TESTS FOR LINEAR TRENDS IN PROPORTIONS AND FREQUENCIES

P. Armitage

Statistical Research Unit of the Medical Research Council,
London School of Hygiene and Tropical Medicine

1. *Introduction*

One frequently encounters data consisting of a series of proportions, occurring in groups which fall into some natural order. The question usually asked is then not so much whether the proportions differ significantly, but whether they show a significant trend, upwards or downwards, with the ordering of the groups. In the data shown in Table 1, for instance, the usual test for a 2 × 3 contingency table yields a χ^2 equal to 7.89 on 2 degrees of freedom, corresponding to a probability of about 0.02. But this calculation takes no account of the fact that the carrier rate increases with the tonsil size, and it is reasonable to believe that a test specifically designed to detect a trend in the carrier rate as the tonsil size increases would show a much higher degree of significance.

TABLE 1

Relationship between nasal carrier rate for Streptococcus pyogenes and size of tonsils, among 1398 children aged 0–15 years. (Data from Drs. M. C. Holmes and R. E. O. Williams, summarised by Holmes and Williams, 1954)

	Present, but not enlarged	Enlarged tonsils		Total
	+	++	+++	
Carriers	19	29	24	72
Non-carriers	497	560	269	1326
	516	589	293	1398
Carrier-rate	0.0368	0.0492	0.0819	

No originality is claimed for the tests discussed in this paper. They will be familiar to many statisticians, and may be derived as particular cases of procedures already published for contingency tables with any number of rows and columns. Since the situation in which one of the classifications in a contingency table is a dichotomy (so that the data form a series of proportions) occurs so frequently, it is hoped that an explicit discussion of this case may be of interest.

We shall regard the data as forming a $2 \times k$ contingency table, and use the following notation:

	Column 1	2	3	$\cdots$	k	Total
Row 1	n_1	n_2	n_3	$\cdots$	n_k	t
Row 2	$N_1 - n_1$	$N_2 - n_2$	$N_3 - n_3$	$\cdots$	$N_k - n_k$	$T - t$
	N_1	N_2	N_3	$\cdots$	N_k	T

The proportion of individuals in the i-th column, which fall into the first row, is denoted by $p_i = n_i/N_i$, and the overall proportion is $P = t/T$. In summations (which are always over the k columns), we shall omit the suffix i. Thus, $\sum Nx$ will denote $\sum_{i=1}^{k} N_i x_i$.

2. *A test based on scores*

To measure and test the significance of the trend in the p_i, a natural procedure is to allot a score x_i to the i-th column $(x_1 < x_2 < \cdots < x_k)$, and to perform some sort of regression analysis of p on x. In addition to the column scores x_i, let us allot to each of the T observations a row score, y, taking the values $y = 1$ for each observation in Row 1, and $y = 0$ for Row 2. Then the mean value of y for the i-th column is clearly $n_i/N_i = p_i$, and the overall mean of y is $t/T = P$. Thus, a regression analysis of y on x will be equivalent to one of p on x (p_i being weighted in proportion to N_i). The T values of y could now be subjected to a formal analysis of variance, between and within columns, as follows:

	Degrees of freedom	Sum of squares
Between columns		
Due to linear regression	1	S_1
Departures from linearity	$k - 2$	S_2
	$k - 1$	$S_1 + S_2$
Within columns	$T - k$	S_3
Total	$T - 1$	$S_1 + S_2 + S_3$

where
$$S_1 = \{\sum Np(x - \bar{x})\}^2 / \sum N(x - \bar{x})^2, \tag{1}$$
$$S_1 + S_2 = \sum N(p - P)^2,$$
$$S_3 = \sum Np(1 - p),$$
$$S_1 + S_2 + S_3 = TP(1 - P), \tag{2}$$
and
$$\bar{x} = \sum Nx/T.$$

