Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

Contingency Tables Involving Small Numbers and the χ^2 Test.

By F. Yates, B.A.

Introduction.

There has in the past been a good deal of argument as to the appropriate statistical tests of independence for contingency tables, particularly those in which each classification is a simple dichotomy (*i.e.* 2×2 tables). It is probably now almost universally admitted that when the numbers in the various cells are large the χ^2 test, introduced by K. Pearson in 1900,[4] with the vital modification as to degrees of freedom established by R. A. Fisher in 1922,[1] and confirmed by Yule,[5] is the appropriate one. The necessity for this modification has been amply demonstrated both from theoretical considerations[1, 5] and by actual tests on material known to be almost if not quite free from association,[2, 5] and there is no need to discuss the matter further here. Several other forms of test for 2×2 tables have been shown to be equivalent to the χ^2 test.[1]

The χ^2 test is admittedly approximate, for in order to establish the test it is necessary to regard each cell value as normally dis tributed with a variance equal to the expected value, the whole set of values being subject to certain restrictions. The accuracy of this approximation depends on the numbers in the various cells, and in practice it has been customary to regard χ^2 as sufficiently accurate if no cell has an expectancy of less than 5. It is with the question of the applicability of χ^2 to 2×2 contingency tables involving small expectancies that we are directly concerned in this paper.

It was suggested to me by Professor Fisher that the probability of any observed set of values in a 2×2 contingency table with given marginal totals can be exactly determined. The method will be explained in the next section. Armed with the exact distribution the divergence of the χ^2 test in any special case can be tested. It will be shown that although the test as ordinarily applied becomes inaccurate even with moderately small numbers in the cells, a simple modification enables the range of usefulness to be considerably extended.

The problem of testing the independence of contingency tables involving small numbers is of considerable practical importance in many branches of biological experimentation, where supplies of experimental material are limited. Medicine offers an extreme example, where considerations quite other than those of expense serve to limit the number of subjects available for testing a new treatment or method of operation. We have therefore thought it worth while to investigate the errors of the modified test in some detail. The results of this investigation are presented in the form of a table which serves to extend the range of this test still further.

In the case of contingency tables involving more than one degree of freedom, reasons are given for believing that the ordinary χ^2 test is considerably more reliable. As an example the exact distribution is worked out for a 2×3 table, and the agreement with the distribution given by χ^2 is shown to be good.

2×2 *Contingency Tables.*

We will define the table by the following notation :

	A	not A	Total
B	a	b	$N - n$
not B	c	d	n
Total	$N - n'$	n'	N

where $n \leq n' \leq \frac{1}{2}N$. It is easily seen that d can take all integral values from 0 to n. We now propose to find the frequency distribution of these values when the table is regarded as consisting of two samples of $N - n$ and n respectively from a binomial distribution with probability p of A occurring, subject to the restriction that the total of A in both samples shall be $N - n'$.

The probability of obtaining the values a and b in the sample of $N - n$ is

$$\frac{(N-n)!}{a!\,b!} p^a (1-p)^b,$$

and of obtaining the values c and d in the sample of n is

$$\frac{n!}{c!\,d!} p^c (1-p)^d.$$

The combined probability is therefore

$$\frac{(N-n)!\,n!}{a!\,b!\,c!\,d!} p^{a+c} (1-p)^{b+d},$$

where a, b, c, and d are subject to the restrictions

$$a + b = N - n, \text{ and } c + d = n.$$

If the additional restriction that $a + c = N - n'$ or $b + d = n'$ is imposed, only those terms of the combined probability series must be selected which satisfy the restriction. But the factors

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

containing p are now constant, and therefore the required probabilities are equal to

$$\frac{1}{a\,!\,b\,!\,c\,!\,d\,!}\Big/\sum_{d=0}^{d=n}\frac{1}{a\,!\,b\,!\,c\,!\,d\,!}$$

subject to the above restrictions.

The summation is easily performed, for it will be given by the coefficient of p^{a+c} in the expansion of

$$(p+q)^{N-n}\,(p+q)^{n}/(N-n)\,!\,n\,!,$$

i.e. of $p^{N-n'}$ in

$$(p+q)^{N}/(N-n)\,!\,n\,!.$$

This is

$$\frac{N\,!}{n\,!\,n'\,!\,(N-n)\,!\,(N-n')\,!}.$$

The successive probabilities from $d = 0$ to $d = n$ are therefore

$$\frac{(N-n)\,!\,(N-n')\,!}{N\,!\,(N-n-n')\,!},\ \frac{(N-n)\,!\,(N-n')\,!\,n\,.\,n'}{N\,!\,(N-n-n'+1)\,!\,1\,!},$$
$$\frac{(N-n)\,!\,(N-n')\,!\,n(n-1)\,n'(n'-1)}{N\,!\,(N-n-n'+2)\,!\,2\,!},\ \ldots\ \frac{(N-n)\,!\,n'\,!}{N\,!\,(n'-n)\,!}.$$

The successive probabilities are therefore proportional to the terms of the hypergeometric series $F(-n, -n', N-n-n'+1, 1)$.

Alternatively stated the probability corresponding to any term a, b, c, d, is

$$\frac{n\,!\,n'\,!\,(N-n)\,!\,(N-n')\,!}{N\,!\,a\,!\,b\,!\,c\,!\,d\,!},$$

i.e., the product of the factorials of the four marginal totals divided by the product of the factorials of the grand total and the four cell numbers.

In cases where N is not too large the distribution with any particular numerical values of the marginal totals can be computed quite quickly, using a table of factorials to determine some convenient term, and working out the rest of the distribution term by term, by simple multiplications and divisions. If a table of factorials is not available we may start with any convenient term as unity, and divide by the sum of the terms so obtained.

