jectionable to the senior age group. Delinquency must be defined in terms of "the expectation in the culture,"[8] and adults, of course, have a dominant role in producing cultural expectations. We do not claim that adults expect more of youths in racially heterogeneous lower socio-economic areas, but rather that they tolerate youthful deviance less, and, therefore, may report youths to law-enforcement officers more frequently than in racially homogeneous areas.

Evidence supporting the supposition, that low-income mixed neighborhoods tend to rely upon formal law-enforcement agents more frequently than the mixed neighborhoods of higher status, is obtained from an unpublished study by Washington Action for Youth which analyzes all youth contacted by police as well as those referred to Juvenile Court in fiscal year 1963. This study finds the highest rate of police contacts in the mixed lower socio-economic areas but the lowest ratio of Court-referrals there. This is largely because more minor offenses of youth are brought to the attention of police in the lower as compared with the higher status mixed areas in Washington.

In the racially mixed areas of higher socio-economic status, there is apparently less need to rely on formal external agents of social control to deal with deviance and apply local community sanctions. This may account for the similarity among delinquency rates in higher status neighborhoods of varying racial compositions. Our field observations and general information about the racially mixed higher socio-economic areas in Washington indicate that at this status level the variation in delinquency rates by racial composition is due not so much to socio-economic characteristics as to community organizational patterns. It is beyond the scope of this paper to analyze these patterns, but we suspect they involve the presence of associations that create a sense of community among neighbors.

[8] Wilkins, op. cit., p. 107.

LIMITATIONS IN THE APPLICATION OF NON-PARAMETRIC COEFFICIENTS OF CORRELATION *

Stanley Lieberson
University of Wisconsin

Kendall's *tau* and Spearman's *rho* are appropriate for certain non-linear functions, but for other curvilinear functions it is as inappropriate to use these non-parametric correlation coefficients as it would be to use Pearson's *r* without some transformation. Since these limitations of *tau* and *rho* are usually not mentioned in sociological statistics textbooks, this paper specifies the conditions of curvilinearity under which the use of *tau* or *rho* is appropriate and the conditions which do not meet the assumptions underlying these statistics. Procedures are presented for avoiding the erroneous inference of no association when applying non-parametric correlation coefficients to the latter class of curvilinear functions.

* Support from the University Research Committee of the University of Wisconsin is gratefully acknowledged.

Although these non-parametric coefficients can only measure the association of ranks, rather than the association of the underlying variables, curvilinear associations are not always lost when the variables are converted into an ordinal scale. Consider separately the two scaling conditions under which *tau* or *rho* might be used: first, where the data are available in interval or ratio scales; second, where the data are available only in ordinal form or where the values of the ratio or interval scales are not reliable. In the first case it is clear that conversion to ranks need not arbitrarily misrepresent a non-linear function. Consider, for example, a parabolic function of the simple form, $Y=a+bX+cX^2$. If the variables were converted to ranks, Y would still rise with increasing X up to a maximum point and then decline with further increases in X. Similarly, data based solely on ranks could represent a non-linear function. Suppose we wish to determine the influence of variations in talkativeness on variations in "attractiveness." We might well find a curvilinear function by inspection of ranks. For example, attractiveness might rise with talkativeness for the lower part of the X scale, but then fall off toward the very talkative end. Again our ranked data would graphically describe this non-linear function.

DIFFERENCES BETWEEN RANKING MONOTONIC AND NON-MONOTONIC CURVILINEAR FUNCTIONS

Since rank order correlations are sensitive to changes in a function's direction but not to changes in the slope within the same direction, all monotonic curvilinear functions are suitably measured by *tau* or *rho* and cannot be discriminated from one another on the basis of their rank scatter diagrams. For example, $Y=\frac{1}{a+bX}$ in rank form would appear identical to $Y=a+bX$, $\log Y=a+b \log X$, or any other monotonic function converted to ranks. Since all completely monotonic functions yield a *rho* or *tau* of ± 1.00, these coefficients may be viewed as measures of the extent to which a set of ranks can be described as a monotonic function.[1] The coefficients will also be ± 1.00 if the ranks are based on underlying variables that are monotonically related. In addition to these curvilinear monotones, some slightly non-monotonic functions can be described or measured reasonably well by utilizing *tau* or *rho* without further adjustment, although this procedure does not permit the investigator to describe the particular form of the function.

