggested* the general result. For a table with r rows and c columns n' should be taken as rc only for judging the probability of a given divergence from *a priori* frequencies. When it is desired to determine the probability of a given divergence from the independence-values calculated from the totals of rows and columns the number of algebraically independent values of δ is only $(r - 1)(c - 1)$ and n' should be taken as given by

$$n' = (r - 1)(c - 1) + 1.$$

But I failed to get any proof of this result, reasonable though it seemed; the most I succeeded in doing was to reduce the general expression for χ^2 in terms of the correlations and standard deviations, in the case of a table with three rows and three columns. A full proof seems, in fact, still to be lacking. It is not clear that the expression for χ^2 in terms of the standard deviations and correlations of errors must, in the more general case, reduce to the simple form—the sum of the squares of the differences from independence-values each divided by the independence-value. I am glad, however, that Mr. Fisher has now entered the field, as, failing proof, the matter might have stood over for some time. The conjecture and the result of the first experiments had stimulated my curiosity so much that a further series of experiments was devised and carried out during 1916 and 1917. A friend kindly constructed for me a slightly more elaborate apparatus, consisting of a circular wooden tray divided into sixteen sectors by tinplate divisions, spinning on a vertical pivot. A hundred beans were thrown into this while it was spinning, and the numbers counted in each sector. For a large number of the later experiments a simple addition was employed to avoid any risk of regularity in throwing or other formations of habit having any influence on the results, an addition which may be regarded as a spatial-generalization of Galton's apparatus for illustrating the formation of the normal curve. A number of wedges were formed by folding some strips of stiff paper, 6 or 8 inches long (the apparatus has been lost during post-war removals, so that actual dimensions cannot be given). These were laid parallel to each other on a table, some half inch or so apart. Another series of strips was then cemented on top of them, at right angles to the first set, another on top of this, and so on; and on top of the whole pyramid of six or eight layers of wedges a short tube of millboard formed the summit. This apparatus could then be held in one hand over the spinning tray, and the beans dropped through the tube. They scattered through the layers of wedges as they fell, and the result could hardly be other than random.

The frequencies thus obtained could be grouped in a number of different ways:—(*a*) Into the sixteen compartments of a table with four rows and four columns; (*b*) into the sixteen compartments of a table with two rows and eight columns; (*c*) into further tables with only two rows and two columns, for adding to the earlier series—or, of course, into other forms, which, however, I did not attempt. The results completely confirmed, it may be said at once, what was at the time little more than a happy and reasonable guess. The results obtained will now be given.

A.—Fourfold Tables.

Two hundred and twenty tables were made from the frequencies given by the later apparatus which, with the 130 tables of the earlier experiments, gave 350 tables altogether, each table having 100 observations. For each table the value of χ^2 was computed (*a*) from the expected distribution given by uniformity, *i.e.* 25 observations to each compartment; (*b*) from the independence-values given by the row and column totals, *i.e.* equation (4). The theoretical distribution of χ^2 is given by differencing the table showing P in terms of χ^2; for example, in the case of $n' = 4$ we can turn up Table XII on p. 26 of *Tables for Statisticians and Biometricians*, and find the values of P for the first few values of χ^2 as below. The differences are as shown. We should then expect the fraction 0·198 . . . of all observed values of χ^2 (calculated for divergence from uniformity) to lie between 0 and 1 : 0·228 . . . of all values between 1 and 2: 0·180 . . . between 2 and 3, and so on.

χ^2.	P.	Δ.
0	1·	0·198747
1	0·801253	0·228846
2	0·572407	0·180782
3	0·391625	0·130161
4	0·261464	

A column showing the values of P corresponding to assigned values of χ^2 for $n' = 2$ is not given in Table XII of *Tables for Statisticians*, that table not being carried below $n' = 3$. The necessary function, equation (6), will, however, be found tabulated for values of χ^2 proceeding by units in Table XV, p. 30, of the same volume. Unfortunately this table is not sufficiently detailed for present purposes, as regards the earlier part of the range, and I have calculated a more detailed table which is given in the Appendix. A description of the calculation of this table will be found below.

The following table summarizes the results for the case when χ^2 is calculated from the expectation of uniform distribution, and consequently n' is 4.