Consider first the problem of testing for general heterogeneity between columns. As in the usual model for the analysis of variance, we assume that in repeated sampling the column totals N_i are fixed. The null hypothesis is that the expected value of y (and hence of p_i) is the same for all columns. The usual analysis of variance test is to calculate the variance ratio $\{(S_1 + S_2)/(k - 1)\}/\{S_3/(T - k)\}$. However, with a variate such as y, taking only the values 0 or 1, the normal theory is strictly valid only for large samples, and in these circumstances a number of alternative approximate tests are available. In particular the usual formula for χ^2 on $k - 1$ degrees of freedom can be expressed as
$$(S_1 + S_2)/\{(S_1 + S_2 + S_3)/T\}. \tag{3}$$
Here the denominator is taken from the "Total" row in the analysis of variance table, but with the divisor T instead of the total degrees of freedom $T - 1$. In all these alternative tests, the tabulated χ^2 distribution is strictly valid only asymptotically for large sample sizes, and the tests become equivalent as the N_i increase, provided that the null hypothesis is true.

Similarly, to test the significance of the regression, the usual analysis of variance procedure would be to compare S_1 with S_3 (or perhaps with $S_2 + S_3$ if we ignored the possibility of departures from linearity). An alternative test, equivalent in large samples if the null hypothesis is true, is to calculate
$$\chi_0^2 = S_1/\{(S_1 + S_2 + S_3)/T\}, \tag{4}$$
which is distributed approximately as χ^2 on 1 degree of freedom. If we wish to calculate confidence limits for the regression coefficient, assuming that the true value might differ from zero, we should use $S_3/(T - k)$ as an estimate of variance rather than $(S_1 + S_2 + S_3)/T$.

Which of the various alternative criteria follows most closely its assumed sampling distribution, for small samples, is a matter for further study; (see the Appendix, §6). In the meantime, there seems little objection to the use of (4). This criterion is equivalent to that proposed by Yates (1948) for contingency tables with any number of rows and columns. For $k = 2$, it is equivalent to the usual χ^2 criterion for 2×2 tables (without continuity correction). For $k > 2$, a comparison of (3) and (4) shows that χ_0^2 is a part of the total χ^2, the difference between the two values representing departures from linearity, and having $k - 2$ d.f.

Denoting by b the estimated regression coefficient of y on x, and by $V(b)$ the estimated* sampling variance of b on the null hypothesis, we find that

$$b = \frac{\sum Np(x - \bar{x})}{\sum N(x - \bar{x})^2} = \frac{T \sum nx - t \sum Nx}{T \sum Nx^2 - (\sum Nx)^2}, \qquad (5)$$

$$V(b) = \frac{P(1 - P)}{\sum N(x - \bar{x})^2} = \frac{t(T - t)}{T\{T \sum Nx^2 - (\sum Nx)^2\}}, \qquad (6)$$

and, from (1), (2) and (4),

$$\chi_0^2 = \frac{b^2}{V(b)} = \frac{T\{T \sum nx - t \sum Nx\}^2}{t(T - t)\{T \sum Nx^2 - (\sum Nx)^2\}} \qquad (7)$$

on 1 degree of freedom.

The calculations cannot be performed until the scores x_i have been chosen. In the absence of any *a priori* knowledge of the type of trend to be expected, it seems reasonable to choose the x_i to be equally-spaced, and it will often be convenient to have them centred around zero. This is the procedure advocated by Yates. Thus, for k columns, we should choose $x_1 = -\frac{1}{2}(k - 1)$, $x_2 = -\frac{1}{2}(k - 3)$, $\cdots$, $x_k = \frac{1}{2}(k - 1)$. The choice of scores is discussed further in a later section. It should be emphasized that, whatever scoring system is chosen, the validity of the significance test is not affected; that is, if the null hypothesis is true, a value of χ_0^2 significant at the $\alpha\%$ level will occur only about α times out of 100.

