A Generalized Interpretation of Pearson's r

Rudy A. Gideon
University of Montana
Missoula, MT 59812
email:gideon@selway.umt.edu

Rudy A. Gideon is a Professor , Department of Mathematical Sciences, University of Montana, Missoula, MT 59812

ABSTRACT

There are many interpretations of Pearson's correlation coefficient (see Rodgers and Nicewater (1988)) but maybe the most important one has been missed. The interpretation presented here makes use of the trigonometric identity: $\cos a = \cos^2 a/2 - \sin^2 a/2$ where a is the angle between two vectors in n space. Further, the difference on the right-hand side of the equation can be interpreted as distance from perfect negative correlation minus distance from perfect positive correlation. The so-called generalized form of correlation involving an inner product does not include all correlation coefficients, whereas the difference concept does. This new difference concept permits new correlation coefficients to be defined and encompasses the general framework of nonparametric correlation coefficients.

Key Words: absolute value methods, correlation, eigenvalues, estimating functions, median methods, simple linear regression.

1. **INTRODUCTION**

Gideon and Hollister (1987) introduced a nonparametric correlation coefficient that was based on the concept of greatest deviations (r_{gd}) and they gave an interpretation to the difference of the two functions in the numerator of this correlation. The numerator can be interpreted as "distance from perfect negative correlation minus the distance from perfect positive correlation", dpnc–dppc. In this definition the correlation does not have a linear restriction (dpnc+dppc=constant) in the numerator as the Pearson, Spearman, and Kendall correlation coefficients have. Spearman (1906) tried to develop a robust correlation coefficient based on absolute values, but was only partially successful and this correlation became known as Spearman's footrule in Kendall (1990). However, using the general framework of dpnc-dppc it is easy to define continuous and rank correlation coefficients based on medians and absolute values. This allows a continuous and rank correlation coefficients to be defined and be related in the same way Pearson and Spearman are. It also makes possible an Median Absolute Deviation (MAD) correlation coefficients to be defined so that if x equals y, then the correlation becomes the usual MAD variation estimator. At the end of the paper some general comments on correlation are made.

2. **A NEW INTERPRETATION OF PEARSON'S r**

Let the n dimensional data be the vectors x, y. Let SS_x, SS_y be the sum of squares about the means, $\bar{x}, \bar{y}$. Now define the standardized data as $z_1 = (x - \bar{x}\underline{1}) / (SS_x)^{1/2}, z_2 = (y - \bar{y}\underline{1}) / (SS_y)^{1/2}$, where $\underline{1}$ is a column vector of all 1's. It follows that $|z_1|^2 = z_1' z_1 = 1$, $|z_2|^2 = 1$, $z_1'\underline{1} = z_2'\underline{1} = 0$, and $(z_1 - z_2)'(z_1 + z_2) = 0$. If a is the angle in n dimensional space between z_1 and z_2, $\cos a = z_1' z_2$=r, Pearson's product moment correlation coefficient. In the Figure, z_1 and z_2 are vectors centered at zero and a is the angle between them. The angle a/2 in the Figure is $a / 2$.

We now want to show that

$$\cos a = \cos^2 \frac{a}{2} - \sin^2 \frac{a}{2} = \frac{\|z_1 + z_2\|^2}{4} - \frac{|z_1 - z_2\|^2}{4}$$

For this we need the following trigonometric identities:

$$\cos a = \cos^2 \frac{a}{2} - \sin^2 \frac{a}{2}, \quad \cos \frac{a}{2} = \sqrt{\frac{1 + \cos a}{2}}, \quad \cos(\frac{3p}{2} + \frac{a}{2}) = \sin \frac{a}{2} = \sqrt{\frac{1 - \cos a}{2}}.$$

Let b be the angle between vectors z_1 and $z_1 + z_2$, then

$$\cos b = \frac{z_1'(z_1 + z_2)}{\|z_1\| \, \|z_1 + z_2\|} = \sqrt{\frac{1 + z_1' z_2}{2}} = \sqrt{\frac{1 + \cos a}{2}} = \cos \frac{a}{2}$$

so that $b = \frac{a}{2}$.