TABLE I.—*The theoretical frequency distribution of values of χ^2 for a fourfold table, when χ^2 is calculated from the* a priori *expectation of uniformly distributed frequency, compared with the actual results for* 350 *experimental tables.*

Value of χ^2.	Number of tables giving a value of χ^2 between the limits on the left.	
	Expected.	Observed.
0— 1	69·56	62
1— 2	80·10	75
2— 3	63·27	72
3— 4	45·56	47·5
4— 5	31·38	27·5
5— 6	21·07	23
6— 7	13·90	19
7— 8	9·06	9
8— 9	5·85	6
9—10	3·75	2·5
10—	6·50	6·5
Total	350·00	350

The agreement is on the whole very satisfactory. Taking the table as it stands and testing the agreement between theory and observation by the χ^2 method, n' is 11, χ^2 works out at 5·32, and

P is 0·867, so that we might expect a worse fit eight or nine times in ten.

When we pass to the case where χ^2 is calculated for divergence from the independence-values and n' has to be taken as 2, the distribution is very different. As nearly 70 per cent. of the values of χ^2 so calculated should be less than unity, it is desirable to break up the distribution at the lower end of the scale. I have therefore divided the first unit into fourths. Table II shows the results.

TABLE II.—*The theoretical distribution of values of χ^2 for a fourfold table, when χ^2 is calculated from the independence-values, compared with the actual results for* 350 *experimental tables.*

Value of χ^2.	Number of tables giving a value of χ^2 between the limits on the left.	
	Expected.	Observed.
0· —0·25	134·02	122
0·25—0·50	48·15	54
0·50—0·75	32·56	41
0·75—1·00	24·21	24
1·00—2·00	56·00	62
2 —3	25·91	18
3 —4	13·22	13
4 —5	7·05	6
5 —6	3·86	5
6 —	5·01	5
Total	349·99	350

Again the agreement with theory is excellent. Taking the table as it stands, n' is 10, ^{2} works out at 7·53, and P is therefore 0·583; we would expect a worse agreement nearly six times in ten. It is worth noting how clearly, even if we had no theory to guide us, the experiment alone would exhibit the incorrectness of applying the same law in both cases. Table I shows that, when χ^2 was calculated from uniform distribution, only 62 tables gave a value of χ^2 less than unity, and 43 a value greater than 6. For the case of divergence from independence-values, on the other hand, Table II shows that there were no less than 241 tables with a value of χ^2 less than unity, and only five tables with a value greater than 6. A little consideration shows at once, moreover, the simple reason why the second case tends to give much lower values of χ^2 than the first. If, owing to the chances of sampling, (AB) has too high a value, the frequencies (A) and (B) will also *tend* to be high, consequently (A) (B)/N will also *tend* to be high, and this compensates in some degree for the excess in (AB). The experimental tables illustrate the point very well.

B.—*Tables with Sixteen Compartments.*

Utilizing all the frequencies in the 16 sectors of the later apparatus, (1) we can calculate the value of χ^2 for divergence from the expectation of uniform distribution, (2) we can arrange the frequencies in the compartments of a table with four rows and four columns and calculate χ^2 for the divergence from the independence-values given by the totals of rows and columns, (3) we can arrange the frequencies in the compartments of a table with two rows and eight columns and again calculate χ^2 for the divergence from the independence-values given by the totals of rows and columns. According to the extended theory, the distribution of χ^2 should be given in the first case by taking n' as 16, in the second case by taking n' as 10 ($3 \times 3 + 1$), in the third case by taking n' as 8 ($7 \times 1 + 1$). These were the three cases tested by the experimental results. As the theory assumes that normal distribution can be predicated for the deviations, I judged it hardly fair to deal with tables of only 100 observations, which would give an expectation of only 100/16, or 6·25 observations to each compartment. The original tables were therefore superposed in pairs, so as to give tables of 200 observations each with an expectation of 12·5 observations to the compartment—even so, a figure so small that the assumption of normality can hardly be justified except as the roughest of approximations. Only the first 200 single tables were thus dealt with, giving the round number of 100 tables of 200 observations each. Table III compares the actual distribution in each case with the theoretical distribution.

