

Chapter 11

Contingency Tables



Abstract This chapter introduces permutation methods for the analysis of contingency tables. Included in this chapter are six example analyses illustrating computation of permutation methods for goodness-of-fit tests, analysis of contingency tables composed of two nominal-level (categorical) variables, analysis of contingency tables composed of two ordinal-level (ranked) variables, analysis of contingency tables composed of one nominal-level variable and one ordinal-level variable, analysis of contingency tables composed of one nominal-level variable and one interval-level variable, and analysis of contingency tables composed of one ordinal-level variable and one interval-level variable. Included in this chapter are permutation versions of Pearson chi-squared goodness-of-fit test, Pearson's chi-squared test of independence, Cramér's symmetrical measure of nominal association, Goodman and Kruskal's τ_a and τ_b asymmetric measures of association for two categorical variables, Goodman and Kruskal's G measure of association for two ranked variables, Somers' d_{yx} and d_{xy} asymmetric measures of association for two ranked variables, Freeman's θ measure of association for a categorical independent variable and a ranked dependent variable, Pearson's point-biserial correlation coefficient for one dichotomous variable and one interval-level variable, and Jaspén's correlation coefficient for one ranked variable and one interval-level variable.

This chapter introduces exact and Monte Carlo permutation statistical methods for selected measures of relationship among nominal-, ordinal-, and interval-level variables, commonly called contingency table analysis. The analysis of contingency tables with their associated measures of effect size and tests of significance constitutes a substantial portion of nonparametric statistical methods.

In this last chapter, exact and Monte Carlo permutation statistical methods for the analysis of contingency tables are illustrated with six types of analyses. The first section of the chapter considers permutation statistical methods applied to conventional goodness-of-fit tests; for example, Pearson's chi-squared goodness-of-fit test. The second section is devoted to permutation statistical methods for analyzing contingency tables composed of two cross-classified nominal-level (categorical)

variables; for example, Pearson's symmetric chi-squared test of independence for two categorical variables and Goodman and Kruskal's t_a and t_b asymmetric measures of association for two categorical variables. The third section utilizes permutation statistical methods for analyzing contingency tables composed of two cross-classified ordinal-level (ranked) variables; for example, Goodman and Kruskal's symmetric G measure of association for two ranked variables and Somers' d_{yx} and d_{xy} asymmetric measures of association for two ranked variables. The fourth section is the first of three sections utilizing permutation statistical methods for analyzing contingency tables composed of two cross-classified mixed-level variables. In this fourth section permutation statistical methods are utilized for analyzing contingency tables composed of one nominal-level (categorical) variable cross-classified with one ordinal-level (ranked) variable; for example, Freeman's θ measure for one categorical independent variable and one ranked dependent variable. The fifth section utilizes permutation statistical methods for analyzing contingency tables composed of one nominal-level variable cross-classified with one interval-level variable; for example, Pearson's point-biserial correlation coefficient for one dichotomous variable and one interval-level variable. The sixth section utilizes permutation statistical methods for analyzing contingency tables composed of one ordinal-level variable cross-classified with one interval-level variable; for example, Jaspert's correlation coefficient for one ranked variable and one interval-level variable.¹

There exist a vast array of measures of association and correlation. The few measures described here illustrate the application of permutation statistical methods to the analysis of two-way contingency tables at various levels of measurement and were selected for their popularity in the research literature and inclusion in various introductory textbooks. For a more comprehensive treatment of permutation statistical methods applied to measures of association and correlation see a 2018 book on *The Measurement of Association* by the authors [2].

11.1 Goodness-of-Fit Tests

Goodness-of-fit tests are essential for determining how well observed data conform to hypothetical models. When at all reasonable, exact goodness-of-fit tests are preferred over asymptotic tests. More specifically, goodness-of-fit tests are designed to compare the observed values in k discrete, unordered categories with values that are expected to occur under chance conditions. For example, for a fair coin the expectation for 100 independent trials is 50 heads and 50 tails over many, many trials of 100 tosses. The observed values, say 60 heads and 40 tails, are then compared with expected values under the null hypothesis, $H_0: p(H) = p(T) = 0.50$.

¹There is never any reason to relate a higher-level independent variable with a lower-level dependent variable due to the loss of information from the independent variable.

The most popular goodness-of-fit test for k discrete, unordered categories is Pearson's chi-squared test, although Wald's likelihood-ratio test is occasionally encountered in the contemporary literature. Utilizing the conventional notation presented in many introductory textbooks, Pearson's chi-squared goodness-of-fit test for k discrete, unordered categories is given by

$$\chi^2 = \sum_{i=1}^k \frac{O_i^2}{E_i} - N ,$$

where O_i and E_i denote the observed and expected frequency values, respectively, for $i = 1, \dots, k$. Under the Neyman–Pearson null hypothesis, $H_0: O_i = E_i$ for $i = 1, \dots, k$, χ^2 is asymptotically distributed as Pearson's χ^2 with $k - 1$ degrees of freedom, under the assumption of normality.²

Consider the random assignment of N objects to k discrete, unordered categories where the probability that any one of the N objects occurs in the i th category is $p_i > 0$ for $i = 1, \dots, k$. Then the probability that O_i objects occur in the i th category for $i = 1, \dots, k$ is the multinomial probability given by

$$P(O_1, O_2, \dots, O_k | p_1, p_2, \dots, p_k, N) = \frac{N!}{\prod_{i=1}^k O_i!} \prod_{i=1}^k p_i^{O_i} ,$$

where

$$\sum_{i=1}^k O_i = N \quad \text{and} \quad \sum_{i=1}^k p_i = 1 .$$

11.1.1 Example 1

Two example analyses will serve to illustrate the permutation approach to goodness-of-fit-tests. For the first analysis under the Neyman–Pearson population model of statistical inference, consider a small example set of data with $k = 3$ unordered categories, $N = 6$ total objects, $O_1 = 5$ objects in the first category, $O_2 = 1$ object in the second category, and $O_3 = 0$ objects in the third category. The observed and expected frequencies along with the associated theoretical proportions are listed in Table 11.1. For the example data listed in Table 11.1 with $N = 6$ observations,

²Pearson's χ^2 test statistic is one of several test statistics that utilizes a lower-case Greek letter for both the sample test statistic and the population parameter.

Table 11.1 Example data for Pearson's chi-squared goodness-of-fit test statistic with $k = 3$ discrete, unordered categories and $N = 6$ observations

Category number	Observed frequency	Expected frequency	Theoretical proportion
1	5	2	0.3333
2	1	2	0.3333
3	0	2	0.3333
Sum	6	6	1.0000

Pearson's chi-squared goodness-of-fit test statistic is

$$\chi^2 = \sum_{i=1}^k \frac{O_i^2}{E_i} - N = \frac{5^2}{2} + \frac{1^2}{2} + \frac{0^2}{2} - 6 = 7.00 .$$

Under the Neyman–Pearson null hypothesis, $H_0: O_i = E_i$ for $i = 1, \dots, k$, χ^2 is asymptotically distributed as Pearson's χ^2 with $k - 1$ degrees of freedom. With $k - 1 = 3 - 1 = 2$ degrees of freedom, the asymptotic probability value of $\chi^2 = 7.00$ is $P = 0.0302$, under the assumption of normality.

11.1.2 An Exact Permutation Analysis

For the example data listed in Table 11.1 under the Fisher–Pitman permutation model of statistical inference there are exactly

$$M = \binom{N + k - 1}{k - 1} = \binom{6 + 3 - 1}{3 - 1} = \binom{8}{2} = 28$$

possible, equally-likely arrangements in the reference set of all permutations of the example data listed in Table 11.1. Table 11.2 lists the $M = 28$ arrangements of the observed data, the associated χ^2 values, and the multinomial point probability values to six decimal places, ordered by the χ^2 values from lowest ($\chi_1^2 = 0.00$) to highest ($\chi_{28}^2 = 12.00$). The exact probability value of $\chi^2 = 7.00$ is the sum of the multinomial point probability values associated with values of χ^2 that are equal to or greater than the observed χ^2 value. There are only nine arrangements of the observed data with χ^2 test statistic values that are equal to or greater than the observed value of $\chi^2 = 7.00$: six values of $\chi^2 = 7.00$ and three values of $\chi^2 = 12.00$, all in rows indicated with asterisks in Table 11.2. Thus if all M arrangements of the $N = 6$ observations listed in Table 11.1 occur with equal chance under the Fisher–Pitman null hypothesis, the exact probability value of $\chi^2 = 7.00$ computed on the $M = 28$ possible arrangements of the observed data with $k = 4$ categories preserved for each arrangement is

$$P = 6(0.008230) + 3(0.001372) = 0.053496 .$$

Table 11.2 Example discrete probability distribution for the data given in Table 11.1 with χ^2 test statistic values and associated multinomial probability values

Order	Frequencies	χ^2 value	Probability
1	2 2 2	0.00	0.123457
2	3 2 1	1.00	0.082305
3	3 1 2	1.00	0.082305
4	2 3 1	1.00	0.082305
5	2 1 3	1.00	0.082305
6	1 3 2	1.00	0.082305
7	1 2 3	3.00	0.082305
8	3 3 0	3.00	0.027435
9	3 0 3	3.00	0.027435
10	0 3 3	3.00	0.027435
11	4 1 1	3.00	0.041152
12	1 4 1	3.00	0.041152
13	1 1 4	3.00	0.041152
14	4 2 0	4.00	0.020576
15	4 0 2	4.00	0.020576
16	2 4 0	4.00	0.020576
17	2 0 4	4.00	0.020576
18	0 4 2	4.00	0.020576
19	0 2 4	4.00	0.020576
20*	5 1 0	7.00	0.008230
21*	5 0 1	7.00	0.008230
22*	1 5 0	7.00	0.008230
23*	1 0 5	7.00	0.008230
24*	0 5 1	7.00	0.008230
25*	0 1 5	7.00	0.008230
26*	6 0 0	12.00	0.001372
27*	0 6 0	12.00	0.001372
28*	0 0 6	12.00	0.001372
Sum			1.000000

There is a substantial difference between the exact probability value of $P = 0.0535$ and the asymptotic probability value of $P = 0.0302$; that is,

$$\Delta_P = 0.0535 - 0.0302 = 0.0233 .$$

With the sparse data given in Table 11.1 there are only $M = 28$ possible arrangements of cell frequencies given the marginal frequency totals and it would be unreasonable to expect a continuous mathematical function such as Pearson's χ^2 to fit such a small discrete distribution consisting of only six different values with any precision.

Fig. 11.1 Punnett square depicting RYRY, RYRy, RYrY, RYry, RyRY, RYrY, rYRY, rYRy, and ryRY hybrids with nine black circles, RyRy, Ryry, and ryRy hybrids with three dark gray circles, rYrY, rYry, and ryrY hybrids with three light gray circles, and the sole rryy hybrid with a single white circle

		Type of Egg			
		RY	Ry	rY	ry
Type of Sperm	RY	RYRY	RYRy	RYrY	RYry
	Ry	RyRY	RyRy	RyRY	Ryry
	rY	rYRY	rYRy	rYrY	rYry
	ry	ryRY	ryRy	ryrY	ryry

11.1.3 Example 2

Gregor Mendel (1822–1884) is notable for his studies of hybridization utilizing the common garden pea while he resided in the Augustinian monastery of St. Thomas at Brünn in Austrian Silesia.^{3,4} In one of his many studies of garden peas, Mendel crossed hybrid plants producing round yellow peas with hybrid plants producing wrinkled green peas. To produce his hybrids, Mendel carefully brushed the pollen of one pea plant onto the pistils of another plant. The first generation, as expected, produced all round yellow peas—both dominant characteristics. However, the second generation yielded four varieties of peas: round yellow, wrinkled yellow, round green, and wrinkled green.⁵

Figure 11.1 displays the different varieties of peas in a Punnett square where RY denotes round yellow peas, Ry denotes round green peas, rY denotes wrinkled yellow peas, and ry denotes wrinkled green peas.⁶ In the Punnett diagram in Fig. 11.1, the round-yellow hybrids, RYRY, RYRy, RYrY, RYry, RyRY, RYrY, rYRY, rYRy, and ryRY, are indicated by nine black circles (●), the round-green hybrids, RyRy, Ryry, and ryRy, are indicated by three dark gray circles (◐), the wrinkled-

³Presently the region of Silesia is located largely in Poland with smaller parts in the Czech Republic and in Germany.

⁴Mendel’s birth name was Johann, but he adopted the name Gregor when he entered the monastery in 1843 at the age of 21.

⁵Mendel was elected abbot of the monastery in 1868 at the age of 46, the administrative duties of which precluded any further research. Mendel passed away in 1884 at the age of 62.

⁶More technically, RY denotes round and yellow, Ry denotes round and not yellow, rY denotes not-round and yellow, and ry denotes not-round and not-yellow.

Table 11.3 Mendel's second-generation hybridization frequency data for $N = 556$ common garden peas

Category	Ratio	Frequency	
		Observed	Expected
RY	9	315	312.75
Ry	3	101	104.25
rY	3	108	104.25
ry	1	32	34.75
Sum		556	556.00

yellow hybrids, rYrY, rYry, and ryrY, are indicated by three light gray circles (●), and the single wrinkled-green ryrY hybrid is indicated by a white circle (○).

Mendel's 15 double-hybrid plants produced a sample of $N = 556$ peas. Mendel's data for the $N = 556$ second-generation hybrids are listed in Table 11.3, along with the expected values which approximate the ratios 9:3:3:1.