The numerical distribution could be used to provide a direct test of significance, but even when the marginal totals are quite small the evaluation of χ^2 is much more expeditious, and it is therefore of some interest to determine the limits within which χ^2 may be safely employed.

Validity of Assumption of Constancy of Marginal Totals.

No demonstration has yet been given that we are justified in assuming the constancy of the marginal totals. From one point of view the matter is indeed almost obvious, for there is no inherent distinction between the rows and columns of the table, and if the marginal totals of the rows may be assumed fixed, then as an alternative the totals of the columns may equally be assumed fixed. If the two tests of significance arising from these separate assumptions are to be identical, then the test must ultimately involve the fixity of all marginal totals. But since it is perhaps not immediately clear that the tests of significance need be identical it may be worth while to examine the matter a little more closely.

All possible sets of observations having $N - n$ and n as the totals of the rows can be classified according to the value of n'. For each n' a 2·5 per cent. (or other) level of significance can be assigned for each tail (subject to discontinuity) and all sets of observations falling outside these limits will be counted significant. If 5 per cent. of each class is judged significant, then 5 per cent. of the whole population of sets will be judged significant, whatever the differences in the relative frequencies of the various classes. Clearly, therefore, the proposed test of significance is a valid test, in the sense that if there is no association a fixed proportion of the sets of observation will be judged significant, whatever p.

Leaving aside the question of discrimination between the two tails it is clear that when all the marginal totals are fixed the test is efficient, in the sense that if p and p' differ, the probability of obtaining a verdict of significance is as great as possible. For there is then only one degree of freedom, which must give rise to a definite distribution, of which the tails will provide the appropriate regions of significance.

That the marginal totals of the columns are necessarily assumed constant can be established without appeal to symmetry. Instead of regarding the table as generated by two random samples of $N - n$ and n respectively from a population with probability p, we may regard it as formed of a single sample N from such a population. In this sample each observation will be either A or not -A. The observations of class B may then be obtained by a random drawing of $N - n$ from this sample, for if there is no association there can be nothing in the attributes of any single observation which will determine whether it is also B or not -B. The remaining n observations will be consigned to class not -B. By this procedure of constructing the table we secure the fixity of all the marginal totals.

It is interesting to note that the marginal totals are in the nature of ancillary statistics as defined by Fisher. They determine the accuracy of the information supplied, but they are not subject to error or other variation themselves. From what has already been

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

said it will be apparent that there is no need to estimate p in order to make a test of significance. The distribution established in the previous section is independent of the actual value of p.

If a verdict of non-significance is obtained and this is regarded as conclusive, *i.e.* if we are prepared to assume that p and p' are identical, then the efficient (indeed sufficient) estimate of p is given by $(N - n')/N$. If we are not prepared to make this assumption, or if the probabilities are judged significantly different, the sufficient estimates of p and p' are $a/(N - n)$ and c/n, and the interpretation of the difference between them can be determined in the light of their numerical values and the nature of the data.

Binomial Distributions with known p.

The examination of the χ^2 test in the case of simple binomial distributions with known p will illustrate some important points.

We will first consider the symmetrical distribution

$$(\tfrac{1}{2} + \tfrac{1}{2})^{10}.$$

Such a distribution would be obtained, for example, for the number of heads in groups of 10 tosses with a coin.

TABLE I.

Successes.		p'.	P.	$P(\chi)$.	Discrepancy.	$P(\chi')$.	Discrepancy.
0	10	0·0010	0·0010	0·0008	−0·0002	0·0022	+0·0012
1	9	0·0098	0·0108	0·0057	−0·0051	0·0134	+0·0027
2	8	0·0439	0·0547	0·0290	−0·0257	0·0569	+0·0022
3	7	0·1172	0·1719	0·1030	−0·0689	0·1714	−0·0005
4	6	0·2051	0·3770	0·2635	−0·1135	0·3759	−0·0011
5		0·2461	—	—	—	—	—

Table I shows this distribution (which is symmetrical about 5 successes). p' is the probability of obtaining a given number of successes, and P the summation of p' from either tail, *i.e.* the probability of obtaining a number of successes deviating from the expected number 5 by a given or any greater amount in one direction. Thus the probability of obtaining exactly 8 successes is 0·0439, and the probability of obtaining 8 or more successes is 0·0547.

The probabilities given by the ordinary χ^2 test are given under the heading $P(\chi)$. These are one-half the ordinary χ^2 probabilities, so as to be comparable with a single tail of the distribution. In the computation of χ^2 the number of failures as well as the number of successes must be taken into account, giving a value for 8 successes, for example, of $3^2/5 + 3^2/5$ or 3·6. Since χ is distributed normally for one degree of freedom the exact probability correspond-

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

ing to this value of χ can be obtained by taking the square root and using a table of normal probability integral.

It will be seen that χ^2 always under-estimates the probability, the discrepancies being quite large except at the extreme tails. The true probability of 8 or more successes, for example, is 0·0547, whereas χ gives a value of 0·0290. We should thus on the χ^2 test judge such a deviation as almost reaching the 5 per cent. level of significance (2·5 per cent. for one tail), whereas in reality its probability of occurrence is over 10 per cent.

These discrepancies are primarily due to the fact that χ is a continuous distribution, whereas the distribution it is endeavouring to approximate is discontinuous. If we group the χ distribution, taking the half units of deviation from expectation as the group boundaries, we may expect to obtain a much closer approximation to the true distribution. This is equivalent to computing the values of χ^2 for deviations half a unit less than the true deviations, 8 successes, for example, being reckoned as $7\frac{1}{2}$, 2 as $2\frac{1}{2}$. This correction may be styled the *correction for continuity*, and the resultant value of χ denoted by χ'.