As an example, using the data of Table 1, we shall allot equally-spaced scores as follows: $x_1 = -1$, $x_2 = 0$, $x_3 = 1$. We obtain

$$\sum nx = 5, \qquad \sum Nx = -223, \qquad \sum Nx^2 = 809,$$

whence, from (5), (6) and (7),

$$b = 0.02131,$$

$$V(b) = 0.000063160; \qquad \sqrt{V(b)} = 0.00795,$$

and $$\chi_0^2 = 7.19 \quad \text{on} \quad 1 \text{ d.f.} \qquad (P = 0.007).$$

*In repeated sampling with both sets of marginal totals fixed, the expression (6) is $(T - 1)/T$ times the *exact* variance of b. This can be shown from results given by Haldane (1940).

The test for trend indicates, as expected, a considerably higher degree of significance than the total χ^2 of 7.89 on 2 d.f. The test for departures from linear regression gives $\chi^2 = 7.89 - 7.19 = 0.70$ on 1 d.f., which is non-significant. In this particular example, the association between carrier rate and tonsil size may be due to the association of both factors with the age or social class of the child.

Yates (1948) points out that the same formula for χ_0^2 is obtained whether one considers the regression of row score on column score, or that of column score on row score. Now, when there are only two rows, a test for the regression of column score on row score is equivalent to a test for the difference between the mean column score for the first row and that for the second row. For some types of data, particularly where the row totals are fixed beforehand, it will be more natural to think of the χ_0^2 test in this way, rather than in terms of the regression of p on x. In the data shown in Table 2, for instance, the row totals, 32 and 32, were fixed by the experimental design, and it seems more natural to ask whether the mean scores in the two treatment groups differ significantly, rather than whether the proportion of patients in group A, in each column, shows a linear trend with the score. In this example, the total $\chi^2 = 5.91$ on 3 d.f. ($P = 0.12$), whereas $\chi_0^2 = 5.26$ on 1 d.f. ($P = 0.02$), showing a fairly definite improvement in group A as compared with group B.

TABLE 2

Changes in size of ulcer crater, 3 months after start of treatment, for patients in two treatment groups (From Table IV of Doll and Pygott, 1952)

Treatment group	Number of cases with crater				Total
	Larger	Less than 2/3 healed	2/3 or more healed	Healed	
A	6	4	10	12	32
B	11	8	8	5	32
	17	12	18	17	64
Score, x_i	−1.5	−0.5	+0.5	+1.5	

A test criterion exactly equivalent to χ_0^2 has been used in genetical applications by Fisher and Ford (1947, p. 163) and by Holt (1948, p. 148). A recent example of the use of this test, in a 2 × 3 table, is given by Grüneberg (1955). He compares the proportions of animals in two stocks which show some effect on 0, 1 or 2 sides of the body. The formula for χ^2 given by C. A. B. Smith in the Appendix to Grüneberg's paper is equivalent to our (7). The more general problem in which more than two stocks are compared could be treated by Yates's methods.

3. *Trends in frequencies*

If $P = t/T$ is very small, we may substitute $T/(T - t) \sim 1$ in (7). Defining $e_i = tN_i/T$, the "expected" frequency corresponding to the observed frequency n_i, we find from (7) that

$$\chi_0^2 = \frac{\{\sum x(n - e)\}^2}{\sum ex^2 - (\sum ex)^2/t}. \tag{8}$$

The numerator of (8) is the square of the cross-product, U, of the scores x_i with the discrepancies $n_i - e_i$. The denominator is equal to $\sum e\,(x - \bar{x})^2$, where $\bar{x} = \sum ex/t$, i.e. a weighted sum of squares of the x_i about their mean, the weights being the expected numbers. The test is thus based entirely on the frequencies in the first row, and is clearly valid only when the sampling errors of the frequencies in the second row are relatively negligible. The frequencies in the first row may be thought of as those occurring in a sample of size t from a multinomial distribution. The denominator of (8) is then obtained directly as the variance of U in repeated sampling, with t and the expected frequencies e_i kept constant. The expression (8) may thus be written as $\chi_0^2 = U^2/V(U)$.