TABLE III.—*The theoretical distributions of χ^2 in tables with* 16 *compartments compared with the actual distributions given by* 100 *experimental tables in three different cases.*

χ^2	From uniform distribution: $n' = 16$.		From independence: 4×4 table, $n' = 10$.		From independence: 8×2 table, $n' = 8$.	
	Expectation.	Observation.	Expectation.	Observation.	Expectation.	Observation.
0— 5	0·8	— }	16·6	17	34·0	29·5
5—10	17·2	20 }	48·4	44	47·1	56·5
10—15	36·8	36	26·0	32	15·3	10
15—20	27·9	30·5	7·3	6 }	3·0	3 }
20—	17·2	13·5	1·8	1 }	0·6	1 }
	99·9	100	100·1	100	100·0	100
χ^2	For 4 groups as shown by brackets	1·28	—	2·27	—	4·36
P		0·74	—	0·52	—	0·22

Again, the agreement with the extended theory is very satisfactory as shown by the values of χ^2 and P for testing this agreement in the last line of the table. The third case gives the poorest agreement, and even here P for four groups is 0·22, so that we might expect a worse agreement once in some five trials. It is again worth while to emphasise how the experiment alone would clearly indicate the difference between the three cases. The first case yielded no values of χ^2 under 5; the second case yielded 17 tables with a value of χ^2 less than that limit; the third case 29·5 such tables. Again, the first case yielded 80 tables giving a value of χ^2 exceeding 10; the second case 39; the third case only 14.

Having regard to Mr. Fisher's note and the experimental tests here given, there can be little doubt as to the correctness of the present theory. Table XII of *Tables for Statisticians and Biometricians* accordingly requires supplementing by a column for $n' = 2$ for use with fourfold tables. It is true that values of P can fairly readily be obtained from χ by Sheppard's table of the normal integral, but a table in the standard form is a great convenience. As already mentioned, the necessary function is already tabulated for values of χ^2 proceeding by steps of a unit in Table XV, p. 30, of that volume, but that table is insufficiently detailed for practical use. As P varies very rapidly with χ^2 at the commencement of the table, I have broken it up into two parts. For values of χ^2 under unity the interval has been taken as 0·01; for values of χ^2 between 1 and 10 the interval has been taken as 0·1. For values of χ^2 over 10, Table XV in *Tables for Statisticians* is quite sufficient. The table has been calculated (1) by taking the square root of χ^2 so as to obtain χ, and then (2) interpolating for the value of the integral of equation (6) in Mr. Sheppard's table of the normal integral, Table II of *Tables for Statisticians*. Seven digits were retained in the work, but since the last digit, as calculated, may have an error of at least 2, these have been reduced to five for tabulation, a number which is more than sufficient for practical purposes. The original values as calculated to seven decimal places were tested by differencing, and although some small unexplained irregularities remained, the five digits finally retained should be accurate. At the commencement of the table even differencing is inadequate as a test. I have to thank Miss Allen and Miss Newbold, of the Medical Research Council's staff, for further checking over parts of the table.

APPENDIX.

Table of the values of P *for divergence from independence in the fourfold table.*

A.—$\chi^2 = 0$ to $\chi^2 = 1$ by steps of 0·01.