The resultant probabilities are shown under the heading $P(\chi')$. The agreement is now remarkably good, although the expectation of successes is only 5. χ' somewhat over-estimates the probability at the tails, and under-estimates it near the centre of the distribution. Only for the extreme values is the discrepancy really large, relative to the true probability, but inasmuch as in tests of significance we are generally concerned with the region (of one tail) of between 0·5 per cent. and 2·5 per cent., this is not of importance. In the critical region the discrepancies of this distribution are almost sufficiently small to be neglected in ordinary tests of significance, and were it only with distributions of this type that we had to deal we might leave the matter here.

There is, however, another source of discrepancy which does not appear in the above example. The χ distribution for one degree of freedom is necessarily symmetrical in terms of deviations from the expected values, whereas many of the distributions which we wish to test are not. Table II exhibits the binomial distribution

$$\left(\tfrac{3}{4} + \tfrac{1}{4}\right)^{20}.$$

Here, as before, $P(\chi)$, in general, seriously under-estimates the probabilities. $P(\chi')$, however, is not now an adequate approximation in the critical region.

The magnitude of the errors to which $P(\chi')$ is subject in the various types of distribution will be discussed in the next section. Here it is sufficient to point out that $P(\chi')$ over-estimates the probability on the shorter tail, and under-estimates it on the longer tail, these

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

TABLE II.

Successes.	p'.	P.	$P(\chi)$.	Discrepancy.	$P(\chi')$.	Discrepancy.
0	0·0032	0·0032	0·0049	+0·0017	0·0102	+0·0070
1	0·0211	0·0243	0·0192	−0·0051	0·0351	+0·0108
2	0·0669	0·0912	0·0618	−0·0294	0·0985	+0·0073
3	0·1339	0·2251	0·1515	−0·0736	0·2192	−0·0059
4	0·1897	0·4148	0·3129	−0·1019	0·3983	−0·0165
5	0·2023	—	—	—	—	—
6	0·1686	0·3829	0·3129	−0·0700	0·3983	+0·0154
7	0·1124	0·2143	0·1515	−0·0628	0·2192	+0·0049
8	0·0609	0·1019	0·0618	−0·0401	0·0985	−0·0034
9	0·0271	0·0410	0·0192	−0·0218	0·0351	−0·0059
10	0·0099	0·0139	0·0049	−0·0090	0·0102	−0·0037
11	0·0030	0·0040	0·0010	−0·0030	0·0023	−0·0017
12	0·0008	0·0010	0·0001	−0·0009	0·0004	−0·0006
13	0·0002	0·0002	—	—	—	—

discrepancies being reversed near the centre of the distribution. In symmetrical and nearly symmetrical distributions $P(\chi')$ over-estimates the probabilities at both tails and under-estimates them near the centre of the distribution. Such discrepancies, however, are small compared with those arising in violently unsymmetrical cases.

From the nature of these discrepancies it is clear that no simple modification of the correction for continuity will materially improve the approximations obtained.

Discrepancies of the χ^2 Test after correcting for Continuity.

If any given contingency distribution is calculated by means of the appropriate hypergeometric series the discrepancies between the values of χ corresponding to the true probabilities and the equivalent values of χ' can be evaluated. These discrepancies are more convenient to work with than the discrepancies between the true and χ' probabilities which were considered in the last section: clearly they are immediately convertible the one into the other by reference to a table of the normal probability integral.

There are, of course, in general no discrepancies corresponding to the exact 2·5 per cent. and 0·5 per cent. points, but it is possible to determine approximate hypothetical discrepancies corresponding to these points on the true probability scale by interpolation. If the discrepancies are plotted against the logarithms of the true probabilities the resultant points lie on remarkably regular curves, making graphical or other interpolation easy. Fig. 1 shows the graph of these discrepancies in the case of the binomial of Table II.

For every distribution generated by a 2 × 2 contingency table with fixed marginal totals but variable class numbers, therefore, a

Downloaded from https://academic.oup.com/jrssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

hypothetical discrepancy in χ' for any given level of significance can be obtained. Clearly the variety of contingency tables met with in practice is very large, but the variation of the χ' discrepancy can be exhibited in quite a compact table in sufficient detail for practical purposes.