Because $z_1 + z_2$ and $z_1 - z_2$ are orthogonal vectors, the angle in a counter-clockwise direction from z_1 to $z_1 - z_2$ is $(3p)/2 + a/2 = b$, for a moment. So we have

$$\cos b = \frac{z_1'(z_1 - z_2)}{\|z_1\| \, \|z_1 - z_2\|} = \sqrt{\frac{1 - z_1'z_2}{2}} = \sqrt{\frac{1 - \cos a}{2}} = \sin\frac{a}{2}.$$

Now $\frac{|z_1 + z_2\|^2}{4} = \frac{1 + z_1'z_2}{2} = \cos^2\frac{a}{2}$ and $\frac{|z_1 - z_2\|^2}{4} = \frac{1 - z_1'z_2}{2} = \sin^2\frac{a}{2}$.

Thus, it follows that Pearson's r can be written as

$$\text{r} = z_1'z_2 = \frac{|z_1 + z_2\|^2}{4} - \frac{|z_1 - z_2\|^2}{4} = \frac{1 + z_1'z_2}{2} - \frac{1 - z_1'z_2}{2} = \cos^2\frac{a}{2} - \sin^2\frac{a}{2}. \quad (1)$$

The distances from perfect positive and negative correlation are defined as

$$\text{dpnc} = \frac{|z_1 + z_2\|^2}{4} = \cos^2\frac{a}{2} \quad \text{and dppc} = \frac{|z_1 - z_2\|^2}{4} = \sin^2\frac{a}{2}.$$

For standardized data the covariance matrix is the correlation matrix and the eigenvalues of this 2 x 2 matrix are $l_1 = 1 + r, \, l_2 = 1 - r$. Thus, since

$$\cos^2\frac{a}{2} = \frac{1+r}{2} = \frac{l_1}{2} \text{ and } \sin^2\frac{a}{2} = \frac{1-r}{2} = \frac{l_2}{2},$$

another form or interpretation of this decomposition of r is that $r = (l_1 - l_2)/2$, and half the difference between the eigenvalues is the distance between dpnc and dppc.

Now for the interpretation of distance from perfect negative correlation (dpnc) minus the distance from perfect positive correlation (dppc), it is easily seen that dpnc is maximum when

$$z_1 = z_2 \text{ or } a = 0, \cos^2\frac{a}{2} = 1, \sin^2\frac{a}{2} = 0.$$

In this case dpnc= 1 and dppc = 0. Thus, r= dpnc – dppc =1–0=1. For the perfect negative correlation case, we have

$$z_2 = -z_1 \text{ and } \; a = p, \; \cos^2\frac{a}{2} = 0, \; \sin^2\frac{a}{2} = 1$$

so that r= dpnc – dppc = 0–1=–1.

We do two more cases. First, when z_1 and z_2 are orthogonal,

$$a = \frac{p}{2}, \sin\frac{a}{2} = \cos\frac{a}{2} = \frac{\sqrt{2}}{2},$$

so that dpnc=dnpc and r=0. Finally, let

$$a = \frac{3p}{4}, \cos^2\frac{3p}{8} = 0.1464 \text{ and } \sin^2\frac{3p}{8} = 0.8536.$$

In this case = dpnc – dppc = 0.1464 – 0.8536 = –0.7072.

Clearly, from the Figure it can be seen that dpnc $= \|z_1 + z_2\|/4$ is maximum and dppc $= \|z_1 - z_2\|/4$ is minimum when $z_1 = z_2$ and the reader can easily visualize the geometry of the examples above.

We now show that the nonparametric analogue of Pearson's correlation which is known as Spearman's correlation has the same interpretation. Let the x,y data be ordered and replaced by ranks (i, p_i) , $i = 1,2,\cdots,n$. Now use the ranks in the dpnc–dppc definition of r and get as the definition of Spearman's correlation coefficient

$$r_s = \frac{\sum (n+1- p_i - i)^2 - \sum (p_i - i)^2}{(n(n^2-1)/3)}. \qquad (2)$$

The second term $\sum (p_i - i)^2$ certainly measures a distance from perfect positive correlation and the first term is obtained from the second by replacing p_i by $n+1-p_i$, and it measures distance from perfect negative correlation; replace p_i by i and a maximum is obtained, replace p_i by n+1–i and a minimum is obtained.