On the other hand, non-parametric coefficients of correlation will underestimate the existing degree of association if they are applied directly to a non-monotonic function. Because both rank measures are based on the assumption of a monotone, one or more major changes in the direction of a function will reduce the applicability of the non-parametric coefficients so that their maximum possible values will be less than ± 1.00.

If one wishes to test the null hypothesis that there is no *monotonic* association between two sets of ordinal numbers, then one may employ *tau* or *rho* directly with any function and a P greater than the specified α will constitute an adequate test. But if one's null hypothesis is that there is no association between two sets of ordinal numbers (whether the function be monotonic or otherwise), then P greater than α is not an adequate test because a high P may be due either to the validity of the null hypothesis or merely to the existence of a non-monotonic function.

In brief, rank-order correlation coefficients assume the existence of a monotone; they cannot be used to measure a non-monotone unless the investigator specifically means to reject higher-order associations. Otherwise, the investigator must normally guard against Type II errors when he obtains a low rank-order correlation: such a result could reflect either a lack of association between two variables or a non-monotonic association. Further, there is no justification for using *tau* or *rho* for non-monotonic functions based on interval or ratio data. That is, conversion to ranks will not make rank-order coefficients any more appropriate to such functions than an adjusted least squares measure would be. (*Tau* or *rho,* however, would be appropriate for interval data that form non-linear monotones.)

[1] Spearman's coefficient may also be described as "the slope of the best-fitting straight line relating two sets of ranks." See Virginia L. Senders, *Measurement and Statistics,* New York: Oxford University Press, 1958, p. 136.

TESTS FOR THE EXISTENCE OF A NON-MONOTONIC RANK ORDER FUNCTION

One fundamental procedure is to graph the ranks or underlying values and, by inspection, determine whether there are one or more major shifts in the direction of a function. Slight reversals in direction over a short span have little effect, and a statistically significant coefficient can still be obtained if the remainder of the curve fits a straight line. A slight reversal in one part when the remainder of the curve is a weak monotonic function may also be ignored, since the apparent change in direction may simply be due to random variations.

Investigators may prefer a formal test more rigorous than graphic inspection for avoiding Type II errors in applying *tau* or *rho* to a non-monotonic function. I am unaware of a suitable rank test, but I shall describe two characteristics that may be used to identify non-monotonic rank-order functions. Where the assumptions can be met, the correlation ratio in conjunction with r^2 provides a suitable test for ratio or interval scaled data; hence the steps described below are intended primarily for data available only in rank form or for data not meeting these assumptions.

First, place the X variable in natural order, i.e., 1, 2, 3, ..., N. Then, starting with the Y rank opposite X_1, mark each increase or decrease in the following Y rank over the preceding one with + or − respectively. In Illustration 1, X is in natural order and the + and − signs are placed appropriately. Since *tau* and *rho* are, respectively, .11 and .15 (in both cases $P > .05$), we would normally accept the null hypothesis of no association in the predicted direction.