The distribution generated by any contingency table with fixed

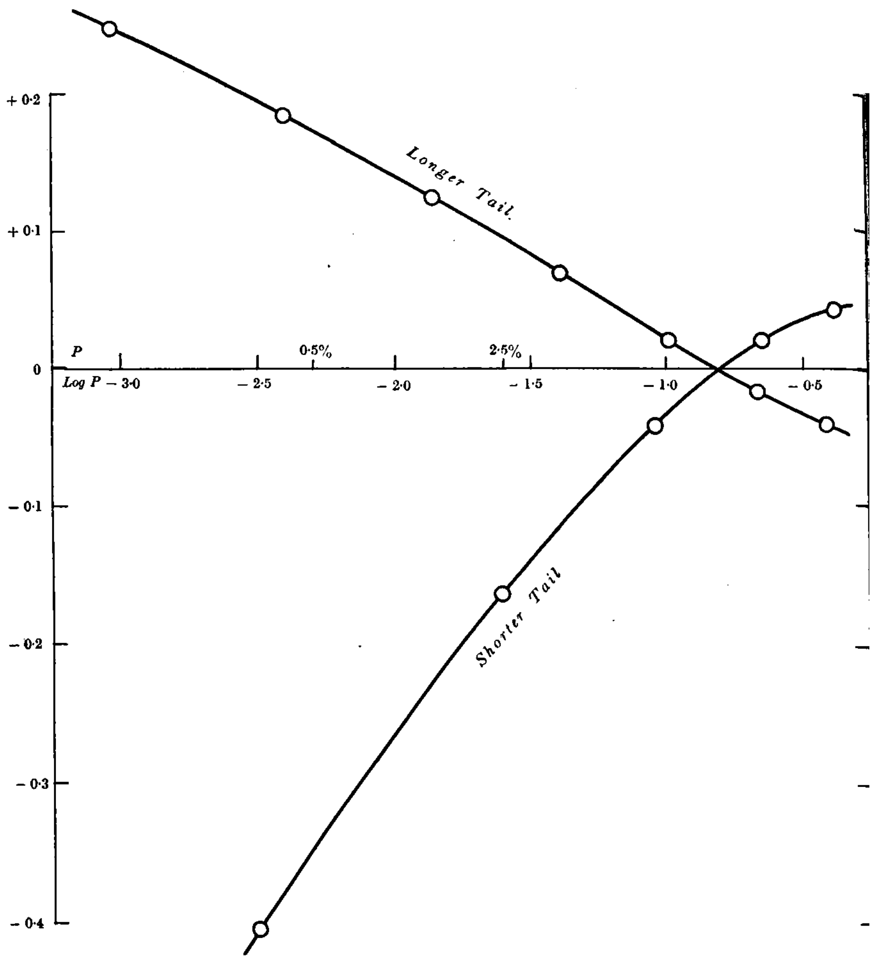


FIG. 1.—Discrepancies of χ' for binomial $(\frac{1}{4} + \frac{3}{4})^{20}$.

marginal totals may be defined in various ways. We might, for instance, use the marginal totals (only three of which are independent) for this purpose, but any three functions of them will do equally well. There is, in fact, a three-dimensional field of distributions and any

Downloaded from https://academic.oup.com/jrssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

position in this field can be defined by three co-ordinates. For our purpose it will be most convenient to use co-ordinates which represent properties of the distribution rather than of the table.

Consider the general table discussed in the second section. The distribution will have $n + 1$ terms corresponding to the $n + 1$ values from 0 to n of d. The expectation for d will be nn'/N, and this will always be less than $\frac{1}{2} n$. All distributions, therefore, may be classified according to the smallest expectation of the table, and according to the number of terms or range. In each class there will be a whole series of distributions having fixed expectation m and range $n + 1$, but with varying N (and n'). N and n' can, in fact, assume all integral solutions of

$$m = nn'/N.$$

with $n' \geq n$. The contingency distribution with smallest possible N will be called the *limiting contingency distribution* of that class. The smallest N must be equal to or greater than n^2/m, but there is no upper limit. As N tends to infinity the distribution approximates to the binomial distribution

$$(p + q)^n,$$

where $p = m/n$, *i.e.* the expectation divided by the range less one. In what follows it will be convenient to use m and p instead of m and $n + 1$ as defining the class.

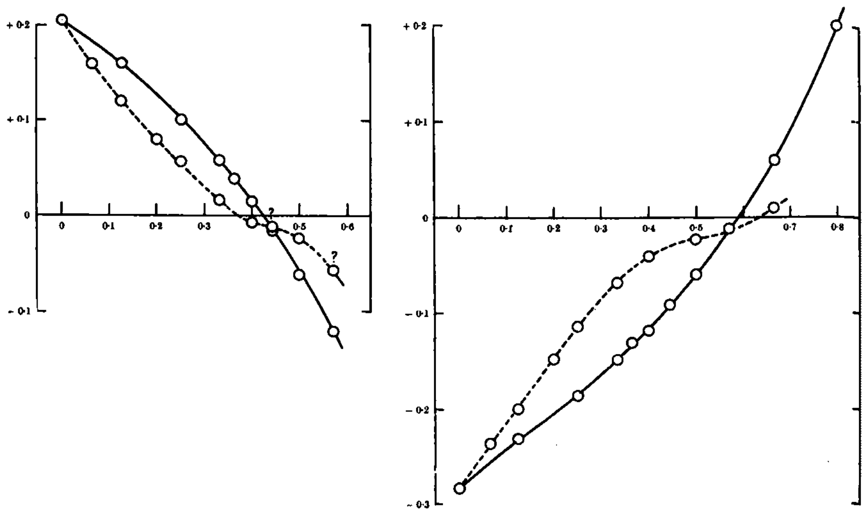
With expectation 4 and range 12 + 1, for example, we obtain the series with sets of marginal totals (24, 12; 24, 12), (27, 12; 26, 13), (30, 12; 28, 14), etc., with the limiting binomial $(\frac{1}{3} + \frac{2}{3})^{12}$. With expectation $3\frac{1}{2}$ and range 10 + 1 we obtain the series (30, 10; 26, 14), (50, 10; 39, 21), etc. In each case the first-named distribution is the limiting contingency distribution.

The utility of the above classification lies in the fact (established by examination of special cases) that the χ' discrepancies are similar for all distributions in any one class, and in general decrease or increase steadily with increasing N. Thus the knowledge of the χ' discrepancy for the limiting contingency distribution and the binomial of any class sets definite limits to its value for any distribution of that class.