For Mendel's hybridization data listed in Table 11.3, Pearson's chi-squared goodness-of-fit test statistic is

$$\chi^2 = \sum_{i=1}^k \frac{O_i^2}{E_i} - N = \frac{315^2}{312.75} + \frac{101^2}{104.25} + \frac{108^2}{104.25} + \frac{32^2}{34.75} - 556 = 0.4700 .$$

Under the Neyman–Pearson null hypothesis, $H_0: O_i = E_i$ for $i = 1, \dots, k$, χ^2 is asymptotically distributed as Pearson's χ^2 with $k - 1$ degrees of freedom. With $k - 1 = 4 - 1 = 3$ degrees of freedom, the asymptotic probability value of $\chi^2 = 0.4700$ is $P = 0.9254$, under the assumption of normality.

In a 1936 paper published in *Annals of Science*, R.A. Fisher, Galton Professor of Eugenics at University College, London, re-examined Mendel's hybridization data, questioned Mendel's recording of his observations, and concluded that the very close agreement between Mendel's observed and expected series was unlikely to have arisen by chance [4]. Fisher submitted his paper at Christmas time in 1936 to *Annals of Science* with a comment to the editor, Dr. Douglas McKie:

I had not expected to find the strong evidence which has appeared that the data had been cooked. This makes my paper far more sensational than ever I had intended... (quoted in Box [3, p. 297]).

11.1.4 An Exact Permutation Analysis

Under the Fisher–Pitman permutation model, the exact probability value of an observed chi-squared value of $\chi^2 = 7.00$ is given by the sum of the multinomial point probability values associated with the values of χ^2 that are equal to or greater than the observed χ^2 value. For the Mendel hybridization data listed in Table 11.3

under the permutation model there are

$$M = \binom{N + k - 1}{k - 1} = \binom{556 + 4 - 1}{4 - 1} = \binom{559}{3} = 28,956,759$$

possible, equally-likely arrangements in the reference set of all permutations of Mendel's hybridization data listed in Table 11.3, making an exact permutation analysis feasible. If all M arrangements of the $N = 556$ observations listed in Table 11.3 occur with equal chance under the Fisher–Pitman null hypothesis, the exact probability value of the observed chi-squared value of $\chi^2 = 0.4700$ computed on the $M = 28,956,759$ possible arrangements of the observed data with $k = 4$ categories preserved for each arrangement is $P = 0.9381$; that is, the sum of the multinomial probability values associated with values of $\chi^2 = 0.4700$ or greater.

11.1.5 A Measure of Effect Size

A chi-squared test of goodness-of-fit and its associated probability value provide no information as to the closeness of the fit between the observed and theoretical values, only whether they are statistically significant under the Neyman–Pearson population-model null hypothesis. Measures of effect size are essential in such cases as they index the magnitude of the fit between the observed and expected frequencies and indicate the practical significance of the research. A maximum-corrected measure of effect size is easily specified for a chi-squared goodness-of-fit test [1].

Define

$$q = \min(E_1, E_2, \dots, E_k)$$

for k disjoint, unordered categories. Then with q determined, the maximum value of χ^2 is given by

$$\chi_{\max}^2 = \frac{N(N - q)}{q} \quad (11.1)$$

and a maximum-corrected measure of effect size for Pearson's chi-squared goodness-of-fit test is given by

$$ES(\chi^2) = \frac{\chi_o^2}{\chi_{\max}^2},$$

where χ_o^2 denotes the observed value of χ^2 [8].

Table 11.4 Maximum arrangement of cell frequencies for Mendel’s second-generation hybridization frequency data with $N = 556$ observations

Category	Ratio	Frequency	
		Observed	Expected
RY	9	0	312.75
Ry	3	0	104.25
rY	3	0	104.25
ry	1	556	34.75
Sum		556	556.00

For Mendel’s hybridization data listed in Table 11.3, the minimum expected frequency value is

$$q = \min(E_1, E_2, E_3, E_4) = \min(312.75, 104.25, 104.25, 34.75) = 34.75 ,$$

and the maximum possible Pearson’s χ^2 test statistic value given $k = 4$, $q = 34.75$, and $N = 556$ is

$$\chi_{\max}^2 = \frac{N(N - q)}{q} = \frac{556(566 - 34.75)}{34.75} = 8340 .$$

To illustrate the function of Eq. (11.1) imagine that all $N = 556$ observations are concentrated in the one category with the smallest expected value and the remaining $k - 1$ categories contain zero observations. In this case all $N = 556$ observations are concentrated in the last category with the minimum expected value of $E_4 = 34.75$, such as depicted in Table 11.4. Then the maximum value of Pearson’s chi-squared goodness-of-fit test statistic is

$$\begin{aligned} \chi_{\max}^2 &= \sum_{i=1}^k \frac{O_i^2}{E_i} - N \\ &= \frac{0^2}{312.75} + \frac{0^2}{104.25} + \frac{0^2}{104.25} + \frac{556^2}{34.75} - 556 = 8340 , \end{aligned}$$

and the maximum-corrected measure of effect size is

$$ES(\chi^2) = \frac{\chi^2}{\chi_{\max}^2} = \frac{0.4700}{8340} = 0.5636 \times 10^{-4} ,$$

indicating that the observed value of $\chi^2 = 0.4700$ is an insignificantly small proportion of the maximum possible χ^2 value, given the expected values $E_1 = 312.75$, $E_2 = 104.25$, $E_3 = 104.25$, and $E_4 = 34.75$.

11.2 Contingency Measures: Nominal by Nominal

The most popular test for the cross-classification of two nominal-level (categorical) variables is Pearson's chi-squared test of independence, which is presented in every introductory textbook. Utilizing the conventional notation presented in many introductory textbooks for a contingency table with r rows and c columns, Pearson's chi-squared test statistic is given by

$$\chi^2 = N \left(\sum_{i=1}^r \sum_{j=1}^c \frac{O_{ij}^2}{R_i C_j} - 1 \right), \quad (11.2)$$

where O_{ij} denotes the observed frequency in the i th row and j th column for $i = 1, \dots, r$ and $j = 1, \dots, c$, R_i denotes a row marginal frequency total for $i = 1, \dots, r$, C_j denotes a column marginal frequency total for $j = 1, \dots, c$, and N denotes the total number of values in the observed contingency table.

11.2.1 Example 1

Two examples will serve to illustrate Pearson's chi-squared test of independence for an $r \times c$ contingency table. For an analysis under the Neyman–Pearson population model, consider the sparse example data given in Table 11.5 with $r = 3$ rows, $c = 3$ columns, and $N = 9$ observations. Under the Neyman–Pearson population model, the chi-squared test statistic value for the example data given in Table 11.5 is

$$\begin{aligned} \chi^2 &= N \left(\sum_{i=1}^r \sum_{j=1}^c \frac{O_{ij}^2}{R_i C_j} - 1 \right) \\ &= 9 \left[\frac{0^2}{(2)(2)} + \frac{0^2}{(2)(3)} + \frac{2^2}{(2)(4)} + \frac{0^2}{(3)(2)} + \frac{3^2}{(3)(3)} + \frac{0^2}{(3)(4)} \right. \\ &\quad \left. + \frac{2^2}{(4)(2)} + \frac{0^2}{(4)(3)} + \frac{2^2}{(4)(4)} - 1 \right] = 11.2500. \end{aligned}$$

Table 11.5 Example data for Pearson's chi-squared test of independence with $r = 3$ rows, $c = 3$ columns, and $N = 9$ cross-classified observations

Row	Column			Total
	1	2	3	
1	0	0	2	2
2	0	3	0	3
3	2	0	2	4
Total	2	3	4	9

Under the Neyman–Pearson null hypothesis, $H_0: O_{ij} = E_{ij}$ for $i = 1, \dots, r$ and $j = 1, \dots, c$, where the expected cell values are given by

$$E_{ij} = \frac{O_{ij}}{R_i C_j}$$

for $i = 1, \dots, r$ and $j = 1, \dots, c$, χ^2 is asymptotically distributed as Pearson's χ^2 with $(r-1)(c-1)$ degrees of freedom. With $(r-1)(c-1) = (3-1)(3-1) = 4$ degrees of freedom, the asymptotic probability value of $\chi^2 = 11.2500$ is $P = 0.0239$, under the assumption of normality.

11.2.2 A Measure of Effect Size

The fact that a chi-squared statistical test produces a low probability value indicates only that there are differences among the response measurement scores between the two variables that (possibly) cannot be attributed to error. The obtained probability value does not indicate whether these differences are of any practical value. Measures of effect size express the practical or clinical significance of an obtained chi-squared value, as contrasted with the statistical significance of a chi-squared value. The most popular measure of effect size for Pearson's chi-squared test of independence is Cramér's V given by

$$V = \sqrt{\frac{\chi^2}{N[\min(r-1, c-1)]}}.$$

For the example data given in Table 11.6 with $\chi^2 = 11.2500$, Cramér's measure of effect size is

$$V = \sqrt{\frac{11.2500}{9[\min(3-1, 3-1)]}} = \sqrt{\frac{11.2500}{18.00}} = 0.7906.$$

For a critical evaluation of Cramér's V measure of effect size, see a discussion in *The Measurement of Association* by the authors [2, pp. 80–82].

Occasionally in the contemporary literature, Cohen's measure of effect size for a chi-squared test of independence is encountered. Cohen's measure is given by

$$w = \sqrt{\frac{\chi^2}{N}}.$$

Table 11.6 All $M = 39$ arrangements of the frequency data given in Table 11.5 with associated chi-squared values and hypergeometric point probability values

Table	Observed frequencies				Chi-squared	Probability
	O_{11}	O_{12}	O_{21}	O_{22}		
1*	2	0	0	3	18.0000	0.793651×10^{-3}
2*	2	0	0	0	14.0625	0.317460×10^{-2}
3*	0	2	0	2	13.0000	0.238095×10^{-2}
4*	0	0	0	3	11.2500	0.476190×10^{-2}
5	2	0	0	2	10.5625	0.952381×10^{-2}
6	0	2	0	0	9.5625	0.952381×10^{-2}
7	0	0	2	0	9.5625	0.952381×10^{-2}
8	0	2	2	0	9.5625	0.952381×10^{-2}
9	1	0	0	3	9.5625	0.634921×10^{-2}
10	1	1	1	2	9.2500	0.476190×10^{-2}
11	2	0	0	1	9.2500	0.14286
12	1	0	0	0	9.0000	0.634921×10^{-2}
13	0	1	0	0	7.8750	0.952381×10^{-2}
14	0	0	1	0	7.8750	0.952381×10^{-2}
15	0	2	0	1	7.7500	0.014286
16	0	0	2	1	7.7500	0.014286
17	0	0	0	1	7.0000	0.014286
18	0	2	1	1	6.4375	0.019048
19	0	1	2	1	6.4375	0.019048
20	1	1	0	0	6.1875	0.019048
21	1	0	1	0	6.1875	0.019048
22	0	0	0	2	6.0625	0.028571
23	1	1	0	2	5.6875	0.019048
24	1	0	1	2	5.6875	0.019048
25	0	2	1	0	5.6250	0.028571
26	0	1	2	0	5.6250	0.028571
27	0	1	1	2	5.3125	0.019048
28	0	1	0	2	5.1250	0.028571
28	0	0	1	2	5.1250	0.028571
30	1	1	1	0	4.5000	0.028571
31	1	1	1	1	3.8125	0.038952
32	0	1	0	1	3.4375	0.057143
33	0	0	1	1	3.4375	0.057143
34	1	0	0	2	3.2500	0.057143
35	1	0	0	1	3.0625	0.057143
36	1	1	0	1	2.8750	0.057143
37	1	0	1	1	2.8750	0.057143
38	0	1	1	0	2.8125	0.057143
39	0	1	1	1	1.0000	0.114285
Sum						1.000000

For the example data given in Table 11.5, Cohen's measure of effect size is

$$w = \sqrt{\frac{11.2500}{9}} = 1.1180 .$$

11.2.3 An Exact Permutation Analysis

Given the observed marginal frequency totals for the example data, there are only $M = 39$ possible, equally-likely arrangements of cell frequencies in the reference set of all permutations of the $N = 9$ observations given in Table 11.5, making an exact permutation analysis possible. Table 11.6 lists the $M = 39$ arrangements of cell frequencies, the associated chi-squared test statistic values, and the hypergeometric point probability values given by

$$p(O_{11}, \dots, O_{rc} | R_1, \dots, R_r, C_1, \dots, C_c, N) = \frac{\left(\prod_{i=1}^r R_i! \right) \left(\prod_{j=1}^c C_j! \right)}{N! \prod_{i=1}^r \prod_{j=1}^c O_{ij}!} .$$

Because the observed marginal frequency totals are fixed, Table 11.6 lists only cell frequencies O_{11} , O_{12} , O_{21} , and O_{22} , as the remaining five cell frequencies can be determined from the observed marginal frequency totals.

Under the Fisher–Pitman permutation model, the exact probability value of $\chi^2 = 11.2500$ is the sum of the hypergeometric point probability values associated with the chi-squared values that are equal to or greater than the observed chi-squared value. For the results listed in Table 11.6, there are four chi-squared test statistic values that are equal to or greater than the observed value of $\chi^2 = 11.2500$: $\chi_1^2 = 18.0000$, $\chi_2^2 = 14.0625$, $\chi_3^2 = 13.0000$, and $\chi_4^2 = 11.2500$, in rows 1, 2, 3, and 4, respectively, and indicated by asterisks. Thus the exact probability value of $\chi^2 = 11.2500$ is

$$0.7937 \times 10^{-3} + 0.3175 \times 10^{-2} + 0.2381 \times 10^{-2} + 0.4762 \times 10^{-2} = 0.0111 .$$

There is a substantial difference between the exact probability value of $P = 0.0111$ and the asymptotic probability value of $P = 0.0239$; that is,

$$\Delta_P = 0.0239 - 0.0111 = 0.0128 .$$

With such sparse data as given in Table 11.5 there are only $M = 39$ possible arrangements of cell frequencies given the marginal frequency totals with only 25 different chi-squared values and it would be unreasonable to expect a continuous

Table 11.7 Example data for Pearson’s chi-squared test of independence with $r = 3$ rows, $c = 5$ columns, and $N = 63$ cross-classified observations

Row	Column					Total
	1	2	3	4	5	
1	6	2	5	7	1	21
2	0	8	5	8	4	25
3	1	1	6	6	3	17
Total	7	11	16	21	8	63

mathematical function such as Pearson’s χ^2 to fit such a small discrete distribution with any precision.