Illustration 1

X	1	2	3	4	5	6	7	8	9	10
Y	1	3	5	7	9	10	8	6	4	2
		+	+	+	+	+	−	−	−	−

Let $n_1 =$ the number of plus signs, $n_2 =$ the number of minus signs, and $V =$ the number of runs of consecutive signs. If the association is non-monotonic, then we would expect V to be relatively small compared to the random distribution of V possible on the basis of the values of n_1 and n_2. Since the + and the − signs are based on successive differences in the Y's when ordered by the corresponding X's, the signs are not independent observations and are affected by a negative serial correlation. This prevents us from applying the usual runs test, which is based on the assumption of statistical independence between successive observations.[2] We can, however, determine the maximum and minimum values of V for a given set of $+$ and $-$ signs, and see whether the observed V is strongly on the low side. The minimum number of runs is $V=1$ only when *tau* or *rho* is $\pm$ 1.00. Otherwise the minimum number of runs would be $V=2$; this would occur when, as in Illustration 1, all of the plus signs are consecutive and all of the minus signs are consecutive. The maximum value of V

$$= 2n_1 + 1, \text{ if } n_1 < n_2, \text{ or}$$
$$= 2n_2 + 1, \text{ if } n_2 < n_1, \text{ or}$$
$$= 2n_1, \text{ if } n_1 = n_2$$

Thus, in Illustration 1, the maximum number of runs is $V=9$. Obtaining an extremely low V in the range possible for the given n_1 and n_2 indicates that there are relatively *persistent* consistencies of trend among the rankings. This can occur only when the association is relatively close, whether it be monotonic or non-monotonic. Since a monotonic function with such a low V would yield a significant *tau* or *rho,* a low V with a low correlation coefficient may be taken as evidence that the function is non-monotonic and hence *tau* and *rho* are not appropriate measures.

Another characteristic of non-monotones, especially useful since it is not shared by monotones, is also based on the runs approach. Suppose we have placed X in natural order and determine the number of runs in Y, as described above. Now, suppose we reverse this procedure, placing Y in natural order and determining the number of X runs. If and only if a strong non-monotonic function exists, the number of runs obtained will shift radically from what it was when X was placed in natural order. The number of runs when X is in natural order will be close to the minimum, and the number of runs when Y is in natural order will be close to the maximum V calculated on the basis of the number of $+$ and $-$ signs obtained in the latter instance.[3] For example, using the data from Illustration 1, we now place Y in natural order and determine the number of runs for X as shown in Illustration 2 below. We now obtain $V=9$, which is the maximum number possible since we again find $n_1=5$ and $n_2=4$.

Illustration 2

Y	1	2	3	4	5	6	7	8	9	10
X	1	10	2	9	3	8	4	7	5	6

$+ \quad - \quad + \quad - \quad + \quad - \quad + \quad - \quad +$

When there is either a strong monotonic function or a random association between the X and Y ranks, the number of runs obtained when X is placed in natural order will be similar to the number of runs obtained when Y is placed in natural order. Hence this procedure has an inherent advantage over the earlier approach, in that it permits us to determine the applicability of rank-order coefficients prior to their calculation.

TIES

There is no difficulty in performing the test when there are ties in Y, if the tied Y ranks are not contiguous after X is placed in natural order, but special procedures are required for contiguous ties in Y or any ties in the X ranks. When contiguous Y ranks are tied, no sign should be placed between the tied ranks. All other signs should be determined in the normal manner. When there are ties in the X ranks, let Y_a and $Y_{a'}$ refer to ranks opposite tied X ranks and let $Y_a > Y_{a'}$. Let Y_p refer to the Y rank preceding Y_a and $Y_{a'}$, and let Y_f refer to the Y rank following Y_a and $Y_{a'}$ when X is placed in natural order. If $Y_p > Y_a$, place a minus sign after Y_p. If $Y_p < Y_{a'}$ place a plus sign after Y_p. If $Y_a < Y_f$, place a plus sign before Y_f. If $Y_{a'} > Y_f$, place a minus sign before Y_f. No sign should be placed under each of the remaining conditions, i.e., $Y_{a'} < Y_p < Y_a$; $Y_{a'} < Y_f < Y_a$. And, again, no sign should be placed between Y_a and $Y_{a'}$. When the Y variable is placed in natural order, then the same procedure should be used, exchanging the X_i and Y_i.

[2] I am indebted to Leo A. Goodman for pointing this out.

[3] The number of $+$ and $-$ signs may differ when Y is in natural order from the number when X is in natural order.