χ^2	P	Δ	χ^2	P	Δ
0	1·00000	7966	0·50	0·47950	436
0·01	0·92034	3280	0·51	0·47514	430
0·02	0·88754	2505	0·52	0·47084	423
0·03	0·86249	2101	0·53	0·46661	418
0·04	0·84148	1842	0·54	0·46243	411
0·05	0·82306	1656	0·55	0·45832	406
0·06	0·80650	1516	0·56	0·45426	400
0·07	0·79134	1404	0·57	0·45026	395
0·08	0·77730	1312	0·58	0·44631	389
0·09	0·76418	1235	0·59	0·44242	384
0·10	0·75183	1169	0·60	0·43858	379
0·11	0·74014	1111	0·61	0·43479	374
0·12	0·72903	1060	0·62	0·43105	369
0·13	0·71843	1015	0·63	0·42736	365
0·14	0·70828	974	0·64	0·42371	360
0·15	0·69854	938	0·65	0·42011	355
0·16	0·68916	905	0·66	0·41656	351
0·17	0·68011	874	0·67	0·41305	346
0·18	0·67137	845	0·68	0·40959	343
0·19	0·66292	820	0·69	0·40616	338
0·20	0·65472	795	0·70	0·40278	334
0·21	0·64677	773	0·71	0·39944	330
0·22	0·63904	752	0·72	0·39614	326
0·23	0·63152	731	0·73	0·39288	322
0·24	0·62421	713	0·74	0·38966	318
0·25	0·61708	696	0·75	0·38648	315
0·26	0·61012	679	0·76	0·38333	311
0·27	0·60333	663	0·77	0·38022	308
0·28	0·59670	648	0·78	0·37714	304
0·29	0·59022	634	0·79	0·37410	301
0·30	0·58388	620	0·80	0·37109	297
0·31	0·57768	607	0·81	0·36812	294
0·32	0·57161	595	0·82	0·36518	291
0·33	0·56566	583	0·83	0·36227	287
0·34	0·55983	572	0·84	0·35940	285
0·35	0·55411	560	0·85	0·35655	281
0·36	0·54851	551	0·86	0·35374	278
0·37	0·54300	540	0·87	0·35096	276
0·38	0·53760	530	0·88	0·34820	272
0·39	0·53230	521	0·89	0·34548	270
0·40	0·52709	512	0·90	0·34278	267
0·41	0·52197	503	0·91	0·34011	264
0·42	0·51694	495	0·92	0·33747	261
0·43	0·51199	487	0·93	0·33486	258
0·44	0·50712	479	0·94	0·33228	256
0·45	0·50233	471	0·95	0·32972	253
0·46	0·49762	463	0·96	0·32719	251
0·47	0·49299	457	0·97	0·32468	248
0·48	0·48842	449	0·98	0·32220	246
0·49	0·48393	443	0·99	0·31974	243
0·50	0·47950	436	1·00	0·31731	241

B.—$\chi^2 = 1$ to $\chi^2 = 10$ by steps of $0 \cdot 1$.

χ^2	P	Δ	χ^2	P	Δ
1·0	0·31731	2304	5·5	0·01902	106
1·1	0·29427	2095	5·6	0·01796	99
1·2	0·27332	1911	5·7	0·01697	94
1·3	0·25421	1749	5·8	0·01603	89
1·4	0·23672	1605	5·9	0·01514	83
1·5	0·22067	1477	6·0	0·01431	79
1·6	0·20590	1361	6·1	0·01352	74
1·7	0·19229	1258	6·2	0·01278	71
1·8	0·17971	1163	6·3	0·01207	66
1·9	0·16808	1078	6·4	0·01141	62
2·0	0·15730	1000	6·5	0·01079	59
2·1	0·14730	929	6·6	0·01020	56
2·2	0·13801	864	6·7	0·00964	52
2·3	0·12937	803	6·8	0·00912	50
2·4	0·12134	749	6·9	0·00862	47
2·5	0·11385	699	7·0	0·00815	44
2·6	0·10686	651	7·1	0·00771	42
2·7	0·10035	609	7·2	0·00729	39
2·8	0·09426	568	7·3	0·00690	38
2·9	0·08858	532	7·4	0·00652	35
3·0	0·08326	497	7·5	0·00617	33
3·1	0·07829	465	7·6	0·00584	32
3·2	0·07364	436	7·7	0·00552	30
3·3	0·06928	408	7·8	0·00522	28
3·4	0·06520	383	7·9	0·00494	26
3·5	0·06137	359	8·0	0·00468	25
3·6	0·05778	337	8·1	0·00443	24
3·7	0·05441	316	8·2	0·00419	23
3·8	0·05125	296	8·3	0·00396	21
3·9	0·04829	279	8·4	0·00375	20
4·0	0·04550	262	8·5	0·00355	19
4·1	0·04288	246	8·6	0·00336	18
4·2	0·04042	231	8·7	0·00318	17
4·3	0·03811	217	8·8	0·00301	16
4·4	0·03594	205	8·9	0·00285	15
4·5	0·03389	192	9·0	0·00270	14
4·6	0·03197	181	9·1	0·00256	14
4·7	0·03016	170	9·2	0·00242	13
4·8	0·02846	160	9·3	0·00229	12
4·9	0·02686	151	9·4	0·00217	12
5·0	0·02535	142	9·5	0·00205	10
5·1	0·02393	134	9·6	0·00195	11
5·2	0·02259	126	9·7	0·00184	10
5·3	0·02133	119	9·8	0·00174	9
5·4	0·02014	112	9·9	0·00165	8
5·5	0·01902	106	10·0	0·00157	8