11.2.4 Example 2

For a second example of a chi-squared analysis of a nominal-nominal contingency table, consider the 3×5 contingency table with cell frequencies given in Table 11.7. Following Eq. (11.2), Pearson’s chi-squared test statistic for the frequency data given in Table 11.7 is

$$\begin{aligned} \chi^2 &= N \left(\sum_{i=1}^r \sum_{j=1}^c \frac{O_{ij}^2}{R_i C_j} - 1 \right) \\ &= 63 \left[\frac{6^2}{(21)(7)} + \frac{2^2}{(21)(11)} + \dots + \frac{6^2}{(17)(21)} + \frac{3^2}{(17)(8)} - 1 \right] = 16.6279 . \end{aligned}$$

Under the Neyman–Pearson null hypothesis the chi-squared test statistic is asymptotically distributed as Pearson’s χ^2 with $(r - 1)(c - 1)$ degrees of freedom. With $(r - 1)(c - 1) = (3 - 1)(5 - 1) = 8$ degrees of freedom, the asymptotic probability value of $\chi^2 = 16.6279$ is $P = 0.0342$, under the assumption of normality.

11.2.5 A Measure of Effect Size

For the frequency data given in Table 11.7, Cramér’s measure of effect size is

$$\begin{aligned} V &= \sqrt{\frac{\chi^2}{N[\min(r - 1, c - 1)]}} \\ &= \sqrt{\frac{16.6279}{63[\min(3 - 1, 5 - 1)]}} = \sqrt{\frac{16.6279}{126.00}} = 0.3633 \end{aligned}$$

and Cohen’s measure of effect size is

$$w = \sqrt{\frac{\chi^2}{N}} = \sqrt{\frac{16.6279}{63}} = 0.2639 .$$

11.2.6 An Exact Permutation Analysis

Given the observed marginal frequency totals for the example data, there are $M = 11,356,797$ possible, equally-likely arrangements of the cell frequencies in the reference set of all permutations of the $N = 63$ observations given in Table 11.7, making an exact permutation analysis possible. Under the Fisher–Pitman permutation model, the exact probability of $\chi^2 = 16.6279$ is the sum of the hypergeometric point probability values associated with the chi-squared values calculated on all M possible arrangements of the cell frequencies, given the observed marginal frequency totals. For the frequency data given in Table 11.7, there are $M = 11,356,797$ possible, equally-likely arrangements of the cell frequencies given the observed marginal frequency totals, of which 10,559,996 chi-squared test statistic values are equal to or greater than the observed chi-squared value of $\chi^2 = 16.6279$, yielding an exact hypergeometric probability value of $P = 0.0306$. Note that with $M = 11,356,797$ possible arrangements of the data given in Table 11.7, the asymptotic χ^2 probability value of $P = 0.0342$ closely approximates the exact hypergeometric probability value of $P = 0.0306$.

11.2.7 Goodman–Kruskal’s t_a and t_b Measures

While all measures of association based on Pearson’s chi-squared are symmetric measures, Goodman and Kruskal’s two asymmetric proportional-reduction-in-error measures (t_a and t_b) allow researchers to specify an independent and a dependent variable. Consider two cross-classified, unordered polytomies, A and B , with variable A the dependent variable and variable B the independent variable. Table 11.8

Table 11.8 Notation for the cross-classification of two categorical variables, A_j for $j = 1, \dots, c$ and B_i for $i = 1, \dots, r$

B	A				Total
	a_1	a_2	\dots	a_c	
b_1	n_{11}	n_{12}	\dots	n_{1c}	$n_{1.}$
b_2	n_{21}	n_{22}	\dots	n_{2c}	$n_{2.}$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
b_r	n_{r1}	n_{r2}	\dots	n_{rc}	$n_{r.}$
Total	$n_{.1}$	$n_{.2}$	\dots	$n_{.c}$	N

provides notation for the cross-classification, where a_j for $j = 1, \dots, c$ denotes the c categories for dependent variable A , b_i for $i = 1, \dots, r$ denotes the r categories for independent variable B , N denotes the total of cell frequencies in the table, $n_{i.}$ denotes a marginal frequency total for the i th row, $i = 1, \dots, r$, summed over all columns, $n_{.j}$ denotes a marginal frequency total for the j th column, $j = 1, \dots, c$, summed over all rows, and n_{ij} denotes a cell frequency for $i = 1, \dots, r$ and $j = 1, \dots, c$.

Goodman and Kruskal's t_a test statistic is a measure of the relative reduction in prediction error where two types of errors are defined. The first type is the error in prediction based solely on knowledge of the distribution of the dependent variable, termed "errors of the first kind" (E_1) and consisting of the expected number of errors when predicting the c dependent variable categories (a_1, \dots, a_c) from the observed distribution of the marginals of the dependent variable ($n_{.1}, \dots, n_{.c}$). The second type is the error in prediction based on knowledge of the distributions of both the independent and dependent variables, termed "errors of the second kind" (E_2) and consisting of the expected number or errors when predicting the c dependent variable categories (a_1, \dots, a_c) from knowledge of the r independent variable categories (b_1, \dots, b_r).

To illustrate the two error types, consider predicting category a_1 only from knowledge of its marginal distribution, $n_{.1}, \dots, n_{.c}$. Clearly, $n_{.1}$ out of the N total cases are in category a_1 , but exactly which $n_{.1}$ of the N cases is unknown. The probability of incorrectly identifying one of the N cases in category a_1 by chance alone is given by

$$\frac{N - n_{.1}}{N}.$$

Since there are $n_{.1}$ such classifications required, the number of expected incorrect classifications is

$$\frac{n_{.1}(N - n_{.1})}{N}$$

and, for all c categories of variable A , the number of expected errors of the first kind is given by

$$E_1 = \sum_{j=1}^c \frac{n_{.j}(N - n_{.j})}{N}.$$

Likewise, to predict n_{11}, \dots, n_{1c} from the independent category b_1 , the probability of incorrectly classifying one of the $n_{1.}$ cases in cell n_{11} by chance alone is

$$\frac{n_{1.} - n_{11}}{n_{1.}}.$$

Since there are n_{11} such classifications required, the number of incorrect classifications is

$$\frac{n_{11}(n_{1.} - n_{11})}{n_{1.}}$$

and, for all cr cells, the number of expected errors of the second kind is given by

$$E_2 = \sum_{j=1}^c \sum_{i=1}^r \frac{n_{ij}(n_{i.} - n_{ij})}{n_{i.}} .$$

Goodman and Kruskal's t_a statistic can then be defined as

$$t_a = \frac{E_1 - E_2}{E_1} .$$

An efficient computation form for Goodman and Kruskal's t_a test statistic is given by

$$t_a = \frac{N \sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{i.}} - \sum_{j=1}^c n_{.j}^2}{N^2 - \sum_{j=1}^c n_{.j}^2} . \quad (11.3)$$

A computed value of t_a indicates the proportional reduction in prediction error given knowledge of the distribution of independent variable B over and above knowledge of only the distribution of dependent variable A . As defined, t_a is a point estimator of Goodman and Kruskal's population parameter τ_a for the population from which the sample of N cases was obtained. If variable B is considered the dependent variable and variable A the independent variable, then Goodman and Kruskal's test statistic t_b and associated population parameter τ_b are analogously defined.

11.2.8 An Example Analysis for t_a

To illustrate Goodman and Kruskal's t_a measure of nominal-nominal association, consider the contingency table given in Table 11.9 with $r = 3$ rows, $c = 4$ columns, and $N = 110$ cross-classified ordered observations. Following Eq. (11.3),

Table 11.9 Example data for Goodman and Kruskal’s t_a and t_b measures of nominal-nominal association with $r = 3$ rows, $c = 4$ columns, and $N = 110$ cross-classified observations

Row (B)	Column (A)				Total
	1	2	3	4	
1	24	2	5	6	37
2	0	22	5	8	35
3	1	1	17	19	38
Total	25	25	27	33	110

the observed value of Goodman and Kruskal’s t_a test statistic is

$$\begin{aligned}
 t_a &= \frac{N \sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_i} - \sum_{j=1}^c n_{.j}^2}{N^2 - \sum_{j=1}^c n_{.j}^2} \\
 &= \frac{110 \left(\frac{24^2}{37} + \frac{2^2}{37} + \dots + \frac{17^2}{38} + \frac{19^2}{38} \right) - (25^2 + 25^2 + 27^2 + 33^2)}{110^2 - (25^2 + 25^2 + 27^2 + 33^2)} \\
 &= 0.2797 .
 \end{aligned}$$

Under the Neyman–Pearson null hypothesis, $H_0: \tau_a = 0$, $t_a(N - 1)(r - 1)$ is asymptotically distributed as Pearson’s χ^2 with $(r - 1)(c - 1)$ degrees of freedom. With $(r - 1)(c - 1) = (3 - 1)(4 - 1) = 6$ degrees of freedom, the asymptotic probability value of $t_a = 0.2797$ is $P = 0.2852 \times 10^{-10}$, under the assumption of normality.

11.2.9 An Exact Permutation Analysis for t_a

Under the Fisher–Pitman permutation model, the exact probability value of an observed value of Goodman and Kruskal’s t_a is given by the sum of the hypergeometric point probability values associated with t_a test statistic values that are equal to or greater than the observed value of $t_a = 0.2797$. For the frequency data given in Table 11.9, there are $M = 26,371,127$ possible, equally-likely arrangements in the reference set of all permutations of cell frequencies given the observed row and column marginal frequency distributions, $\{37, 35, 38\}$ and $\{25, 25, 27, 33\}$, respectively, making an exact permutation analysis possible. There are exactly 1,523,131 t_a test statistic values that are equal to or greater than the observed value of $t_a = 0.2797$. The exact probability value of the observed t_a value under the Fisher–Pitman null hypothesis is $P = 0.0578$; that is, the sum of the hypergeometric point probability values associated with values of $t_a = 0.2797$ or greater.

11.2.10 An Example Analysis for t_b

Now consider variable B as the dependent variable. A convenient computing formula for t_b is

$$t_b = \frac{N \sum_{j=1}^c \sum_{i=1}^r \frac{n_{ij}^2}{n_{.j}} - \sum_{i=1}^r n_i^2}{N^2 - \sum_{i=1}^r n_i^2} .$$

Thus, for the frequency data given in Table 11.9 the observed value of t_b is

$$t_b = \frac{110 \left(\frac{24^2}{25} + \frac{2^2}{25} + \cdots + \frac{17^2}{27} + \frac{19^2}{33} \right) - (37^2 + 35^2 + 38^2)}{110^2 - (37^2 + 35^2 + 38^2)} = 0.4428 .$$

Under the Neyman–Pearson null hypothesis, $H_0: \tau_b = 0$, $t_b(N - 1)(c - 1)$ is asymptotically distributed as Pearson's χ^2 with $(r - 1)(c - 1)$ degrees of freedom. With $(r - 1)(c - 1) = (3 - 1)(4 - 1) = 6$ degrees of freedom, the asymptotic probability value of $t_b = 0.4428$ is $P = 0.9738 \times 10^{-28}$, under the assumption of normality.

11.2.11 An Exact Permutation Analysis for t_b

Under the Fisher–Pitman permutation model, the exact probability value of an observed value of Goodman and Kruskal's t_b is given by the sum of the hypergeometric point probability values associated with t_b test statistic values that are equal to or greater than the observed value of $t_b = 0.4428$. For the frequency data given in Table 11.9, there are $M = 26,371,127$ possible, equally-likely arrangements in the reference set of all permutations of cell frequencies given the observed row and column marginal frequency distributions, $\{37, 35, 38\}$ and $\{25, 25, 27, 33\}$, respectively, making an exact permutation analysis possible. There are exactly 991,488 t_b test statistic values that are equal to or greater than the observed value of $t_b = 0.4428$. The exact probability value of the observed t_b value under the Fisher–Pitman null hypothesis is $P = 0.0376$; that is, the sum of the hypergeometric point probability values associated with values of $t_b = 0.4428$ or greater.

11.2.12 The Relationships Among t_a , t_b , and χ^2

While no general equivalence exists between Goodman and Kruskal's t_a and t_b measures of nominal-nominal association and Pearson's χ^2 test of independence, certain relationships hold among t_a , t_b , and χ^2 under some limited conditions. Four of the relationships can easily be specified.

