Bulletin of the Psychonomic Society
1983, Vol. 21 (1), 76-78

"Spurious nonsignificance" in rank correlation

L. W. BUCKALEW and W. H. PEARSON
Air Force Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio 45433

Spearman's rho and Kendall's tau rank correlations are acknowledged as highly utilitarian for a discipline increasingly responsive to the ordinal properties of much of its measurements. Rho and tau are conceptually reviewed, and concern is noted for the possibility that these rank correlation coefficients could be nonsignificant when underlying variables are strongly associated in a nonlinear fashion. An example of "spurious nonsignificance" is provided in which exploratory techniques involving a scattergram and the one-sample runs test are used to discern the true relationship of variables. Emphasis is placed on avoiding possible loss or misinterpretation of information in many nonsignificant rank correlations.

Both Spearman's rank correlation (rho) and Kendall's rank correlation (tau) hold high utility value for psychological research, particularly with regard to increased recognition of the ordinal properties of much psychological measurement. With this recognition has come concern for extending the information value and interpretation potentials of popular nonparametric techniques (Buckalew & Pearson, 1981, 1982a, 1982b). Present concern is with the possible loss or misinterpretation of information associated with a nonsignificant rho or tau predicated on an H_0 assumption of linearity in two sets of ranks. Pointedly, it is contended that a rho or tau correlation coefficient could be nonsignificant, indeed 0, when a curvilinear relationship exists between the variables. Typically, this relationship would go undetected, resulting in a case of "spurious nonsignificance."

Siegel (1956) suggested that, of all statistics based on ranks, rho is the earliest developed and perhaps best known. The computation of rho requires that the two variables measured are at least ordinal to facilitate ranking in two ordered series. Typically, N individuals are ranked according to two variables, X and Y. The actual mathematical value of rho, free to vary from −1 to +1, is a function of the ratio of squared differences between paired ranks (d^2) to the number of pairs of ranks (N). The major prerequisite in computation is that the data of X and Y have been transformed into ranks. Tau is a suitable measure of correlation for the same sort of data. It provides a measure of the degree of association between two sets of ranks, and is typically applied to reflect the correspondence between judgments (rankings) of two judges, X and Y. The actual mathematical value of tau is a function of the statistic S, an index of disarray between two sets of ranks. In essence, tau is a ratio of how many pairs of ranks in the Y set, following ordering of the X set, are in natural order, compared to the maximum possible.

Reprint requests may be addressed to L. W. Buckalew, Department of Psychology, Box 200, Alabama A & M University, Normal, Alabama 35762, or W. H. Pearson, AFAMRL/HEG, Wright-Patterson AFB, Ohio 45433.

As discussed by Siegel (1956), both rho and tau use the same amount of information in the data and have the same power to detect the existence of association in the population. However, these statistics are not numerically comparable, as they have different underlying data: Rho deals directly with differences in pairs of ranks, whereas tau deals with the number of inversions needed to transform one ranking into another. Consequently, when rho and tau are computed from the same data (pairs of rankings), numerical values are dissimilar but highly correlated (Ferguson, 1980), and their respective sampling distributions are such that the null hypothesis would be rejected at the same level of significance. Ferguson notes that the distribution of tau approaches normal form more rapidly than that of rho, tau is more amenable to mathematical manipulation, and its S component has applications beyond correlation.

Given the functional and power similarity of rho and tau, it is of interest to explore the "spurious nonsignificance" potential. Although rho will be used as an example, it is suggested that tau may be no less vulnerable. Assume that for 16 persons there exist IQ (X) and anxiety (Y) test scores. The hypothesis, H_1, is entertained that a relationship exists between these variables. Test scores are replaced with ranks, with 1 assigned the lowest score, 16 the highest, and ties conventionally treated. Spearman rank correlation for this data, displayed in Table 1, will be −.043, which is clearly nonsignificant. From the conventional (linear) perspective, the only conclusion warranted is that there is no relationship between intelligence and anxiety scores, hence rejection of H_1 and closure to analysis.

Next, consider the possibility of "spurious nonsignificance." The true relationship between X and Y is dramatically seen if scores are plotted in a scattergram. Figure 1 provides a very different perspective on this "nonsignificant" association indicated by rho, as clearly there appears to be the potential of a meaningful curvilinear-type relationship. To test whether these scores are truly randomly related, as per rho, the ranks from Table 1 must first be arranged so that one of the variables is presented in ordered fashion. Table 2 reflects this operation. Next, assuming X is so arranged, the sign (+ or –) reflecting how each X rank relates to its corresponding Y rank is entered. If the nonsignificant association indicated by rho is legitimate, there should be an essentially random series of signs, as each would be equally likely.