In Table 3, the expected frequencies e_i, have been obtained by sub-dividing the total number of maternal deaths, 127, in proportion to the number of mothers at risk during each of the eight periods. The last line of Table 3 suggests, perhaps, a slight tendency for the maternal mortality rates to fall. The scores, x_i, have been taken as the mid-points of the different periods, minus 1900. The total χ^2, calculated from the observed and expected frequencies is 3.91 on 7 d.f. ($P = 0.79$); even if the whole of this quantity were ascribed to regression it would barely reach the 5% level of significance on 1.d.f. In fact, application of (8) gives $\chi_0^2 = 1.27$ on 1 d.f. ($P = 0.26$). The data, therefore, do not provide any evidence for a gradual decline in maternal mortality amongst women of this particular parity and age-group.

4. *Kendall's rank correlation test*

An alternative approach to data like those in Tables 1 and 2 is to apply rank correlation methods (Kendall, 1948; Stuart, 1953). In Table 1, for instance, we could regard the 1398 children as being ranked

TABLE 3

Maternal mortality in New South Wales, for primiparae aged 40 and over. (From Tables I, II and III of Wilcocks and Lancaster, 1951)

	1894–1900	1901–1907	1908–1910	1911–1920
x_i	−2.5	4.5	9.5	16.0
Number of mothers, N_i	346	454	272	1133
Deaths				
Observed, n_i	10	9	3	23
Expected, e_i	6.603	8.664	5.191	21.621
Maternal mortality rate, per 1,000	28.90	19.82	11.03	20.30

	1921–1930	1931–1937	1938–1942	1943–1948	Total
x_i	26.0	34.5	40.5	46.0	
Number of mothers, N_i	1546	909	699	1296	6655
Deaths					
Observed, n_i	32	17	13	20	127
Expected, e_i	29.503	17.347	13.339	24.732	127.000
Maternal mortality rate, per 1,000	20.70	18.70	18.60	15.43	

in two ways. In the first ranking (corresponding to the rows of Table 1), 1326 individuals are tied with a rank of (1 + 1326)/2 = 663.5, and the remaining 72 are tied with a rank of 1326 + (1 + 72)/2 = 1362.5. In the second ranking (for columns), 516 are tied with a rank of (1 + 516)/2 = 258.5, 589 are tied with a rank of 516 + (1 + 589)/2 = 811.0, and 293 are tied with a rank of 516 + 589 + (1 + 293)/2 = 1252.0. To test for a tendency for the carrier-rate to increase or decrease with tonsil size, we could apply the usual techniques of rank correlation, making allowance for the considerable number of ties.

To calculate Kendall's statistic, S (§1.9 of his book), we form the sum of products of each frequency in the second row with the frequencies above and to the right of it, and subtract the sum of products of each frequency in the first row with those below and to the right of it. Thus, in the notation previously used:

$$S = (N_1 - n_1)(n_2 + n_3 + \cdots + n_k) + (N_2 - n_2)(n_3 + \cdots + n_k) + \cdots + (N_{k-1} - n_{k-1})n_k - n_1\{(N_2 - n_2) + \cdots + (N_k - n_k)\} - \cdots - n_{k-1}(N_k - n_k). \tag{9}$$

When the null hypothesis is true (i.e. there is no association), the variance of S is (writing Kendall's (4.5) in the present notation),

$$V(S) = \frac{t(T - t)}{3T(T - 1)} (T^3 - \sum N^3). \tag{10}$$

(Stuart (1953) considers inequalities for the variance when the null hypothesis is not true.) A test for association is thus provided by

$$\chi_1^2 = S^2/V(S) \quad \text{on} \quad 1 \text{ d.f.} \tag{11}$$

If $k = 2$, χ_1^2 is equal to $(T - 1)/T$ times the usual χ^2 for a 2×2 table (without continuity correction). This factor is of no great importance, in view of the asymptotic nature of the assumed χ^2 distribution.