First, if $n_{i.} = N/r$ for $i = 1, \dots, r$, then $\chi^2 = N(r - 1)t_b$ and $t_b = \chi^2/N(r - 1)$. To illustrate the relationship between Goodman and Kruskal's t_b asymmetric measure of nominal-nominal association and Pearson's χ^2 test of independence when $n_{i.} = N/r$ for $i = 1, \dots, r$, consider the frequency data given in Table 11.10 with $r = 3$ rows, $c = 3$ columns, $N = 30$ cross-classified observations, and $n_{i.} = N/r = 10$ for $i = 1, \dots, r$. For the frequency data given in Table 11.10 with $N = 30$ observations,

$$\begin{aligned}
 t_b &= \frac{N \sum_{j=1}^c \sum_{i=1}^r \frac{n_{ij}^2}{n_{.j}} - \sum_{i=1}^r n_{i.}^2}{N^2 - \sum_{i=1}^r n_{i.}^2} \\
 &= \frac{30 \left(\frac{2^2}{5} + \frac{3^2}{10} + \dots + \frac{3^2}{10} + \frac{6^2}{15} \right) - (10^2 + 10^2 + 10^2)}{36^2 - (10^2 + 10^2 + 10^2)} \\
 &= \frac{10}{600} = 0.0167
 \end{aligned}$$

and

$$\begin{aligned}
 \chi^2 &= N \left(\sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{i.}n_{.j}} - 1 \right) \\
 &= 30 \left[\frac{2^2}{(10)(5)} + \frac{3^2}{(10)(10)} + \dots + \frac{3^2}{(10)(10)} + \frac{6^2}{(10)(15)} \right] \\
 &= 30(1.0333 - 1) = 1.00.
 \end{aligned}$$

Table 11.10 Example data for χ^2 and t_b with $r = 3$ rows, $c = 3$ columns, and $N = 30$ cross-classified observations

Row (B)	Column (A)			Total
	1	2	3	
1	2	3	5	10
2	2	4	4	10
3	1	3	6	10
Total	5	10	15	30

Then the observed value of Pearson’s χ^2 test statistic with respect to the observed value of Goodman and Kruskal’s t_b test statistic is

$$\chi^2 = N(r - 1)t_b = 30(3 - 1)(0.0167) = 1.00$$

and the observed value of Goodman and Kruskal’s t_b test statistic with respect to the observed value of Pearson’s χ^2 test statistic is

$$t_b = \frac{\chi^2}{N(r - 1)} = \frac{1.00}{30(3 - 1)} = 0.0167.$$

Second, if $n_{.j} = N/c$ for $j = 1, \dots, c$, then $\chi^2 = N(c - 1)t_a$ and $t_a = \chi^2/N(c - 1)$. To illustrate the relationship between Goodman and Kruskal’s t_a measure of nominal-nominal association and Pearson’s χ^2 test of independence when $n_{.j} = N/c$ for $j = 1, \dots, c$, consider the frequency data given in Table 11.11 with $r = 2$ rows, $c = 4$ columns, $N = 40$ cross-classified observations, and $n_{.j} = N/c = 10$ for $j = 1, \dots, c$. For the frequency data given in Table 11.11 with $N = 40$ observations,

$$\begin{aligned}
 t_a &= \frac{N \sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{i.}} - \sum_{j=1}^c n_{.j}^2}{N^2 - \sum_{j=1}^c n_{.j}^2} \\
 &= \frac{40 \left(\frac{7^2}{25} + \frac{6^2}{25} + \dots + \frac{2^2}{15} + \frac{6^2}{15} \right) - (10^2 + 10^2 + 10^2 + 10^2)}{40^2 - (10^2 + 10^2 + 10^2 + 10^2)} \\
 &= \frac{37.3333}{1200} = 0.0311
 \end{aligned}$$

Table 11.11 Example data for χ^2 and t_a with $r = 2$ rows, $c = 4$ columns, and $N = 40$ cross-classified observations

Row (B)	Column (A)				Total
	1	2	3	4	
1	7	6	8	4	25
2	3	4	2	6	15
Total	10	10	10	10	40

and

$$\begin{aligned}\chi^2 &= N \left(\sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{i.}n_{.j}} - 1 \right) \\ &= 40 \left[\frac{7^2}{(25)(10)} + \frac{6^2}{(25)(10)} + \cdots + \frac{2^2}{(15)(10)} + \frac{6^2}{(10)(15)} \right] \\ &= 40(1.0933 - 1) = 3.7333.\end{aligned}$$

Then the observed value of Pearson's χ^2 test statistic with respect to the observed value of Goodman and Kruskal's t_a test statistic is

$$\chi^2 = N(c-1)t_a = 40(4-1)(0.0311) = 3.7333$$

and the observed value of Goodman and Kruskal's t_a test statistic with respect to the observed value of Pearson's χ^2 test statistic is

$$t_a = \frac{\chi^2}{N(c-1)} = \frac{3.7333}{40(4-1)} = 0.0311.$$

Third, if $r = 2$, then $\chi^2 = Nt_a$ and $t_a = \chi^2/N$, which is Pearson's ϕ^2 coefficient of contingency. Also, if $c = 2$, then $\chi^2 = Nt_b$ and $t_b = \chi^2/N$. Thus, if $r = c = 2$, then $\chi^2 = Nt_a = Nt_b$. To illustrate the relationships between Goodman and Kruskal's t_a and t_b measures of nominal-nominal association and Pearson's χ^2 test of independence with $r = c = 2$, consider the frequency data given in Table 11.12 with $r = 2$ rows, $c = 2$ columns, and $N = 90$ cross-classified observations. For the frequency data given in Table 11.12 with $N = 90$ observations,

$$\begin{aligned}t_a &= \frac{N \sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{i.}} - \sum_{j=1}^c n_{.j}^2}{N^2 - \sum_{j=1}^c n_{.j}^2} \\ &= \frac{90 \left(\frac{20^2}{30} + \frac{10^2}{30} + \frac{20^2}{60} + \frac{40^2}{60} \right) - (40^2 + 50^2)}{90^2 - (40^2 + 50^2)} \\ &= \frac{400}{4000} = 0.10,\end{aligned}$$

Table 11.12 Example data for χ^2 , t_a , and t_b with $r = 2$ rows, $c = 2$ columns, and $N = 36$ cross-classified observations

Row (B)	Column (A)		Total
	1	2	
1	20	10	30
2	20	40	60
Total	40	50	90

$$\begin{aligned}
 t_b &= \frac{N \sum_{j=1}^c \sum_{i=1}^r \frac{n_{ij}^2}{n_{.j}} - \sum_{i=1}^r n_{i.}^2}{N^2 - \sum_{i=1}^r n_{i.}^2} \\
 &= \frac{90 \left(\frac{20^2}{40} + \frac{10^2}{50} + \frac{20^2}{40} + \frac{40^2}{50} \right) - (30^2 + 60^2)}{90^2 - (30^2 + 60^2)} \\
 &= \frac{360}{3600} = 0.10,
 \end{aligned}$$

and

$$\begin{aligned}
 \chi^2 &= N \left(\sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{i.}n_{.j}} - 1 \right) \\
 &= 90 \left[\frac{20^2}{(30)(40)} + \frac{10^2}{(30)(50)} + \frac{20^2}{(60)(40)} + \frac{40^2}{(60)(50)} \right] \\
 &= 90 (1.1000 - 1) = 9.00.
 \end{aligned}$$

Then the observed value of Pearson’s χ^2 test statistic with respect to the observed value of Goodman and Kruskal’s t_a test statistic is

$$\chi^2 = N t_a = 90(0.10) = 9.00$$

and the observed value of Goodman and Kruskal’s t_a test statistic with respect to the observed value of Pearson’s χ^2 test statistic is

$$t_a = \frac{\chi^2}{N} = \frac{9.00}{90} = 0.10.$$

Also, the observed value of Pearson's χ^2 test statistic with respect to Goodman and Kruskal's t_b test statistic is

$$\chi^2 = Nt_b = 90(0.10) = 9.00$$

and the observed value of Goodman and Kruskal's t_b test statistic with respect to the observed value of Pearson's χ^2 test statistic is

$$t_b = \frac{\chi^2}{N} = \frac{9.00}{90} = 0.10 .$$

Fourth, if $n_{i.} = N/r$ and $n_{.j} = N/c$ for $i = 1, \dots, r$ and $j = 1, \dots, c$, then $\chi^2 = N(c - 1)t_a = N(r - 1)t_b$. To illustrate the relationships between Goodman and Kruskal's t_a and t_b asymmetric measures of nominal association and Pearson's χ^2 test of independence with $n_{i.} = N/r = 12$ for $i = 1, \dots, r$ and $n_{.j} = N/c = 9$ for $j = 1, \dots, c$, consider the frequency data given in Table 11.13 with $r = 3$ rows, $c = 4$ columns, and $N = 36$ cross-classified observations. For the frequency data given in Table 11.13 with $N = 36$ observations,

$$\begin{aligned}
 t_a &= \frac{N \sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{i.}} - \sum_{j=1}^c n_{.j}^2}{N^2 - \sum_{j=1}^c n_{.j}^2} \\
 &= \frac{36 \left(\frac{3^2}{12} + \frac{2^2}{12} + \dots + \frac{3^2}{12} + \frac{4^2}{12} \right) - (9^2 + 9^2 + 9^2 + 9^2)}{36^2 - (9^2 + 9^2 + 9^2 + 9^2)} \\
 &= \frac{24}{972} = 0.0247 ,
 \end{aligned}$$

Table 11.13 Example data for χ^2 , t_a , and t_b with $r = 3$ rows, $c = 4$ columns, and $N = 36$ cross-classified observations

Row (B)	Column (A)				Total
	1	2	3	4	
1	3	2	4	3	12
2	4	4	2	2	12
3	2	3	3	4	12
Total	9	9	9	9	36

$$\begin{aligned}
 t_b &= \frac{N \sum_{j=1}^c \sum_{i=1}^r \frac{n_{ij}^2}{n_{.j}} - \sum_{i=1}^r n_i^2}{N^2 - \sum_{i=1}^r n_i^2} \\
 &= \frac{36 \left(\frac{3^2}{9} + \frac{2^2}{9} + \cdots + \frac{3^2}{9} + \frac{4^2}{9} \right) - (12^2 + 12^2 + 12^2)}{36^2 - (12^2 + 12^2 + 12^2)} \\
 &= \frac{32}{864} = 0.0370,
 \end{aligned}$$

and

$$\begin{aligned}
 \chi^2 &= N \left(\sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{i.} n_{.j}} - 1 \right) \\
 &= 36 \left[\frac{3^2}{(12)(9)} + \frac{2^2}{(12)(9)} + \cdots + \frac{3^2}{(12)(9)} + \frac{4^2}{(12)(9)} \right] \\
 &= 36 \left(\frac{116}{108} - 1 \right) = 2.6667.
 \end{aligned}$$

Then the observed value of Pearson's χ^2 test statistic with respect to the observed value of Goodman and Kruskal's t_a test statistic is

$$\chi^2 = N(c-1)t_a = 36(4-1)(0.0247) = 2.6667$$

and the observed value of Goodman and Kruskal's t_a test statistic with respect to the observed value of Pearson's χ^2 test statistic is

$$t_a = \frac{\chi^2}{N(c-1)} = \frac{2.6667}{36(4-1)} = 0.0247.$$

Also, the observed value of Pearson's χ^2 test statistic with respect to Goodman and Kruskal's t_b test statistic is

$$\chi^2 = N(r-1)t_b = 36(3-1)(0.0370) = 2.6667$$

and the observed value of Goodman and Kruskal's t_b test statistic with respect to the observed value of Pearson's χ^2 test statistic is

$$t_b = \frac{\chi^2}{N(r-1)} = \frac{2.6667}{36(3-1)} = 0.0370.$$

11.2.13 The Relationships Among t_b , δ , and \mathfrak{R}

Goodman and Kruskal's t_b measure of nominal-nominal association is directly related to the permutation test statistic δ and, hence, to the permutation-based, chance-corrected \mathfrak{R} measure of effect size. To illustrate the relationships among test statistics t_b , δ , and \mathfrak{R} , consider the frequency data given in Table 11.9 on p. 426, replicated in Table 11.14 for convenience. The conventional notation for an $r \times c$ contingency table is given in Table 11.8 on p. 423 where the row marginal frequency totals are denoted by $n_{i\cdot}$ for $i = 1, \dots, r$, the column marginal frequency totals are denoted by $n_{\cdot j}$ for $j = 1, \dots, c$, the cell frequencies are denoted by n_{ij} for $i = 1, \dots, r$ and $j = 1, \dots, c$, and

$$N = \sum_{i=1}^r n_{i\cdot} = \sum_{j=1}^c n_{\cdot j} = \sum_{i=1}^r \sum_{j=1}^c n_{ij} .$$

Then for the frequency data given in Table 11.14, Goodman and Kruskal's t_b test statistic is

$$\begin{aligned} t_b &= \frac{N \sum_{j=1}^c \sum_{i=1}^r \frac{n_{ij}^2}{n_{\cdot j}} - \sum_{i=1}^r n_{i\cdot}^2}{N^2 - \sum_{i=1}^r n_{i\cdot}^2} \\ &= \frac{110 \left(\frac{24^2}{25} + \frac{2^2}{25} + \dots + \frac{17^2}{27} + \frac{19^2}{33} \right) - (37^2 + 35^2 + 38^2)}{110^2 - (37^2 + 35^2 + 38^2)} = 0.4428 . \end{aligned}$$

In 1971 Richard Light and Barry Margolin developed test statistic R^2 , based on an analysis of variance technique for categorical response variables [6]. Light and Margolin were unaware that R^2 was identical to Goodman and Kruskal's t_b test statistic and that they had asymptotically solved the long-standing problem of testing the null hypothesis that the population parameter corresponding to Goodman and Kruskal's t_b was zero; that is, $H_0: \tau_b = 0$. The identity between R^2 and t_b was first recognized by Särndal in 1974 [9] and later discussed by Margolin and

Table 11.14 Example data for illustrating the relationships among t_b , δ , and \mathfrak{R} with $r = 3$ rows, $c = 4$ columns, and $N = 110$ cross-classified observations

Row (B)	Column (A)				Total
	1	2	3	4	
1	24	2	5	6	37
2	0	22	5	8	35
3	1	1	17	19	38
Total	25	25	27	33	110

Light [7], where they showed that $t_b(N - 1)(c - 1)$ was distributed as Pearson's chi-squared with $(r - 1)(c - 1)$ degrees of freedom.