It is offered that the one-sample runs test, as described by Siegel (1956), may be applied to this series of signs to test its randomness. The number of runs, as a succession of identical symbols, may be counted, as shown in Table 2. The total number of runs in a sample of any given size gives an indication of whether or not the sample is random. In this test, n_1 and n_2 represent the number of elements (signs) of each kind that exist. In Table 2 data, there are 4 runs (r), 6 pluses (n_1), and 10 minuses (n_2). In the tables of critical values of r provided in Siegel (1956), the obtained r is found to be significant at the .05 level. It must be concluded that this series of signs, and hence the underlying data, are not random, as suggested by the nonsignificant rho, but that it indicates grouping due to a lack of independence.

Table 1
Data for Calculation of Spearman's Coefficient of Correlation

Subject	Intelligence (X)	Rank	Anxiety (Y)	Rank
1	72	2	1	2.5
2	119	13	3	8.5
3	84	4	3	8.5
4	93	6	4	12
5	131	15	1	2.5
6	98	7	5	15
7	89	5	4	12
8	68	1	1	2.5
9	102	8	5	15
10	111	11	4	12
11	137	16	1	2.5
12	116	12	3	8.5
13	78	3	2	5.5
14	125	14	2	5.5
15	106	9	5	15
16	108	10	3	8.5

Table 2
Ordered Pairs for One-Sample Runs Test

Subject	X Rank	Y Rank	Sign	Runs
8	1	2.5	–	
1	2	2.5	–	
13	3	5.5	–	
3	4	8.5	–	
7	5	12	–	1
4	6	12	–	
6	7	15	–	
9	8	15	–	
15	9	15	–	
16	10	8.5	+	2
10	11	12	–	3
12	12	8.5	+	
2	13	8.5	+	
14	14	5.5	+	4
5	15	2.5	+	
11	16	2.5	+	

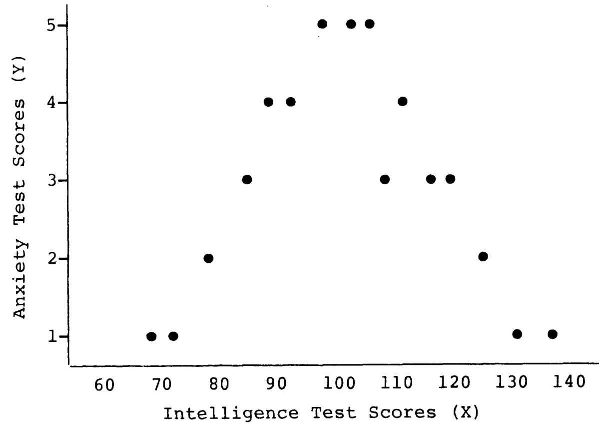

Figure 1. Scattergram of hypothetical intelligence and anxiety test scores for rho = –.043 (N = 16).

The conflict of information provided by the rank correlation coefficient and one-sample runs test gives rise to contentions of "spurious nonsignificance," and the scattergram suggests a resolution. Unfortunately, no test for curvilinear rank correlation is readily available, although an interesting possibility emerges. The data of Table 2 could be divided into two separate groups, with logical dichotomy between the second and third run (or the midpoint/apogee of Figure 1). Within each new group, data of X and Y could be ranked and conventional rank correlations computed. For the group comprised of Runs 1 and 2 (N = 10), rho would be +.767, and for that of Runs 3 and 4 (N = 6), rho would be –.914. Both coefficients are significant at the .05 level and offer interesting information about the true relationship between the two variables.

The procedure described to explore a nonsignificant rho (or tau) for relationships other than linear is both simple and direct and may hold promise for salvaging information potentially available in many cases of nonsignificant rank correlation. With a discordant finding of a rho or tau and the one-sample runs test, exploratory techniques seem warranted to determine the true relationship(s) between variables in question.

REFERENCES

Buckalew, L. W., & Pearson, W. H. Determination of critical observed frequencies in chi square. *Bulletin of the Psychonomic Society*, 1981, 18, 289-290.

Buckalew, L. W., & Pearson, W. H. Critical factors in the chi-square test of independence: A technique for exploratory data analysis. *Bulletin of the Psychonomic Society*, 1982, **19**, 225-226. (a)

Buckalew, L. W., & Pearson, W. H. Deviance probabilities: Determination of judgmental bias within Kendall's coefficient of concordance data. *Bulletin of the Psychonomic Society*, 1982, **20**, 187-189. (b)

Ferguson, G. A. *Statistical analysis in psychology and education* (5th ed.). New York: McGraw-Hill, 1980.

Siegel, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

(Received for publication October 7, 1982.)