At first sight the approach of §2 seems to bear little relationship to that of the present section. In point of fact the two methods are quite closely related. It is known (Hemelrijk, 1952) that when one of the classifications in a rank correlation table is a dichotomy, Kendall's test based on S is equivalent to Wilcoxon's test for the sum of the ranks in one of the sub-groups (see Kruskal and Wallis, 1952, for references). This, in turn, is equivalent to a test for the difference between the mean ranks in the two sub-groups, since the overall sum of ranks is constant. This difference would be the same as the difference in mean column scores, discussed in §1, if we chose the score for each column to be equal to the mid-rank for that column. Thus, we should have $x_1 = (1 + N_1)/2$, $x_2 = (1 + 2N_1 + N_2)/2$, $x_3 = (1 + 2N_1 + 2N_2 + N_3)/2$, etc. It would, therefore, not be surprising if the χ_1^2 test were closely related to the χ_0^2 test with the x_i chosen in this way, or at least chosen so as to be linearly related to these values. It is not difficult to show directly that this is so.

Rearranging the terms in (9), and writing $p_i = n_i/N_i$, we find that

$$S = \sum nx, \tag{12}$$

where
$$\begin{aligned} x_i &= N_1 + N_2 + \cdots + N_{i-1} - N_{i+1} - \cdots - N_k \\ &= (1 + 2N_1 + \cdots + 2N_{i-1} + N_i) - (T + 1). \end{aligned} \tag{13}$$

These scores x_i are linearly related to the mid-ranks given above, and it can easily be verified that

$$\sum Nx = 0 \qquad \text{and} \qquad \sum Nx^2 = (T^3 - \sum N^3)/3. \tag{14}$$

Hence, from (7) and (14),

$$\begin{aligned} \chi_0^2 &= \frac{3T^2S^2}{t(T - t)(T^3 - \sum N^3)} \\ &= \{T/(T - 1)\}\chi_1^2, \qquad \text{from (10) and (11)} \end{aligned} \tag{15}$$

When the N_i are equal, the x_i are clearly equally-spaced. The rank correlation test is then equivalent to the regression test with equally-spaced scores, except for the factor $T/(T - 1)$ in (15). As already stated, this factor is asymptotically unimportant. The tests would have been exactly equivalent if, in the formula (4) from which (7) is derived, the total degrees of freedom, $T - 1$, had been used as a divisor in the denominator, instead of T. As we have seen, when $k = 2$, χ_0^2 agrees with the usual χ^2 for a 2×2 table, whereas χ_1^2 differs from it by a factor $(T - 1)/T$.

As examples of the rank correlation test, formulae (9)–(11) have been applied to the data shown in Tables 1 and 2. For Table 1,

$$S = 16229, \qquad V(S) = 38{,}543{,}560.2,$$

and

$$\chi_1^2 = 6.83 \qquad (P = 0.009),$$

as compared with $\chi_0^2 = 7.19$. For Table 2,

$$S = 330 \qquad V(S) = 20720.25$$

$$\chi_1^2 = 5.26 \qquad (P = 0.02),$$

as compared with $\chi_0^2 = 5.26$ (the exact agreement being coincidental).

5. *Choice of test*

Since the rank correlation test has been shown to be equivalent (apart from the factor $T/(T - 1)$) to the regression test, with a particular choice of scores depending on the N_i, the decision whether to use χ_1^2 or χ_0^2 reduces to a choice of the most suitable system of scoring. In most situations there will be no prior reason to expect any particular type of relationship, and it is difficult to formulate any general advice.

If the columns are defined by a measurement, like age, it will often be reasonable to choose scores linearly related to the values assumed by the measurement, taking mid-points of groups where necessary (as in Table 3).

If the columns are defined by a qualitative classification as in Table 1, the choice is more arbitrary. If the problem is primarily thought of as a trend in proportions in well-defined ordered groups, the regression method with equally-spaced x_i seems the most appropriate. An estimate is obtained of the mean change in p_i from group to group, and one avoids the use of scores depending on the N_i which may be difficult to interpret. If, on the other hand, the grouping by columns is arbitrary, there may be little virtue in using equally-spaced x_i, and the rank correlation method is perhaps the more objective. Fortunately, the two tests will usually give fairly close results.