Following Light and Margolin in the context of a completely-randomized analysis of variance for the frequency data given in Table 11.14, the sum-of-squares total is

$$SS_{\text{Total}} = \frac{N}{2} - \frac{1}{2N} \sum_{i=1}^r n_i^2$$

$$= \frac{110}{2} - \frac{1}{(2)(110)}(37^2 + 35^2 + 38^2) = 36.6455 ,$$

the sum-of-squares between treatments is

$$SS_{\text{Between}} = \frac{1}{2} \left(\sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}^2}{n_{.j}} \right) - \frac{1}{2N} \sum_{i=1}^r n_i^2$$

$$= \frac{1}{2} \left(\frac{24^2}{25} + \frac{2^2}{25} + \dots + \frac{19^2}{33} \right) - \frac{1}{(2)(110)}(37^2 + 35^2 + 38^2) = 16.2281 ,$$

the sum-of-squares within treatments is

$$SS_{\text{Within}} = \sum_{j=1}^c \left(\frac{n_{.j}}{2} - \frac{1}{2n_{.j}} \sum_{i=1}^r n_{ij}^2 \right)$$

$$= \frac{25}{2} - \frac{1}{(2)(25)}(24^2 + 0^2 + 1^2) + \dots + \frac{33}{2} - \frac{1}{(2)(33)}(6^2 + 8^2 + 19^2)$$

$$= 20.4174 ,$$

and Light and Margolin's test statistic is

$$R^2 = \frac{SS_{\text{Between}}}{SS_{\text{Total}}} = \frac{16.9857}{36.6455} = 0.4428 ,$$

which is identical to Goodman and Kruskal's $t_b = 0.4428$.

The essential factors, sums of squares (SS), degrees of freedom (df), mean squares (MS), and variance-ratio test statistic (F) are summarized in Table 11.15

Table 11.15 Source table for the data listed in Table 11.14

Factor	SS	df	MS	F
Between	16.2281	3	5.4094	28.0835
Within	20.4174	106	0.1926	
Total	36.6455	109		

where $df_{\text{Between}} = c - 1 = 4 - 1 = 3$, $df_{\text{Within}} = N - c = 110 - 4 = 106$, and $df_{\text{Total}} = N - 1 = 110 - 1 = 109$. Under the Neyman–Pearson null hypothesis, $H_0: n_{ij} = n_{i.}/c$ for $i = 1, \dots, r$ and $j = 1, \dots, c$, where each of the c treatment groups possesses the same multinomial probability structure, test statistic F is asymptotically distributed as Snedecor’s F with $\nu_1 = r - 1$ and $\nu_2 = N - r$ degrees of freedom. With $\nu_1 = r - 1 = 4 - 1 = 3$ and $\nu_2 = N - r = 110 - 4 = 106$ degrees of freedom, the asymptotic probability value of $F = 28.0835$ is $P = 0.1917 \times 10^{-12}$, under the assumptions of normality and homogeneity.

For the frequency data given in Table 11.14, the permutation test statistic is

$$\delta = \frac{2SS_{\text{Within}}}{N - c} = \frac{2(20.4174)}{110 - 4} = 0.3852 ,$$

the exact expected value of test statistic δ under the Fisher–Pitman null hypothesis is

$$\mu_{\delta} = \frac{2SS_{\text{Total}}}{N - 1} = \frac{2(36.6455)}{110 - 1} = 0.6724 ,$$

and Mielke and Berry’s chance-corrected measure of effect size is

$$\mathfrak{R} = 1 - \frac{\delta}{\mu_{\delta}} = 1 - \frac{0.3852}{0.6724} = +0.4271 ,$$

indicating approximately 43% agreement between variables A and B above what is expected by chance.

Alternatively, in terms of a completely-randomized analysis of variance model the chance-corrected measure of effect size is

$$\mathfrak{R} = 1 - \frac{(N - 1)(SS_{\text{Within}})}{(N - c)(SS_{\text{Total}})} = 1 - \frac{(110 - 1)(20.4174)}{(110 - 4)(36.6455)} = +0.4271 .$$

Then the observed value of test statistic δ with respect to the observed value of Goodman and Kruskal’s t_b test statistic is

$$\delta = \frac{2SS_{\text{Between}}(1 - t_b)}{t_b(N - c)} = \frac{2(16.2281)(1 - 0.4428)}{(0.4428)(110 - 4)} = 0.3852$$

and the observed value of Goodman and Kruskal’s t_b test statistic with respect to the observed value of test statistic δ is

$$t_b = \frac{2SS_{\text{Between}}}{\delta(N - c) + 2SS_{\text{Between}}} = \frac{2(16.2281)}{(0.3852)(110 - 4) + 2(16.2281)} = 0.4428 .$$

The observed value of test statistic δ with respect to the observed value of Fisher's F -ratio test statistic is

$$\delta = \frac{2SS_{\text{Between}}}{F(c-1)} = \frac{2(16.2281)}{(28.0835)(4-1)} = 0.3852$$

and the observed value of Fisher's F -ratio test statistic with respect to the observed value of test statistic δ is

$$F = \frac{2SS_{\text{Between}}}{\delta(c-1)} = \frac{2(16.2281)}{(0.3852)(4-1)} = 28.0835 .$$

The observed value of Goodman and Kruskal's t_b test statistic with respect to the observed value of Mielke and Berry's \mathfrak{R} measure of effect size is

$$t_b = \frac{\mathfrak{R}(N-c) + c - 1}{N-1} = \frac{(0.4271)(110-4) + 4 - 1}{110-1} = 0.4428$$

and the observed value of Mielke and Berry's \mathfrak{R} measure of effect size with respect to Goodman and Kruskal's t_b test statistic is

$$\mathfrak{R} = 1 - \frac{(N-1)(1-t_b)}{N-c} = 1 - \frac{(110-1)(1-0.4428)}{110-4} = +0.4271 .$$

The observed value of Mielke and Berry's \mathfrak{R} measure of effect size with respect to the observed value of Fisher's F -ratio test statistic is

$$\mathfrak{R} = 1 - \frac{(N-1)SS_{\text{Between}}}{F(c-1)SS_{\text{Total}}} = 1 - \frac{(110-1)(16.2281)}{(28.0835)(4-1)(36.6455)} = +0.4271$$

and the observed value of Fisher's F -ratio test statistic with respect to the observed value of Mielke and Berry's \mathfrak{R} measure of effect size is

$$F = \frac{SS_{\text{Between}}(N-1)}{SS_{\text{Total}}(c-1)(1-\mathfrak{R})} = \frac{(16.2281)(110-1)}{(36.6455)(4-1)(1-0.4271)} = 28.0835 .$$

11.3 Contingency Measures: Ordinal by Ordinal

There exist numerous measures of association for the cross-classification of two ordinal (ranked) variables. Three popular measures of ordinal-ordinal association are Goodman and Kruskal's symmetric measure of ordinal association denoted by G and two asymmetric measures of ordinal association by Somers denoted by

Table 11.16 Example data for Goodman and Kruskal's G measure of ordinal-ordinal association with $r = 3$ rows, $c = 5$ columns, and $N = 63$ cross-classified observations

Row (x)	Column (y)					Total
	1	2	3	4	5	
1	6	2	5	7	1	21
2	0	8	5	8	4	25
3	1	1	6	6	3	17
Total	7	11	16	21	8	63

Table 11.17 Two sets of $N = 8$ rank scores with no tied scores

Object	Variable	
	x	y
1	1	3
2	3	4
3	2	1
4	4	2
5	5	5
6	7	8
7	8	6
8	6	7

d_{yx} and d_{xy} .⁷ These three measures and several others are based on the numbers of concordant and discordant pairs present in the observed contingency table. To illustrate the calculation of concordant and discordant pairs, consider the 3×5 contingency table given in Table 11.16 with $N = 63$ observations.

For any ordered contingency table there are five types of pairs to be considered: concordant pairs (C), discordant pairs (D), pairs that are tied on variable x but not tied on variable y (T_x), pairs tied on variable y but not tied on variable x (T_y), and pairs tied on both variable x and variable y (T_{xy}). Together they sum to the number of possible pairs in the table; that is,

$$C + D + T_x + T_y + T_{xy} = \frac{N(N - 1)}{2} .$$

To demonstrate the calculation of concordant (C) and discordant (D) pairs, consider the two sets of rank scores listed in Table 11.17, where there are no tied ranks. Consider the first pair of objects: Objects 1 and 2. For Object 1, $x_1 = 1$ and $y_1 = 3$, and for Object 2, $x_2 = 3$ and $y_2 = 4$. Since $x_1 < x_2$ and $y_1 < y_2$ ($1 < 3$ and $3 < 4$), the pair is considered to be *concordant*. Now consider a second pair of objects: Objects 1 and 3. For Object 1, $x_1 = 1$ and $y_1 = 3$, and for Object 3, $x_3 = 2$ and $y_3 = 1$. Since $x_1 < x_3$ and $y_1 > y_3$ ($1 < 2$ and $3 > 1$), the

⁷Goodman and Kruskal's G measure of ordinal association is oftentimes denoted by the lower-case Greek letter γ . In this section γ denotes the population parameter and G denotes the sample test statistic.

Table 11.18 Paired differences: concordant (C) and discordant (D) values for the rank scores listed in Table 11.17

Pair	x_i and x_j	y_i and y_j	Type	Pair	x_i and x_j	y_i and y_j	Type
1	1 < 3	3 < 4	C	15	2 < 5	1 < 5	C
2	1 < 2	3 > 1	D	16	2 < 7	1 < 8	C
3	1 < 4	3 > 2	D	17	2 < 8	1 < 6	C
4	1 < 5	3 < 5	C	18	2 < 6	1 < 7	C
5	1 < 7	3 < 8	C	19	4 < 5	2 < 5	C
6	1 < 8	3 < 6	C	20	4 < 7	2 < 8	C
7	1 < 6	3 < 7	C	21	4 < 8	2 < 6	C
8	3 > 2	4 > 1	C	22	4 < 6	2 < 7	C
9	3 < 4	4 > 2	D	23	5 < 7	5 < 8	C
10	3 < 5	4 < 5	C	24	5 < 8	5 < 6	C
11	3 < 7	4 < 8	C	25	5 < 6	5 < 7	C
12	3 < 8	4 < 6	C	26	7 < 8	8 > 6	D
13	3 < 6	4 < 7	C	27	7 > 6	8 > 7	C
14	2 < 4	1 < 2	C	28	8 > 6	6 < 7	D

Table 11.19 Two sets of rank scores with tied scores

Object	Variable	
	x	y
1	1.5	2
2	1.5	2
3	3.5	4.5
4	5.5	2
5	3.5	4.5
6	5.5	6

pair is considered to be *discordant*. For the untied rank data listed in Table 11.17, the number of concordant pairs is $C = 23$ and the number of discordant pairs is $D = 5$. The

$$\frac{N(N - 1)}{2} = \frac{8(8 - 1)}{2} = 28$$

concordant (C) and discordant (D) pairs for the rank-score data listed in Table 11.17 are listed in Table 11.18.

To illustrate the calculation of the T_x , T_y , and T_{xy} tied pairs, consider the two sets of rank scores listed in Table 11.19, where there are multiple tied rank scores on both variable x and variable y . For the rank scores listed in Table 11.19, $N = 6$, the number of concordant pairs is $C = 8$, the number of discordant pairs is $D = 2$, the number of pairs tied on variable x is $T_x = 1$, the number of pairs tied on variable y is $T_y = 2$, and the number of pairs tied on both variable x and variable y is $T_{xy} = 2$.

Table 11.20 Paired differences: C , D , T_x , T_y , and T_{xy} values for the rank scores listed in Table 11.19

Pair	x_i and x_j	y_i and y_j	Type
1	$1.5 = 1.5$	$2.0 = 2.0$	T_{xy}
2	$1.5 < 3.5$	$2.0 < 4.5$	C
3	$1.5 < 5.5$	$2.0 = 2.0$	T_y
4	$1.5 < 3.5$	$2.0 < 4.5$	C
5	$1.5 < 5.5$	$2.0 < 6.0$	C
6	$1.5 < 3.5$	$2.0 < 4.4$	C
7	$1.5 < 5.5$	$2.0 = 2.0$	T_y
8	$1.5 < 3.5$	$2.0 < 4.5$	C
9	$1.5 < 5.5$	$2.0 < 6.0$	C
10	$3.5 < 5.5$	$4.5 > 2.0$	D
11	$3.5 = 3.5$	$4.5 = 4.5$	T_{xy}
12	$3.5 < 5.5$	$4.5 < 6.0$	C
13	$5.5 > 3.5$	$2.0 < 4.5$	D
14	$5.5 = 5.5$	$2.0 < 6.0$	T_x
15	$3.5 < 5.5$	$4.5 < 6.0$	C

Table 11.20 lists the

$$\frac{N(N - 1)}{2} = \frac{6(6 - 1)}{2} = 15$$

paired differences: concordant pairs (C), discordant pairs (D), pairs tied on variable x (T_x), pairs tied on variable y (T_y), and pairs tied on both variable x and variable y (T_{xy}).