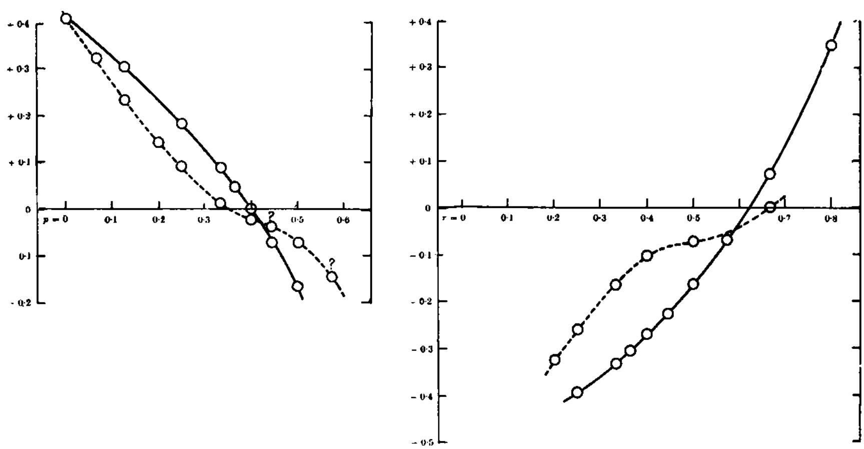
Moreover the χ' discrepancies for the limiting contingency distributions and the binomials vary in a regular manner as m and p are varied. Fig. 2 illustrates this variation with variations of p when m is equal to 4. There are four separate diagrams corresponding to 2·5 per cent. and 0·5 per cent. points of the longer and shorter tails. In each case the values actually calculated are marked. It will be seen that the values fall very satisfactorily on to smooth curves (the curves of binomial values are shown full, those of the limiting contingency series values dotted).

Downloaded from https://academic.oup.com/jrssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

Practically we are only interested in the values of p between 0 and 0·5, but higher values of p have been included where possible to illustrate the continuity. In the series of limiting contingency



2·5% point Longer Tail. 2·5% point Shorter Tail.



0·5% point Longer Tail. 0·5% point Shorter Tail.

FIG. 2.—Discrepancies of χ', expectation 4.

The vertical scales represent discrepancies, the horizontal scales values of p. The full curves are those given by the binomials, the dotted curves those given by the limiting contingency distributions.

distributions only values such that $N = n^2/m$ have been included, except in the case of the points marked with a query, where no

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

integral values of N exactly equal to n^2/m were available, but the points given by the next higher integral values appeared of sufficient interest for conclusion.

These diagrams are similar to those obtained with other expectations, subject to the proviso that the larger the expectation the smaller the discrepancies. Several interesting facts may be at once deduced from them. Thus of all the binomial distributions with given expectation and $p \leq 0{\cdot}5$ the Poisson distribution ($p = 0$) gives the greatest discrepancies, the discrepancies in χ' for expectation 4 being $+ 0{\cdot}205$ and $- 0{\cdot}283$ for the 2·5 per cent. point on the longer and shorter tail respectively, corresponding to χ' probabilities of 1·52 per cent. and 4·68 per cent. respectively. Also the discrepancies on the two tails of distributions with small p are approximately equal but of opposite sign. The discrepancies of the symmetrical binomial ($p = 0{\cdot}5$), on the other hand, are quite small, namely $- 0{\cdot}060$ with a corresponding probability of 2·87 per cent. Thus asymmetry is the most powerful disturbing factor of the χ' test. This confirms the conclusion of the fourth section.

The discrepancies of the limiting contingency series are less than those of the corresponding binomial series except for a small region in the neighbourhood of zero discrepancy. In general, therefore, the binomial may be regarded as furnishing the upper limits, and the limiting contingency distribution the lower limits to the discrepancies with any given m and p.

The application of the above results to the construction of practical tests of significance will now be considered. Clearly if it is desired to test the significance of a table with some particular marginal totals (that is, with some particular m, p and N) and the 2·5 per cent. and 0·5 per cent. values of χ' for the corresponding distribution are known, the test can be made immediately by comparing the value of χ' calculated from the table with the 2·5 per cent. and 0·5 per cent. values. Ideally, therefore, a table of the 2·5 per cent. and 0·5 per cent. values of χ' for each tail is required, covering all possible contingency distributions. This would be a three-dimensional table, and we have not attempted to construct it. Instead, two two-dimensional tables have been computed, one giving the values of χ' for the binomial distributions, the other those for the limiting contingency distributions. Both the 2·5 per cent. and the 0·5 per cent. points are tabulated for both tails for values of m ranging from 1 to 96 and values of p equal to 0, 0·25, and 0·5. Since the longer and shorter tail are identical when $p = 0{\cdot}5$ and the binomial and limiting contingency distribution are identical when $p = 0$, the whole set of values for either the 2·5 per cent. or 0·5 per cent. points can be most conveniently given in a single table. This is done

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

in Table III. In the case of the binomial with $m = 5$, $p = 0{\cdot}25$, for instance, Fig. 1 shows that the discrepancies for the 2·5 per cent. and 0·5 per cent. points are $-0{\cdot}162$ and $-0{\cdot}343$ on the shorter

TABLE III.

The 2·5 *per cent. and* 0·5 *per cent. points of* χ'.

The values for the binomial distributions are shown in ordinary type, those for the limiting contingency distributions in italics.

m = the smallest expectation. $p = \dfrac{\text{the smallest expectation}}{\text{the smallest marginal total}}$.