It may be of interest to conclude with a historical note. The reader will find a number of sets of data suitable for analysis by the methods outlined here, in two papers by Karl Pearson (1909, 1910). In the first paper, Pearson considered situations in which the columns corresponded to a numerical variate; the rows were assumed to represent a dichotomy of an underlying normal variate and the method provided an estimate of the hypothetical correlation coefficient (sometimes referred to as "biserial r"). In the second paper, the method was extended for data in which the columns were qualitatively defined, but might still be ordered; an estimate of the hypothetical correlation ratio ("biserial η"), not dependent on the ordering, was obtained, and the trend was assessed by inspection. These methods have largely fallen into disuse, partly because of difficulties in determining the sampling errors of the coefficients, and partly because the existence of a normal variate underlying the dichotomy by rows was not generally accepted.

6. *Appendix. Exact distribution of χ_0^2 on the null hypothesis*

The exact distribution of χ_0^2 has been determined, by enumeration of all possible results, for the case where $k = 3$, $N_1 = N_2 = N_3 = 10$, and the n_i each follow the binomial distribution $(\frac{1}{2} + \frac{1}{2})^{10}$. The probabilities with which χ_0^2 exceeds various tabulated percentiles of the χ^2 distribution on 1 d.f. are shown in Table 4. This table also shows the cumulative distribution of an alternative test criterion,

$$\chi_2^2 = S_1/\{(S_2 + S_3)/(T - k)\},$$

in the notation of §2. The formula for χ_2^2 differs from that for χ_0^2, (4), in having as denominator the mean square about regression. The two test criteria are connected by the relationship

$$\chi_2^2 = (T - 2)\chi_0^2/(T - \chi_0^2),$$

$$= 28\chi_0^2/(30 - \chi_0^2)$$

since $T = 30$.

Although the expected frequencies in this example are as low as 5, Table 4 shows (a) that there is little to choose between the two tests up to about the 5% level of significance, and (b) that the distribution of either test criterion agrees well with the theoretical χ^2 distribution between the 50% and 5% points. The appreciable discrepancy at the lower end of each distribution is due to there being a probability of 0.176 that $\chi_0^2 = \chi_2^2 = 0$. It would be dangerous to generalize from this example alone, but the results are at least encouraging.

TABLE 4

Cumulative distributions of two alternative test criteria, in case described in text

Values of χ^2	Cumulative probability		
	Tabulated	χ_0^2	χ_2^2
0–	1.000	1.000	1.000
0.0^3157–	0.990	0.824	0.824
0.0^3628–	0.980	0.824	0.824
0.0^2293–	0.950	0.824	0.824
0.0158–	0.900	0.824	0.824
0.0642–	0.800	0.824	0.824
0.148–	0.700	0.824	0.824
0.455–	0.500	0.504	0.504
1.074–	0.300	0.264	0.264
1.642–	0.200	0.263	0.263
2.706–	0.100	0.116	0.116
3.841–	0.050	0.042	0.044
5.412–	0.020	0.014	0.042
6.635–	0.010	0.012	0.013
10.827–	0.0010	0.0005	0.0028

I am indebted to Professor A. Bradford Hill and Dr. J. O. Irwin for commenting on the first draft of this paper; to Dr. M. C. Holmes and Dr. R. E. O. Williams for permission to quote, in Table 1, details not appearing in their paper; and to Miss Irene Allen for computing assistance.

Since this paper was accepted for publication, the regression test based on χ_0^2 has been discussed by W. G. Cochran (1954), *Biometrics 10*: 417–451 §§6.2, 6.3.

REFERENCES

Doll, R. and Pygott, F. (1952). Factors influencing the rate of healing of gastric ulcers. *Lancet*, i, 171–182.

Fisher, R. A. and Ford, E. B. (1947). The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. *Heredity*, *1*, 143–174.

Grüneberg, H. (1955). Genetical studies on the skeleton of the mouse. XV. The interaction between major and minor variants. (With an Appendix by C. A. B. Smith). *J. Genet.* (in press)

Haldane, J. B. S. (1940). The mean and variance of χ^2, when used as a test of homogeneity, when expectations are small. *Biometrika*, *31*, 346–355.