11.3.1 An Example Analysis for G

For the example rank data given in Table 11.16 on p. 438 with $N = 63$ observations, the number of concordant pairs is

$$\begin{aligned}
 C &= \sum_{i=1}^{r-1} \sum_{j=1}^{c-1} n_{ij} \left(\sum_{k=i+1}^r \sum_{l=j+1}^c n_{kl} \right) \\
 &= (6)(8 + 5 + 8 + 4 + 1 + 6 + 6 + 3) + (2)(5 + 8 + 4 + 6 + 6 + 3) \\
 &\quad + \cdots + (5)(6 + 3) + (8)(3) = 653,
 \end{aligned}$$

the number of discordant pairs is

$$\begin{aligned}
 D &= \sum_{i=1}^{r-1} \sum_{j=1}^{c-1} n_{i,c-j+1} \left(\sum_{k=i+1}^r \sum_{l=1}^{c-j} n_{kl} \right) \\
 &= (1)(0 + 8 + 5 + 8 + 1 + 1 + 6 + 6) + (7)(0 + 8 + 5 + 1 + 1 + 6) \\
 &\quad + \cdots + (5)(1 + 1) + (8)(1) = 372,
 \end{aligned}$$

and Goodman and Kruskal's measure of ordinal-ordinal association is

$$G = \frac{C - D}{C + D} = \frac{653 - 372}{653 + 372} = +0.2741.$$

Under the Neyman–Pearson null hypothesis, $H_0: \gamma = 0$, Goodman and Kruskal's G measure of ordinal-ordinal association is asymptotically distributed $N(0, 1)$ as $N \rightarrow \infty$ with a standard error given by

$$s_G = \sqrt{\frac{N(1 - G^2)}{C + D}}.$$

For the frequency data given in Table 11.16,

$$z = \frac{G}{\sqrt{\frac{N(1 - G^2)}{C + D}}} = \frac{+0.2741}{\sqrt{\frac{63[1 - (0.2741)^2]}{653 + 372}}} = +1.1496,$$

yielding an asymptotic upper-tail $N(0, 1)$ probability value of $P = 0.1252$, under the assumption of normality.

11.3.2 An Exact Permutation Analysis for G

Under the Fisher–Pitman permutation model, the exact probability value of an observed value of Goodman and Kruskal's G measure of ordinal-ordinal association is given by the sum of the hypergeometric point probability values associated with values of test statistic G that are equal to or greater than the observed value of $G = +0.2741$. For the frequency data given in Table 11.16 with $N = 63$ observations, there are $M = 11,356,797$ possible, equally-likely arrangements in the reference set of all permutations of cell frequencies given the observed row and column marginal frequency distributions $\{21, 25, 17\}$ and $\{7, 11, 16, 21, 8\}$, respectively, making an exact permutation analysis feasible. The exact probability

value of the observed value of test statistic G is $P = 0.0336$; that is, the sum of the hypergeometric point probability values associated with values of $G = +0.2741$ or greater.

11.3.3 The Relationship Between Statistics G and δ

The functional relationships between test statistic δ and Goodman and Kruskal's G measure of ordinal-ordinal association are given by

$$\delta = \frac{N(N-1) - 2G(C+D)}{2N} \quad \text{and} \quad G = \frac{N \left(\frac{N-1}{2} - \delta \right)}{C+D}.$$

For the frequency data given in Table 11.16, the observed value of test statistic δ with respect to the observed value of Goodman and Kruskal's G measure of ordinal-ordinal association is

$$\delta = \frac{63(63-1) - 1(+0.2741)(653+372)}{(2)(63)} = 26.5404$$

and the observed value of Goodman and Kruskal's G measure of ordinal-ordinal association with respect to the observed value of test statistic δ is

$$G = \frac{63 \left(\frac{63-1}{2} - 26.5404 \right)}{653+372} = +0.2741.$$

11.3.4 Somers' d_{yx} and d_{xy} Measures

While Goodman and Kruskal's G measure of ordinal-ordinal association is a symmetric measure, Somers' two asymmetric proportional-reduction-in-error (PRE) measures (d_{yx} and d_{xy}) allow researchers to specify an independent and a dependent variable. For Somers' d_{yx} , the dependent variable is typically the column variable labeled y and for Somers' d_{xy} , the dependent variable is typically the row variable labeled x . The two asymmetric measures are given by

$$d_{yx} = \frac{C-D}{C+D+T_y} \quad \text{and} \quad d_{xy} = \frac{C-D}{C+D-T_x}, \quad (11.4)$$

where C is the number of concordant pairs, D is the number of discordant pairs, T_x is the number of pairs tied on the row variable, and T_y is the number of pairs

tied on the column variable. As is evident in Eq.(11.4), Somers included in the denominators of d_{yx} and d_{xy} the number of tied pairs on the dependent variable: T_y for d_{yx} and T_x for d_{xy} . The rationale for including the tied pairs is simply that when variable y is the dependent variable (d_{yx}), then if two values of the independent variable x differ, but the corresponding two values of the dependent variable y do not differ (are tied), there is evidence of a lack of association and the ties on dependent variable y (T_y) should be included in the denominator where they act to decrease the value of d_{yx} . The same rationale holds for Somers' d_{xy} where the ties on dependent variable x (T_x) are included in the denominator.

11.3.5 An Example Analysis for d_{yx}

For the frequency data given in Table 11.16 on p. 438, replicated in Table 11.21 for convenience, the number of concordant pairs is

$$\begin{aligned}
 C &= \sum_{i=1}^{r-1} \sum_{j=1}^{c-1} n_{ij} \left(\sum_{k=i+1}^r \sum_{l=j+1}^c n_{kl} \right) \\
 &= (6)(8 + 5 + 8 + 4 + 1 + 6 + 6 + 3) + (2)(5 + 8 + 4 + 6 + 6 + 3) \\
 &\quad + \dots + (5)(6 + 3) + (8)(3) = 653 ,
 \end{aligned}$$

the number of discordant pairs is

$$\begin{aligned}
 D &= \sum_{i=1}^{r-1} \sum_{j=1}^{c-1} n_{i,c-j+1} \left(\sum_{k=i+1}^r \sum_{l=1}^{c-j} n_{kl} \right) \\
 &= (1)(0 + 8 + 5 + 8 + 1 + 1 + 6 + 6) + (7)(0 + 8 + 5 + 1 + 1 + 6) \\
 &\quad + \dots + (5)(1 + 1) + (8)(1) = 372 ,
 \end{aligned}$$

Table 11.21 Example data for Somers' d_{yx} and d_{xy} measures of ordinal-ordinal association with $r = 3$ rows, $c = 5$ columns, and $N = 63$ cross-classified observations

Row (x)	Column (y)					Total
	1	2	3	4	5	
1	6	2	5	7	1	21
2	0	8	5	8	4	25
3	1	1	6	6	3	17
Total	7	11	16	21	8	63

the number of pairs tied on variable x is

$$\begin{aligned}
 T_x &= \sum_{i=1}^r \sum_{j=1}^{c-1} n_{ij} \left(\sum_{k=j+1}^c n_{ik} \right) \\
 &= (6)(2 + 5 + 7 + 1) + (2)(5 + 7 + 1) + (5)(7 + 1) + (7)(1) \\
 &\quad + \cdots + (1)(6 + 6 + 3) + (6)(6 + 3) + (6)(3) = 494,
 \end{aligned}$$

the number of pairs tied on dependent variable y is

$$\begin{aligned}
 T_y &= \sum_{j=1}^c \sum_{i=1}^{r-1} n_{ij} \left(\sum_{k=i+1}^r n_{kj} \right) \\
 &= (6)(0 + 1) + (0)(1) + (2)(8 + 1) + (8)(1) \\
 &\quad + \cdots + (1)(4 + 3) + (4)(3) = 282,
 \end{aligned}$$

and Somers' d_{yx} asymmetric measure of ordinal-ordinal association is

$$d_{yx} = \frac{C - D}{C + D + T_y} = \frac{653 - 372}{653 + 372 + 282} = +0.2150.$$

For an $r \times c$ contingency table, d_{yx} is asymptotically distributed $N(0, 1)$ under the Neyman–Pearson null hypothesis as $N \rightarrow \infty$ with a standard error given by

$$s_{d_{yx}} = \frac{2}{3r} \sqrt{\frac{(r^2 - 1)(c + 1)}{N(c - 1)}}.$$

For the frequency data given in Table 11.21,

$$z = \frac{d_{yx}}{\frac{2}{3r} \sqrt{\frac{(r^2 - 1)(c + 1)}{N(c - 1)}}} = \frac{+0.2150}{\frac{2}{(3)(3)} \sqrt{\frac{(3^2 - 1)(5 + 1)}{(63)(5 - 1)}}} = +2.2168,$$

yielding an asymptotic upper-tail $N(0, 1)$ probability value of $P = 0.0133$, under the assumption of normality.

11.3.6 An Exact Permutation Analysis for d_{yx}

Under the Fisher–Pitman permutation model, the exact probability value of an observed value of Somers' d_{yx} is given by the sum of the hypergeometric point

probability values associated with values of test statistic d_{yx} that are equal to or greater than the observed value of $d_{yx} = +0.2150$. For the frequency data given in Table 11.21, there are $M = 11,356,797$ possible, equally-likely arrangements in the reference set of all permutation of cell frequencies given the observed row and column marginal frequency distributions $\{21, 25, 17\}$ and $\{7, 11, 16, 21, 8\}$, respectively, making an exact permutation analysis feasible. The exact probability value of $d_{yx} = +0.2150$ is $P = 0.0331$; that is, the sum of the hypergeometric point probability values associated with values of $d_{yx} = +0.2150$ or greater.

11.3.7 The Relationship Between Statistics d_{yx} and δ

The functional relationships between test statistic δ and Somers' d_{yx} asymmetric measure of ordinal-ordinal association are given by

$$\delta = \frac{N-1}{2} - \frac{d_{yx}(C+D+T_y)}{N} \quad \text{and} \quad d_{yx} = \frac{N\left(\frac{N-1}{2} - \delta\right)}{C+D+T_y}.$$

For the frequency data given in Table 11.21, the observed value of test statistic δ with respect to the observed value of Somers' d_{yx} measure of ordinal-ordinal association is

$$\delta = \frac{63-1}{2} - \frac{+0.2150(653+372+282)}{63} = 26.5396$$

and the observed value of Somers' d_{yx} measure of ordinal-ordinal association with respect to the observed value of test statistic δ is

$$d_{yx} = \frac{63\left(\frac{63-1}{2} - 26.5396\right)}{653+372+282} = +0.2150.$$

11.3.8 An Example Analysis for d_{xy}

For the frequency data given in Table 11.21, the number of concordant pairs is $C = 653$, the number of discordant pairs is $D = 372$, the number of pairs tied on dependent variable x is $T_x = 494$, and Somers' d_{xy} asymmetric measure of ordinal-ordinal association is

$$d_{xy} = \frac{C-D}{C+D+T_x} = \frac{653-372}{653+372+494} = +0.1850.$$

For an $r \times c$ contingency table, d_{yx} is asymptotically distributed $N(0, 1)$ under the Neyman–Pearson null hypothesis as $N \rightarrow \infty$ with a standard error given by

$$s_{d_{yx}} = \frac{2}{3c} \sqrt{\frac{(c^2 - 1)(r + 1)}{N(r - 1)}}.$$

For the frequency data given in Table 11.21,

$$z = \frac{d_{yx}}{\frac{2}{3c} \sqrt{\frac{(c^2 - 1)(r + 1)}{N(r - 1)}}} = \frac{+0.1850}{\frac{2}{(3)(5)} \sqrt{\frac{(5^2 - 1)(3 + 1)}{(63)(5 - 1)}}} = +2.2480,$$

yielding an asymptotic upper-tail $N(0, 1)$ probability value of $P = 0.0123$, under the assumption of normality.

11.3.9 An Exact Permutation Analysis for d_{xy}

Under the Fisher–Pitman permutation model, the exact probability value of an observed value of Somers' d_{xy} is given by the sum of the hypergeometric point probability values associated with values of test statistic d_{xy} that are equal to or greater than the observed value of $d_{xy} = +0.1850$. For the frequency data given in Table 11.21, there are $M = 11,356,797$ possible, equally-likely arrangements in the reference set of all permutation of cell frequencies given the observed row and column marginal frequency distributions $\{21, 25, 17\}$ and $\{7, 11, 16, 21, 8\}$, respectively, making an exact permutation analysis feasible. The exact probability value of $d_{xy} = +0.1850$ is $P = 0.0331$; that is, the sum of the hypergeometric point probability values associated with values of $d_{xy} = +0.1850$ or greater.

11.3.10 The Relationship Between d_{xy} and δ

The functional relationships between test statistic δ and Somers' d_{xy} asymmetric measure of ordinal-ordinal association are given by

$$\delta = \frac{N - 1}{2} - \frac{d_{yx}(C + D + T_x)}{N} \quad \text{and} \quad d_{xy} = \frac{N \left(\frac{N - 1}{2} - \delta \right)}{C + D + T_x}.$$

For the frequency data given in Table 11.21, the observed value of test statistic δ with respect to the observed value of Somers' d_{xy} measure of ordinal-ordinal association is

$$\delta = \frac{63 - 1}{2} - \frac{+0.1850(653 + 372 + 494)}{63} = 26.5394$$

and the observed value of Somers' d_{xy} measure of ordinal association with respect to the observed value of test statistic δ is

$$d_{xy} = \frac{63 \left(\frac{63 - 1}{2} - 26.5394 \right)}{653 + 372 + 494} = +0.1850 .$$

11.3.11 Probability Values for d_{yx} and d_{xy}

It may appear inconsistent that while Somers' two measures of effect size differ ($d_{yx} = +0.2150$ and $d_{xy} = 0.1850$), they both yield the same probability value of $P = 0.0331$. It follows from the fact that the denominators of d_{yx} and d_{xy} ($C + D + T_y$ and $C + D + T_x$, respectively) can be computed from just the marginal frequency distributions, which are fixed for all possible arrangements of cell frequencies and are, therefore, invariant under permutation.