	2·5 per cent. points.					0·5 per cent. points.				
	Longer tail.		Shorter tail.			Longer tail.		Shorter tail.		
m \ p	0.	0·25.	0·5.	0·25.	0.	0.	0·25.	0·5.	0·25.	0.
1	2·32	2·10 *2·04*	— —	— —	—	3·33	2·79 *2·67*	— —	— —	—
2	2·24	2·08 *2·03*	— *1·91*	— —	—	3·13	2·79 *2·67*	— —	— —	—
3	2·19	2·07 *2·02*	1·88 *1·93*	1·73 *1·83*	—	3·05	2·78 *2·67*	— *2·48*	— —	—
4	2·16	2·06 *2·02*	1·90 *1·94*	1·77 *1·85*	1·68	2·97	2·76 *2·67*	2·41 *2·50*	2·18 *2·32*	—
5	2·14	2·06 *2·01*	1·91 *1·94*	1·80 *1·86*	1·71	2·95	2·75 *2·66*	2·44 *2·52*	2·23 *2·36*	2·06
6	2·13	2·05 *2·01*	1·92 *1·94*	1·82 *1·87*	1·74	2·92	2·73 *2·66*	2·47 *2·53*	2·27 *2·38*	2·13
8	2·11	2·04 *2·00*	1·93 *1·95*	1·84 *1·89*	1·77	2·88	2·72 *2·65*	2·50 *2·54*	2·32 *2·42*	2·19
12	2·08	2·02 *1·99*	1·94 *1·95*	1·87 *1·90*	1·81	2·83	2·70 *2·64*	2·52 *2·55*	2·38 *2·46*	2·27
24	2·05	2·01 *1·99*	1·95 *1·95*	1·90 *1·92*	1·86	2·76	2·67 *2·63*	2·55 *2·56*	2·45 *2·50*	2·37
48	2·03	2·00 *1·98*	1·96 *1·96*	1·91 *1·94*	1·89	2·70	2·64 *2·62*	2·56 *2·57*	2·49 *2·52*	2·43
96	2·01	1·99 *1·97*	1·96 *1·96*	1·93 *1·94*	1·91	2·67	2·63 *2·60*	2·57 *2·57*	2·52 *2·54*	2·48

tail, and $+0{\cdot}095$ and $+0{\cdot}173$ on the longer tail. Adding the 2·5 per cent. discrepancies to 1·960 (the 2·5 per cent. point of χ) and the 0·5 per cent. discrepancies to 2·576 (the 0·5 per cent. point of χ) gives the tabulated values of 1·80, 2·23, 2·06 and 2·75 respectively.

Since the table only contains values for the binomials and the limiting contingency distributions it only serves to provide upper and lower limits to the actual 2·5 per cent. and 0·5 per cent points of χ' for other contingency distributions. In testing any particular

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

contingency table interpolation will be necessary to determine these upper and lower limits, unless the m and p of the table to be tested happen to coincide with the tabulated values. There is, however, no need for any great accuracy in this interpolation, which can be made more or less by inspection.

Should the actual value of χ' calculated from the particular contingency table being tested happen to fall between the upper and lower limits of either the 2·5 per cent. or 0·5 per cent. points, then the appropriate hypergeometric series must be calculated, at east if an exact test is required.

Table III also serves to determine the limits of applicability of the χ' test when the standard χ^2 distribution for one degree of freedom is used (*i.e.* of the ordinary χ^2 test after correction for continuity). Thus if it is considered that errors in the χ' test up to $\pm$ 0·5 per cent. may be tolerated when the true probability is 2·5 per cent., the limits of the 2·5 per cent. point of χ' given by Table III must be 2·054 and 1·881, these being the 2·0 per cent. and 3·0 per cent. points of the ordinary χ distribution. When p equals 0 (the worst case), these limits are exceeded when m is less than 40. In the symmetrical distributions ($p = 0·5$) m can be as small as 3. At these limits the relative discrepancies at the 0·5 per cent. points are decidedly greater, for the χ' probabilities range between 0·8 per cent. and 0·3 per cent. roughly.

Table IV gives the limits for permissible discrepancies of $\pm$ 0·5 per cent. and $\pm$ 0·25 per cent. at the 2·5 per cent. points.

TABLE IV.

Limits of m *within which the standard χ^2 table may be used.*

Permissible discrepancy at 2·5% point.	$p = 0·5$.	$p \geq 0·375$.	$p \geq 0·25$.	$p \geq 0$.
$\pm$ 0·5%	$m \geq 3$	$m \geq 7$	$m \geq 16$	$m \geq 40$
$\pm$ 0·25%	$m \geq 6$	$m \geq 20$	$m \geq 75$	$m \geq 125$

In view of the wide variation in the limits of m for different p it will probably be more convenient in practice to use Table III if it is available whenever the smallest expectation is less than 100. On the other hand, when Table III is not available the worker will not be led badly astray if he applies the ordinary χ^2 test (after correcting for continuity) to tables giving expectations as low as 10, so long as the corresponding distributions are reasonably symmetrical. Moreover, in view of the fact that the discrepancies are of opposite signs on the two tails of the distribution, no consistent under- or over-estimates of significance will be made when applying

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

the ordinary test corrected for continuity to a heterogeneous collection of data, even with very small expectations.

The correction for continuity involves practically no extra work and should therefore be applied unless the expectations are very large. Even with a smallest expectation of 500 the discrepancy introduced into the χ^2 test at the 2·5 per cent. point of either tail by omitting this correction may be as much as 0·26 per cent.

The sequence of operations in making a practical test of any 2×2 table is therefore as follows :

1. If the smallest expectation is less than 500, calculate χ'^2, *i.e.* χ^2 corrected for continuity, instead of χ^2.

2. If χ'^2 is in the neighbourhood of the 5 per cent. or 1 per cent. point and the smallest expectation is less than 100, calculate p, take the square root of χ'^2, and refer to Table III, interpolating if necessary. (Note that interpolation will often be unnecessary because the value of χ' calculated from the observations is greater or less than all the values of Table III between which interpolation would be made.) Alternatively Table IV may first be consulted to decide if reference to Table III is really necessary.