Hemelrijk, J. (1952). Note on Wilcoxon's two-sample test when ties are present. *Ann. Math. Statist.*, *23*, 133–135.

Holmes, M. C. and Williams, R. E. O. (1954). The distribution of carriers of *Streptococcus pyogenes* among 2,413 healthy children. *J. Hyg., Camb.*, *52*, 165–179.

Holt, S. B. (1948). The effect of maternal age on the manifestation of a polydactyl gene in mice. *Ann. Eugen., Lond.*, *14*, 144–157.

Kendall, M. G. (1948). *Rank correlation methods*. London: Griffin.

Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion analysis of variance. *J. Amer. Statist. Ass.*, *47*, 583–621.

Pearson, K. (1909). On a new method of determining correlation between a measured character *A*, and a character *B*, of which only the percentage of cases wherein *B* exceeds (or falls short of) a given intensity is recorded for each grade of *A*. *Biometrika*, *7*, 96–105.

Pearson, K. (1910). On a new method of determining correlation, when one variable is given by an alternative and the other by multiple categories. *Biometrika*, *7*, 248–257.

Stuart, A. (1953). The estimation and comparison of strengths of association in contingency tables. *Biometrika*, *40*, 105–110.

Wilcocks, W. J. and Lancaster, H. O. (1951). Maternal mortality in New South Wales with special reference to age and parity. *J. Obstet. Gynaec., Brit. Emp.*, *58*, 945–960.

Yates, F. (1948). The analysis of contingency tables with groupings based on quantitative characters. *Biometrika*, *35*, 176–181.

LINKED CITATIONS

Tests for Linear Trends in Proportions and Frequencies
P. Armitage
Biometrics, Vol. 11, No. 3. (Sep., 1955), pp. 375-386.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28195509%2911%3A3%3C375%3ATFLTIP%3E2.0.CO%3B2-K

[Footnotes]

[*] **The Mean and Variance of |chi2, When Used as a Test of Homogeneity, When Expectations are Small**
J. B. S. Haldane
Biometrika, Vol. 31, No. 3/4. (Mar., 1940), pp. 346-355.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194003%2931%3A3%2F4%3C346%3ATMAVOW%3E2.0.CO%3B2-W

References

The Mean and Variance of |chi2, When Used as a Test of Homogeneity, When Expectations are Small
J. B. S. Haldane
Biometrika, Vol. 31, No. 3/4. (Mar., 1940), pp. 346-355.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194003%2931%3A3%2F4%3C346%3ATMAVOW%3E2.0.CO%3B2-W

On a New Method of Determining Correlation Between a Measured Character A, and a Character B, of which Only the Percentage of Cases Wherein B Exceeds (or Falls Short of) a Given Intensity is Recorded for Each Grade of A
Karl Pearson
Biometrika, Vol. 7, No. 1/2. (Jul. - Oct., 1909), pp. 96-105.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28190907%2F10%297%3A1%2F2%3C96%3AOANMOD%3E2.0.CO%3B2-R

LINKED CITATIONS

On a New Method of Determining Correlation, when One Variable is Given by Alternative and the Other by Multiple Categories
Karl Pearson
Biometrika, Vol. 7, No. 3. (Apr., 1910), pp. 248-257.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28191004%297%3A3%3C248%3AOANMOD%3E2.0.CO%3B2-X

The Estimation and Comparison of Strengths of Association in Contingency Tables
A. Stuart
Biometrika, Vol. 40, No. 1/2. (Jun., 1953), pp. 105-110.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195306%2940%3A1%2F2%3C105%3ATEACOS%3E2.0.CO%3B2-C

The Analysis of Contingency Tables with Groupings Based on Quantitative Characters
F. Yates
Biometrika, Vol. 35, No. 1/2. (May, 1948), pp. 176-181.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194805%2935%3A1%2F2%3C176%3ATAOCTW%3E2.0.CO%3B2-W

NOTE: *The reference numbering from the original has been maintained in this citation list.*