It is easily shown that $C + D + T_y$ can be obtained from N and the row marginal frequency distribution. Recall that for the frequency data listed in Table 11.21 on p. 443, the number of concordant pairs is $C = 653$, the number of discordant pairs is $D = 372$, the number of pairs tied on variable y is $T_y = 282$, and $C + D + T_y = 653 + 372 + 282 = 1307$. Then with $N = 63$,

$$C + D + T_y = \frac{1}{2} \left(N^2 - \sum_{i=1}^r n_i^2 \right) = \frac{1}{2} \left[63^2 - (21^2 + 25^2 + 17^2) \right] = 1307 .$$

In such manner $C + D + T_x$ can be obtained from N and the column marginal frequency distribution. For the frequency data listed in Table 11.21, the number of concordant pairs is $C = 653$, the number of discordant pairs is $D = 372$, the number of pairs tied on variable x is $T_x = 494$, and $C + D + T_x = 653 + 372 + 494 = 1519$. Then with $N = 63$,

$$\begin{aligned} C + D + T_x &= \frac{1}{2} \left(N^2 - \sum_{j=1}^c n_j^2 \right) \\ &= \frac{1}{2} \left[63^2 - (7^2 + 11^2 + 16^2 + 21^2 + 8^2) \right] = 1519 . \end{aligned}$$

11.4 Contingency Measures: Nominal by Ordinal

There exist any number of measures of association for which the standard error is unknown. Permutation statistical methods do not rely on knowledge of standard errors and therefore provide much-needed probability values for a number of otherwise very useful measures of association. One measure without a known standard error is Freeman’s θ measure of nominal-ordinal association [5, pp. 108–119].

Consider an $r \times c$ contingency table where the r rows are a nominal-level (categorical) independent variable (x) and the c columns are an ordinal-level (ranked) dependent variable (y). For Freeman’s θ it is necessary to calculate the absolute sum of the number of concordant pairs and number of discordant pairs for all combinations of the nominal-level independent variable (rows) considered two at a time. Assuming that the column ordered variable (y) is underlying continuous and that ties in ranking result simply from a crude classification of the continuous variable, Freeman’s nominal-ordinal measure of association is given by

$$\theta = \frac{\sum_{i=1}^{r-1} \sum_{j=i+1}^r |C_{ij} - D_{ij}|}{C + D + T_y} .$$

11.4.1 An Example Analysis for θ

To illustrate the calculation of Freeman’s θ measure of nominal-ordinal association, consider the 4×5 contingency table given in Table 11.22 with $N = 40$ observations. For the frequency data given in Table 11.22 with $N = 40$ observations, the number of concordant pairs is

$$C = \sum_{i=1}^{r-1} \sum_{j=1}^{c-1} n_{ij} \left(\sum_{k=i+1}^r \sum_{l=j+1}^c n_{kl} \right) = (1)(5 + 5 + 0 + 0 + 0 + 2 + 2 + 1 + 0 + 0 + 2 + 3)$$

Table 11.22 Example data for Freeman’s θ measure of nominal-ordinal association with $r = 4$ rows, $c = 5$ columns, and $N = 40$ cross-classified observations

Row (x)	Column (y)					Total
	1	2	3	4	5	
1	1	2	5	2	0	10
2	10	5	5	0	0	20
3	0	0	2	2	1	5
4	0	0	0	2	3	5
Total	11	7	12	6	4	40

$$\begin{aligned}
 &+ (2)(5 + 0 + 0 + 2 + 2 + 1 + 0 + 2 + 3) \\
 &+ \cdots + (2)(2 + 3) + (2)(3) = 304 ,
 \end{aligned}$$

the number of discordant pairs is

$$\begin{aligned}
 D &= \sum_{i=1}^{r-1} \sum_{j=1}^{c-1} n_{i,c-j+1} \left(\sum_{k=i+1}^r \sum_{l=1}^{c-j} n_{kl} \right) \\
 &= (0)(10 + 5 + 5 + 0 + 0 + 0 + 2 + 2 + 0 + 0 + 0 + 2) \\
 &+ (2)(10 + 5 + 6 + 0 + 0 + 2 + 0 + 0 + 0) \\
 &+ \cdots + (2)(0 + 0) + (0)(0) = 141 ,
 \end{aligned}$$

the number of pairs tied on variable y is

$$\begin{aligned}
 T_y &= \sum_{j=1}^c \sum_{i=1}^{r-1} n_{ij} \left(\sum_{k=i+1}^r n_{kj} \right) \\
 &= (1)(10 + 0 + 0) + (10)(0 + 0) + (0)(0) \\
 &+ \cdots + (0)(0 + 1 + 3) + (1)(1 + 3) + (1)(3) = 80 ,
 \end{aligned}$$

the concordant and discordant pairs for the $r = 4$ rows considered two at a time are

$$\begin{aligned}
 C_{12} &= (1)(5 + 5 + 0 + 0) + (2)(5 + 0 + 0) + (5)(0 + 0) + (2)(0) = 20 , \\
 D_{12} &= (0)(10 + 5 + 5 + 0) + (2)(10 + 5 + 5) + (5)(10 + 5) + (2)(10) = 135 , \\
 C_{13} &= (1)(0 + 2 + 2 + 1) + (2)(2 + 2 + 1) + (5)(2 + 1) + (2)(1) = 32 , \\
 D_{13} &= (0)(0 + 0 + 2 + 2) + (2)(0 + 0 + 2) + (5)(0 + 0) + (2)(0) = 4 , \\
 C_{14} &= (1)(0 + 0 + 2 + 3) + (2)(0 + 2 + 3) + (5)(2 + 3) + (2)(3) = 46 , \\
 D_{14} &= (0)(0 + 0 + 0 + 2) + (2)(0 + 0 + 0) + (5)(0 + 0) + (2)(0) = 0 , \\
 C_{23} &= (10)(0 + 2 + 2 + 1) + (5)(2 + 2 + 1) + (5)(2 + 1) + (0)(1) = 90 , \\
 D_{23} &= (0)(0 + 0 + 2 + 2) + (0)(0 + 0 + 2) + (5)(0 + 0) + (5)(0) = 0 , \\
 C_{24} &= (10)(0 + 0 + 2 + 3) + (5)(0 + 2 + 3) + (5)(2 + 3) + (0)(3) = 100 . \\
 D_{24} &= (0)(0 + 0 + 0 + 2) + (0)(0 + 0 + 0) + (5)(0 + 0) + (5)(0) = 0 , \\
 C_{34} &= (0)(0 + 0 + 2 + 3) + (0)(0 + 2 + 3) + (2)(2 + 3) + (2)(3) = 16 , \\
 D_{34} &= (1)(0 + 0 + 0 + 2) + (2)(0 + 0 + 0) + (2)(0 + 0) + (0)(0) = 2 ,
 \end{aligned}$$

and Freeman's θ is

$$\begin{aligned}\theta &= \frac{\sum_{i=1}^{r-1} \sum_{j=i+1}^r |C_{ij} - D_{ij}|}{C + D + T_y} \\ &= \frac{|20 - 135| + |32 - 4| + |46 - 0| + |90 - 0| + |100 - 0| + |16 - 2|}{304 + 141 + 80} \\ &= 0.7486.\end{aligned}$$

11.4.2 An Exact Permutation Analysis for θ

Under the Fisher–Pitman permutation model, the exact probability value of an observed value of $\theta = 0.7486$ is given by the sum of the hypergeometric point probability values associated with the values of test statistic θ calculated on all M possible arrangements of the cell frequencies that are equal to or greater than the observed value of $\theta = 0.7486$. For the frequency data given in Table 11.22, there are only $M = 6,340,588$ possible arrangements in the reference set of all permutations of cell frequencies consistent with the observed row and column marginal frequency distributions, $\{10, 20, 5, 5\}$ and $\{11, 7, 12, 6, 4\}$, respectively, making an exact permutation analysis feasible.

If all M possible arrangements of the observed data occur with equal chance, the exact probability value of Freeman's θ under the Fisher–Pitman null hypothesis is the sum of the hypergeometric point probability values associated with the arrangements of cell frequencies with values of θ that are equal to or greater than the observed value of $\theta = 0.7486$. Based on the underlying hypergeometric probability distribution, the exact probability value of $\theta = 0.7486$ is $P = 0.2105 \times 10^{-10}$.

11.5 Contingency Measures: Nominal by Interval

Pearson's point-biserial correlation coefficient, denoted by r_{pb} , measures the association between a nominal-level (categorical) variable with two categories and an interval-level variable. Pearson's point-biserial correlation coefficient is an important measure in fields such as education and educational psychology where it is typically used to measure the correlation between test questions scored as correct (1) or incorrect (0) and the overall score on the test for N test takers. A low or negative point-biserial correlation coefficient indicates that the test takers with the highest scores on the test answered the question incorrectly and the test takers with the lowest scores on the test answered the question correctly, alerting the instructor to the possibility that the question failed to discriminate properly and may be faulty.

Table 11.23 Example (0, 1) coded data for Pearson's point-biserial correlation coefficient

Object	Variable		Object	Variable	
	x	y		x	y
1	0	99	11	1	86
2	0	99	12	1	90
3	1	98	13	0	97
4	1	98	14	0	95
5	1	97	15	1	92
6	0	89	16	0	98
7	0	95	17	1	86
8	0	94	18	1	85
9	1	92	19	0	94
10	1	60	20	0	96

11.5.1 An Example Analysis for r_{pb}

To illustrate the calculation of Pearson's point-biserial correlation coefficient, consider the dichotomous data listed in Table 11.23 for $N = 20$ observations where variable x is the dichotomous variable and variable y is an unspecified interval-level variable. The point-biserial correlation coefficient is often expressed as

$$r_{pb} = \frac{\bar{y}_0 - \bar{y}_1}{s_y} \sqrt{\frac{n_0 n_1}{N(N-1)}},$$

where n_0 and n_1 denote the number of y values coded 0 and 1, respectively, $N = n_0 + n_1$, \bar{y}_0 and \bar{y}_1 denote the means of the y values coded 0 and 1, respectively, and s_y is the sample standard deviation of the y values given by

$$s_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2}.$$

For the data listed in Table 11.23, $n_0 = n_1 = 10$,

$$\bar{y}_0 = \frac{1}{n_0} \sum_{i=1}^{n_0} y_i = \frac{99 + 99 + \cdots + 89}{10} = 88.40,$$

$$\bar{y}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} y_i = \frac{98 + 98 + \cdots + 60}{10} = 95.60,$$

$$s_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2} = \sqrt{\frac{1456}{20-1}} = 8.7539,$$

and Pearson's point-biserial correlation coefficient is

$$r_{pb} = \frac{\bar{y}_0 - \bar{y}_1}{s_y} \sqrt{\frac{n_0 n_1}{N(N-1)}} = \frac{88.40 - 95.60}{8.7539} \sqrt{\frac{(10)(10)}{20(20-1)}} = -0.4219.$$

Alternatively, with

$$\begin{aligned} \sum_{i=1}^N x_i &= 10, & \sum_{i=1}^N x_i^2 &= 10, & \sum_{i=1}^N y_i &= 1840, & \sum_{i=1}^N y_i^2 &= 170,736, \\ & & & & \text{and} & & \sum_{i=1}^N x_i y_i &= 884, \end{aligned}$$

Pearson's point-biserial correlation coefficient is simply the product-moment correlation between dichotomous variable x and interval-level variable y . Thus,

$$\begin{aligned} r_{pb} &= \frac{N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{\sqrt{\left[N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2 \right] \left[N \sum_{i=1}^N y_i^2 - \left(\sum_{i=1}^N y_i \right)^2 \right]}} \\ &= \frac{(20)(884) - (10)(1840)}{\sqrt{[(20)(10) - 10^2][(20)(170,736) - 1840^2]}} = -0.4219. \end{aligned}$$

The conventional test of significance for Pearson's point-biserial correlation coefficient is

$$t = r_{pb} \sqrt{\frac{N-2}{1-r_{pb}^2}} = -0.4219 \sqrt{\frac{20-2}{1-(-0.4219)^2}} = -1.9743.$$

Under the Neyman–Pearson null hypothesis, $H_0: \rho_{pb} = 0$, test statistic t is asymptotically distributed as Student's t with $N - 2$ degrees of freedom. With $N - 2 = 20 - 2 = 18$ degrees of freedom, the asymptotic two-tail probability value of $t = -1.9743$ is $P = 0.0639$, under the assumption of normality. For a critical evaluation of the point-biserial correlation coefficient, see a discussion in *The Measurement of Association* by the authors [2, pp. 417–424].

11.5.2 An Exact Permutation Analysis for r_{pb}

For the bivariate observations listed in Table 11.23, there are only

$$M = \frac{(n_0 + n_1)!}{n_0! n_1!} = \frac{(10 + 10)!}{10! 10!} = 184,756$$

possible, equally-likely arrangements in the reference set of all permutations of the observed scores, making an exact permutation analysis possible. Under the Fisher–Pitman permutation model, the exact probability of an observed value of Pearson’s $|r_{pb}|$ is the proportion of $|r_{pb}|$ values calculated on all possible arrangements of the observed data that are equal to or greater than the observed value of $|r_{pb}| = 0.4219$. There are exactly 11,296 $|r_{pb}|$ values that are equal to or greater than the observed value of $|r_{pb}| = 0.4219$. If all arrangements of the $N = 20$ observed scores occur with equal chance, the exact probability value of $|r_{pb}| = 0.4219$ computed on the $M = 184,756$ possible arrangements of the observed data with $n_0 = n_1 = 10$ preserved for each arrangement is

$$P(r_{pb} \geq |r_0|) = \frac{\text{number of } r_{pb} \text{ values } \geq |r_0|}{M} = \frac{11,296}{184,756} = 0.0611 ,$$

where $|r_0|$ denotes the observed absolute value of test statistic r_{pb} and M is the number of possible, equally-likely arrangements of the $N = 20$ bivariate observations listed in Table 11.23.