3. If the value of χ' calculated from the observations lies between the upper and lower limits given by Table III and a more precise test is desired, calculate the hypergeometric series and perform the exact test.

Example. The following figures for malocclusion of the teeth in infants were obtained by M. Hellman,[3] who draws from them the conclusion that bottle-feeding is one of the factors causing malocclusion.

	Normal Teeth.	Malocclusion.	Total.
Breast-fed	4	16	20
Bottle-fed	1	21	22
Breast and bottle fed	3	47	50
Total	8	84	92

Considering first the wholly breast-fed and wholly bottle-fed only we have the fourfold table :

	Normal Teeth.	Malocclusion.	Total.
Breast-fed	4	16	20
Bottle-fed	1	21	22
Total	5	37	42

Downloaded from https://academic.oup.com/jrssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

Here
$$\chi^2 = \frac{(4 \times 21 - 1 \times 16)^2 . 42}{5 . 37 . 20 . 22} = 2{\cdot}386,$$

$$\chi = 1{\cdot}545,\ P(\chi) = 0{\cdot}0612,\ P(\chi^2) = 2P(\chi) = 0{\cdot}1224$$

$$\chi'^2 = \frac{(3\tfrac{1}{2} \times 20\tfrac{1}{2} - 1\tfrac{1}{2} \times 16\tfrac{1}{2})^2 . 42}{5 . 37 . 20 . 22} = 1{\cdot}140,$$

$$\chi' = 1{\cdot}068,\ P(\chi') = 0{\cdot}1427,\ P(\chi'^2) = 2P(\chi') = 0{\cdot}2854.$$

Note that $P(\chi)$ and $P(\chi')$ are obtained from χ and χ' by reference to a table of the normal probability integral.

The exact distribution is computed as follows. The probability of obtaining no normal breast-fed children is

$$\frac{5\,!\,37\,!\,20\,!\,22\,!}{42\,!\,0\,!\,20\,!\,5\,!\,17\,!} = 0{\cdot}0309568.$$

Using the successive multipliers shown we obtain the complete table :

No. of Normal Breast-fed Children.	Multiplier.	Probability.	
0	—	0·0309568	
1	5.20/1.18	0·171982	
2	4.19/2.19	0·343964	
3	3.18/3.20	0·309568	
4	2.17/4.21	0·125301	0·143527
5	1.16/5.22	0·018226	
		0·999998	

Thus the true probability of obtaining four or more normal breast-fed children is 0·1435. $P(\chi')$ gives 0·1427, an excellent approximation, whereas $P(\chi)$ gives 0·0612, which, though not in itself attaining significance, is less than half the true value ; this would be exceedingly misleading if a number of such probabilities from different classes of experiment were to be combined.

χ' is so far from the 2·5 per cent. point for any m and p that exact reference to Table III is unnecessary. The smallest value of the 2·5 per cent. of χ' in Table III is 1·68.

If malocclusion is due to mechanical action on the teeth it may be considered that children both breast-fed and bottle-fed are exposed to much the same risk of damage as those wholly bottle-fed. Combining these two classes we have the following table :

	Normal Teeth.	Malocclusion.	Total.
Breast-fed	4	16	20
Bottle or breast and bottle fed ...	4	68	72
Total	8	84	92

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

This table gives the following results :

$$\chi^2 = 4\cdot113,\ \chi = 2\cdot028,\ P(\chi) = 0\cdot0213.$$
$$\chi'^2 = 2\cdot495,\ \chi' = 1\cdot580,\ P(\chi') = 0\cdot0571.$$

No. of Normal Breast-fed Children.	Probability.	
0	0·12859	
1	0·31652	
2	0·31892	
3	0·17136	
4	0·05355	
5	0·00993	
6	0·00106	0·06460
7	0·00006	
8	0·00000	
	0·99999	

Here the ordinary χ^2 test attains the 5 per cent. level of significance (2·5 per cent. on one tail). The true probability, however, 0·0646 on the one tail, is nowhere near this, and again the correction for continuity, which gives a probability of 0·0571, is a good approximation. Here also χ' is sufficiently small to make reference to Table III unnecessary.

Thus it will be seen that even on the most favourable grouping of the data association of the degree observed might have arisen by chance about once in eight times, so that Hillman's conclusions cannot be regarded as established.

In neither of these examples has it been necessary to make any exact reference to Table III. This is the case with the great majority of tests, but to illustrate the application of the table we will suppose that the results obtained in the second case were as follows :

	Normal Teeth.	Malocclusion.	Total.
Breast-fed	5	15	20
Breast or breast and bottle fed ...	3	69	72
Total	8	84	92

Here $\chi' = 2\cdot477$ and $P(\chi') = 0\cdot0066$, which on the face of it indicates significance approaching the 0·5 per cent. point, but since the smallest expectation $m = 8 \times 20/92 = 1\cdot74$, which is very small, and $p = 1\cdot74/8 = 0\cdot22$, so that the distribution will be decidedly skew, some assurance is needed that the approximation is good enough. Referring to Table III we see that for $m = 1$ and $p = 0$ the 2·5 per cent. point of χ' on the longer tail is 2·32, for $p = 0\cdot25$ it lies between

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

2·10 and 2·04, and for $m = 2$ the corresponding values are 2·24, 2·08, and 2·03. The value 2·477 calculated from the observations is greater than all these, and therefore the 2·5 per cent. level of significance is attained. The 0·5 per cent. level of significance, on the other hand, is not attained, since the value 2·477 is less than all the values of the 0·5 per cent. point between which interpolation would have to be made. The true value of the probability given by the exact distribution already worked out is 0·0110.