11.6 Contingency Measures: Ordinal by Interval

The best-known and most-widely reported measure of ordinal-by-interval association is Jaspén’s multiserial correlation coefficient, which is simply the Pearson product-moment correlation coefficient between an interval-level variable, Y , and a transformation of an ordinal-level variable, X . Given N values on the interval variable and k disjoint, ordered categories on the ordinal variable, the mean standard score of the underlying scale for a given category is given by

$$\bar{Z}_j = \frac{Y_{L_j} - Y_{U_j}}{p_j} \quad \text{for } j = 1, \dots, k ,$$

where Y_{L_j} and Y_{U_j} are the lower and upper ordinates of the segment of the $N(0, 1)$ distribution corresponding to the j th ordered category, and where p_j is the proportion of cases in the j th of k ordered categories. Given the obtained values of \bar{Z}_j , $j = 1, \dots, k$, and the original N values of the interval-level variable, a standard Pearson product-moment correlation between the Y and \bar{Z} values yields the multiserial correlation coefficient.

11.6.1 An Example Analysis for $r_{Y\bar{Z}}$

To illustrate the calculation of Jaspens’s multiserial correlation coefficient, consider the small set of data given in Table 11.24 where $N = 32$ interval-level variables are listed in $k = 4$ disjoint, ordered categories: $A, B, C,$ and D . Table 11.25 illustrates the calculation of Jaspens’s multiserial correlation coefficient. The first column, headed X in Table 11.25, lists the $k = 4$ ordered categories of variable X . The second column, headed n , lists the number of observations in each of the k ordered categories. The third column, headed p , lists the proportion of observations in each of the k ordered categories. The fourth column, headed P , lists the cumulative proportion of observations in each of the k ordered categories. The fifth column, headed z , lists the standard score that defines the cumulative proportion from the fourth column under the unit-normal distribution for each of the k ordered categories. For example, for category A the standard score that defines the lowest (left-tail) of the normal distribution with proportion $P = 0.1250$ is $z = -1.1503$. The sixth column, headed Y_L , lists the height of the ordinate at the standard score listed in the fifth column *below* the specified segment of the unit-normal distribution. For example, for category A ,

$$Y_{L_A} = \frac{\exp(-z^2/2)}{\sqrt{2\pi}} = \frac{\exp[-(-1.1503)^2/2]}{\sqrt{2(3.1416)}} = 0.2059 .$$

Table 11.24 Example ordinal-by-interval data for Jaspens’s correlation coefficient with $N = 32$ observations

Category			
A	B	C	D
83	91	86	75
78	84	81	58
73	81	80	51
63	78	79	50
	76	77	50
	73	76	48
	69	70	48
	64	64	
	58	63	
	56	59	
		53	

Table 11.25 Calculation of the mean standard scores for the $k = 4$ ordinal categories

X	n	p	P	z	Y_L	Y_U	\bar{Z}
A	4	0.1250	0.1250	-1.1503	0.2059	0.0000	+1.6472
B	10	0.3125	0.4375	-0.1573	0.3940	0.2059	+0.6019
C	11	0.3438	0.7813	+0.7766	0.2951	0.3940	-0.2877
D	7	0.2188	1.0000	+1.0000	0.0000	0.2951	-1.3487
Total	32	1.0000					

The seventh column, headed Y_U , lists the height of the ordinate at the standard score listed in the fifth column *above* the specified segment of the unit-normal distribution. For example, for category C ,

$$Y_{U_C} = \frac{\exp(-z^2/2)}{\sqrt{2\pi}} = \frac{\exp[-(-0.1573)^2/2]}{\sqrt{2(3.1416)}} = 0.3940.$$

The last column, headed \bar{Z} , lists the average standard scores for the k ordered categories. For example, for category B ,

$$\bar{Z}_B = \frac{Y_{L_B} - Y_{U_B}}{p_B} = \frac{0.3940 - 0.2059}{0.3125} = +0.6019$$

Jaspen's multiserial correlation coefficient is the Pearson product-moment correlation between the Y interval-level values given in Table 11.24 and the transformed \bar{Z} values given in Table 11.25. Table 11.26 lists the Y , \bar{Z} , Y^2 , \bar{Z}^2 , and $Y\bar{Z}$ values, along with the corresponding sums.

For the summations given in Table 11.26, the Pearson product-moment correlation between the Y and \bar{Z} values is

$$\begin{aligned} r_{Y\bar{Z}} &= \frac{N \sum_{i=1}^N Y_i \bar{Z}_i - \sum_{i=1}^N Y_i \sum_{i=1}^N \bar{Z}_i}{\sqrt{\left[N \sum_{i=1}^N Y_i^2 - \left(\sum_{i=1}^N Y_i \right)^2 \right] \left[N \sum_{i=1}^N \bar{Z}_i^2 - \left(\sum_{i=1}^N \bar{Z}_i \right)^2 \right]}} \\ &= \frac{(32)(189.3918) - (2195)(0.0000)}{\sqrt{[(32)(155,471) - 2195^2][(32)(28.1193) - 0.0000^2]}} = +0.5094. \end{aligned}$$

Jaspen's multiserial correlation coefficient is known to be biased. The bias is due to the grouping of the values into k categories. When k is small, the bias can be pronounced. For the example data listed in Table 11.24, the correction for grouping is

$$\begin{aligned} S_{\bar{Z}} &= \left(\frac{1}{N} \sum_{j=1}^k n_j \bar{Z}_j^2 \right)^{1/2} \\ &= \left\{ \frac{1}{32} \left[(4)(+1.6472)^2 + (10)(+0.6019)^2 + (11)(-0.2877)^2 \right. \right. \\ &\quad \left. \left. + (7)(-1.3487)^2 \right] \right\}^{1/2} = 0.9374 \end{aligned}$$

Table 11.26 Calculation of the sums needed for the Pearson product-moment correlation between variables Y and \bar{Z}

Category	Y	\bar{Z}	Y^2	\bar{Z}^2	$Y\bar{Z}$
A	83	+1.6472	6889	2.7133	+136.7176
	78	+1.6472	6084	2.7133	+128.4816
	73	+1.6472	5329	2.7133	+120.2456
	63	+1.6472	3969	2.7133	+103.7736
B	91	+0.6019	8281	0.3623	+54.7729
	84	+0.6019	7056	0.3623	+50.5596
	81	+0.6019	6561	0.3623	+48.7539
	78	+0.6019	6084	0.3623	+46.9482
	76	+0.6019	5776	0.3623	+45.7444
	73	+0.6019	5329	0.3623	+43.9387
	69	+0.6019	4761	0.3623	+41.5311
	64	+0.6019	4096	0.3623	+38.5216
	58	+0.6019	3364	0.3623	+34.9102
	56	+0.6019	3136	0.3623	+33.7064
C	86	-0.2877	7396	0.0828	-23.8220
	81	-0.2877	6561	0.0828	-22.4370
	80	-0.2877	6400	0.0828	-22.1600
	79	-0.2877	6241	0.0828	-21.8830
	77	-0.2877	5929	0.0828	-21.3290
	76	-0.2877	5776	0.0828	-21.0520
	70	-0.2877	4900	0.0828	-19.3900
	64	-0.2877	4096	0.0828	-17.7280
	63	-0.2877	3969	0.0828	-17.4510
	59	-0.2877	3481	0.0828	-16.3430
	53	-0.2877	2809	0.0828	-14.6810
D	75	-1.3487	5625	1.8190	-101.1525
	58	-1.3487	3364	1.8190	-78.2246
	51	-1.3487	2601	1.8190	-68.7837
	50	-1.3487	2500	1.8190	-67.4350
	50	-1.3487	2500	1.8190	-67.4350
	48	-1.3487	2304	1.8190	-64.7376
	48	-1.3487	2304	1.8190	-64.7376
Sum	2195	0.0000	155,471	28.1193	+189.3918

and the corrected multiserial correlation coefficient is

$$r_c = \frac{r_{Y\bar{Z}}}{S_{\bar{Z}}} = \frac{+0.5094}{0.9374} = +0.5434 .$$

Jaspen's r_c is asymptotically distributed as Student's t under the Neyman-Pearson null hypothesis with $N - 2$ degrees of freedom. If the population parameter,

ρ_c , is assumed to be zero, then for the observed data in Table 11.24,

$$t = \frac{r_c - \rho_c}{\sqrt{\frac{1 - r_c^2}{N - 2}}} = \frac{+0.5434 - 0.00}{\sqrt{\frac{1 - (0.5434)^2}{32 - 2}}} = +3.5455 ,$$

and with $N - 2 = 32 - 2 = 30$ degrees of freedom the asymptotic two-tail probability value of $r_c = +0.5434$ is $P = 0.1308 \times 10^{-2}$, under the assumption of normality.

11.6.2 A Monte Carlo Permutation Analysis for $r_{Y\bar{Z}}$

Because there are

$$M = N! = 32! = 263,130,836,933,693,530,167,218,012,160,000,000$$

possible, equally-likely arrangements in the reference set of all permutations of the observed values listed in Table 11.24, an exact permutation analysis is not possible and a Monte Carlo analysis is mandated. Let r_o indicate the observed value of r_c . Based on $L = 1,000,000$ randomly-selected arrangements of the observed data, there are 3069 $|r_c|$ values that are equal to or greater than $|r_o| = 0.5434$, yielding a Monte Carlo probability value of

$$P(|r_c| \geq |r_o|) = \frac{\text{number of } |r_c| \text{ values } \geq |r_o|}{L} = \frac{3069}{1,000,000} = 0.3069 \times 10^{-2} ,$$

where r_o denotes the observed value of r_c and L is the number of randomly-selected, equally-likely arrangements of the ordinal-interval data listed in Table 11.24.

11.7 Summary

Under the Neyman–Pearson model of statistical inference, this chapter examined various measures of nominal-nominal, ordinal-ordinal, nominal-ordinal, nominal-interval, and ordinal-interval association. Asymptotic probability values were provided under either Pearson's χ^2 probability distribution, Student's t distribution, Snedecor's F distribution, or the $N(0, 1)$ probability distribution. Under the Fisher–Pitman permutation model of statistical inference, procedures for both exact and Monte Carlo probability values were developed.

Six sections provided examples and illustrative analyses of permutation statistical methods for contingency tables. In the first section, goodness-of-fit measures for

k discrete, mutually-exclusive categories were described and permutation methods were utilized to obtain exact probability values. A measure of effect size for the chi-squared goodness-of-fit test was presented and illustrated.

The second section illustrated the use of permutation statistical methods for analyzing contingency tables in which two nominal-level variables have been cross-classified. Three well-known and widely-used measures of nominal-nominal association were introduced and analyzed with permutation methods: Cramér's symmetric V measure, based on Pearson's chi-squared test statistic, and Goodman and Kruskal's t_a and t_b asymmetric measures, based on the differences between concordant and discordant pairs of observations. The relationships between Pearson's χ^2 test statistic and Goodman and Kruskal's t_a and t_b measures were described.

The third section illustrated the use of permutation statistical methods for analyzing contingency tables in which two ordinal-level variables have been cross-classified. Three popular measures of ordinal-ordinal association were introduced and analyzed with permutation methods: Goodman and Kruskal's G symmetric measure of ordinal-ordinal association and Somers' d_{yx} and d_{xy} asymmetric measures of ordinal-ordinal association.

The fourth section illustrated the use of permutation statistical methods for analyzing contingency tables in which a nominal-level variable was cross-classified with an ordinal-level variable. Freeman's θ measure of association for a nominal-level independent variable and an ordinal-level dependent variable was described and illustrated with exact permutation statistical methods.

The fifth section illustrated the use of permutation statistical methods for analyzing contingency tables in which a nominal-level variable was cross-classified with an interval-level variable. Pearson's point-biserial correlation coefficient for a dichotomous nominal-level variable and a continuous interval-level variable was described and analyzed with exact permutation statistical methods.

The sixth section illustrated the use of permutation statistical methods for analyzing contingency tables in which an ordinal-level variable was cross-classified with an interval-level variable. Jaspert's multi-serial correlation coefficient for ordinal-interval association was described and analyzed with Monte Carlo permutation methods.

References

1. Berry, K.J., Johnston, J.E., Mielke, P.W.: Exact goodness-of-fit tests for unordered equiprobable categories. *Percept. Motor Skill.* **98**, 909–918 (2004)
2. Berry, K.J., Johnston, J.E., Mielke, P.W.: *The Measurement of Association: A Permutation Statistical Approach*. Springer, Cham (2018)
3. Box, J.F.: *R. A. Fisher: The Life of a Scientist*. Wiley, New York (1978)
4. Fisher, R.: Has Mendel's work been rediscovered? *Ann. Sci.* **1**, 115–137 (1936)
5. Freeman, L.C.: *Elementary Applied Statistics*. Wiley, New York (1965)
6. Light, R.J.: Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychol. Bull.* **76**, 365–377 (1971)

7. Margolin, B.H., Light, R.J.: An analysis of variance for categorical data, II: small sample comparisons with chi square and other competitors. *J. Am. Stat. Assoc.* **69**, 755–764 (1974)
8. Mielke, P.W., Berry, K.J.: Exact goodness-of-fit probability tests for analyzing categorical data. *Educ. Psychol. Meas.* **53**, 707–710 (1993)
9. Särndal, C.E.: A comparative study of association measures. *Psychometrika* **39**, 165–187 (1974)