Contingency Tables involving more than one Degree of Freedom.

The exact method of determining the probabilities of every set of values conforming to given marginal totals can easily be extended to tables with more than two classes in either or both classifications. The probability of obtaining any given set of values is, as in the case of 2×2 tables, the product of the factorials of all the marginal totals divided by the product of the factorials of the grand total and the individual cell numbers.

In the case of a 3×2 table, for instance, which involves two degrees of freedom, the probabilities of each particular set of values can be set out in a two-dimensional diagram which consists of two sets of interlocking hypergeometric series. The computation naturally becomes tedious if the number of possible sets of values is at all large, but it involves no new principle.

The discontinuous nature of the distribution has in general not so serious an effect on the χ^2 test. For even with quite small marginal totals there are far more possible sets of values, and it is only in exceptional cases that more than one set of values has the same χ^2. If we regard the probability of obtaining any one set of values as the result of grouping a continuous frequency distribution, the χ^2 distribution may be considered as an approximation to this continuous distribution. With any considerable deviation from expectation the surfaces of equal χ^2 cut across many of the cells representing the sets of values; some of these cells will be included in the true probability for that χ^2, and others excluded. Inasmuch, however, as the probability density falls off with increasing χ^2, χ^2 may as before be expected to under-estimate the true probability except at the tails, where other sources of disturbance become important.

As an illustration of what is likely to occur in practice take the contingency table—

a_1	a_2	a_3	17
b_1	b_2	b_3	13
13	11	6	30

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

TABLE V.

Discrepancies in a 2 × 3 Contingency Table.

a_2.	a_3.	χ^2.	P.	$P(\chi^2)$.
9	2	4·750	0·1067	0·0930
7	1	4·963	0·0948	0·0836
8	1	5·192	0·0830	0·0746
5	6	5·736	0·0724	0·0568
6	1	6·102	0·0658	0·0473
3	4	6·285	0·0592	0·0432
6	6	6·363	0·0532	0·0415
3 8	5 5 }	6·442	0·0483	0·0399
4	6	6·476	0·0365	0·0392
9	4	6·630	0·0317	0·0363
9	1	6·786	0·0268	0·0336
4	2	7·311	0·0221	0·0258
3	3	8·113	0·0188	0·0173
10	2	8·338	0·0167	0·0155
7	6	8·358	0·0149	0·0153
3	6	8·583	0·0130	0·0137
5	1	8·608	0·0120	0·0135
10	3	8·914	0·0094	0·0116
10	1	9·748	0·0081	0·0076
8	0	9·838	0·0071	0·0073
7	0	10·236	0·0061	0·0060
2 9	5 5 }	10·546	0·0053	0·0051

Table V gives the true probabilities in the critical region of obtaining as great or greater χ^2 than that given in the first column. $P(\chi^2)$ is given in the third column. It will be noted that χ^2 here serves as a criterion by which deviations from expectation may be arranged in order of magnitude, and performs an additional and entirely different function to that in the case of a single degree of freedom where its *only* function has been to serve as an approximation to the actual distribution. It could, however, have been made to perform this additional function in the case of a single degree of freedom, when it would serve to combine the deviations of the two tails on the basis of equal χ^2. This would eliminate the effects of skewness in the true distribution and the agreement would be much improved. The same elimination is undoubtedly effected in the case of more than one degree of freedom. Whether χ^2 is really the best criterion of a deviation from expectation we do not propose to discuss here.

The distribution given here is likely to be an example which is fairly favourable to χ^2. Cases where some of the marginal totals are large, and others small, so that certain degrees of freedom approximate to the binomial or Poisson distributions, may be expected, in the light of the results obtained for one degree of freedom,

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026

to give much more unfavourable results. Another source of discrepancy between the χ^2 and the exact probability is that due to several cells having the same value of χ^2. This will occur in symmetrical tables, particularly those with integral expectations in the various cells.

Summary.

1. A method of obtaining the exact probability distribution associated with a 2×2 contingency table with given marginal totals is developed.

2. It is shown that the ordinary χ^2 test is liable to considerable errors when the expectations are moderately small. A simple modification is suggested which considerably increases the accuracy of χ^2. Tables are given which enable the limits of applicability of the modified test to be determined, and serve as a means of increasing the accuracy of the modified test.

3. The applicability of the χ^2 test to contingency tables involving more than one degree of freedom is briefly discussed.

References.

[1] R. A. Fisher (1922). "On the Interpretation of χ^2 from Contingency Tables, and the Calculation of *P*." *Journ. Roy. Stat. Soc.*, Vol. LXXXV, pp. 87–94.

[2] R. A. Fisher (1926). "Bayes' Theorem and the Fourfold Table." *Eugenics Rev.*, 18, pp. 32–33.

[3] M. Hellman (1914). "A Study of some Etiological Factors of Malocclusion." *Dental Cosmos*, Vol. LVI, pp. 1017–1032.

[4] K. Pearson (1900). "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling." *Phil. Mag.*, 1, Series 5, Vol. L, pp. 157–175.

[5] G. U. Yule (1922). "On the Application of the χ^2 Method of Association and Contingency Tables, with Experimental Illustrations." *Journ. Roy. Stat. Soc.*, Vol. LXXXV, pp. 95–104.

Downloaded from https://academic.oup.com/jrsssb/article/1/2/217/6905373 by ROTHAMSTEAD RESEARCH WEST COMMON user on 27 February 2026