An historical note and contrast will now be made. The above n-dimensional geometrical view of Pearson's r apparently allows a general geometrical way of viewing any correlation coefficient as will be seen in succeeding sections. Equation (1) can be viewed as the differences of scale measures, $\|z_1 + z_2\|^2 - \|z_1 - z_2\|^2$, and Gnanadesikan and Kettenring did this in their 1972 paper. They then went on to propose robust correlation coefficients by using robust measures of scale on $z_1 + z_2$ and $z_1 - z_2$. Huber (1981) gives a theoretical development of this in sections 8.2 and 8.3. The scale approach seems to have led to mainly the "subsets of data" approach in constructing robust correlation coefficients (see, Gnanadesikan (1977)).

The dnpc-dppc interpretation was first used for the Greatest Deviation correlation coefficient (GDCC,see equation (5). The n-dimensional geometrical view comes with this paper and connects all correlation coefficients with a common interpretation. GDCC currently has not been shown as the differences between two measures of scale. However, GDCC or any of the seven correlation coefficients in this paper (see equations 1 through 7) can be used to estimate scale, location, and parameters from many statistical problems, see section 5 for a brief introduction. GDCC is a very good robust estimator of scale and the method utilizes the concepts in Gnanadesikan (1977) to construct the estimates of parameters. For GDCC, location is more effectively estimated after the scale estimator which it uses in its construction. Thus, the development in this paper goes from robust estimates of correlation to robust estimates of scale, whereas, the earlier approach of Gnanadesikan and Kettenring(1972) went from robust estimates of scale to robust estimates of correlation.

To define a robust correlation coefficient both Gnanadesikan and Kettenring(1972) in equation (4) page 91 and Huber (1981) equation (2.6) page 203 utilized the sum of the robust measures of scale in the denominator to make the correlation coefficients lie between -1 and +1. Again this approach would be unsatisfactory for GDCC and many of the other correlation coefficients, because, for example, the maximum of GDCC is the greatest integer in the sample size divided by 2, and not the sum of the numerator parts. Note that to correctly define a correlation coefficient such as r_{mad} in equation (6) the standardization factors MAD_x and MAD_y cannot be factored out as is done with Pearson's r. In other words, other correlation coefficients need their own geometry, not the usual least squares type geometry.

3. **ABSOLUTE VALUE AND OTHER CORRELATIONS**

Rather than use squared distance in our correlation, since correlation is dpnc-dppc, we can use absolute distance. SS_x gets replaced by $SA_x = \sum |x_i - \bar{x}|$, and similarly for y. Then an absolute value correlation coefficient can be defined as

$$r_{av} = \frac{1}{2}\left(\sum \left| \frac{x_i - \bar{x}}{SA_x} + \frac{y_i - \bar{y}}{SA_y} \right| - \sum \left| \frac{x_i - \bar{x}}{SA_x} - \frac{y_i - \bar{y}}{SA_y} \right|\right). \quad (3)$$

It is very easy to show that this absolute value correlation coefficient is between -1 and +1. Let $x_i^* = \frac{x_i - \bar{x}}{SA_x}, y_i^* = \frac{y_i - \bar{y}}{SA_y}$ and apply the triangle inequality to $|x_i^* \pm y_i^*|$. The values -1 and +1 are achieved at $x_i = -y_i, i = 1,2,\cdots n$ and $x_i = y_i, i = 1,2,\cdots n$ respectively. Because r_{av} can assume the same values as Pearson's correlation coefficient, it should be true that a correlation matrix formed using r_{av} should also be positive definite. A covariance matrix could be formed by deleting the normalizing constants SA_x, SA_y in the definition of r_{av}.

To obtain the nonparametric analogue of this, do the same as for the squared distance measures, replace the data by their ranks. This will be the footrule correlation of Spearman made into a legitimate correlation and we will call it the modified footrule,

$$r_{mf} = \frac{\sum |n + 1 - p_i - i| - \sum |p_i - i|}{\left\lceil \frac{n^2}{2} \right\rfloor}. \quad (4)$$

This correlation coefficient was discovered by Gini (1914) and some of its properties are discussed in Betro (1993) and its population version is discussed in Scarsini (1984).

Because the median is the value of "a" that minimizes $\sum |x_i - a|$, r_{av} in equation (3) could be modified as follows to give yet another correlation coefficient.

$$r_{avm} = \frac{1}{2}\left(\sum \left| \frac{x_i - m_x}{SA_{m_x}} + \frac{y_i - m_y}{SA_{m_y}} \right| - \sum \left| \frac{x_i - m_x}{SA_{m_x}} - \frac{y_i - m_y}{SA_{m_y}} \right|\right) \quad (3^*)$$

where m_x, m_y are sample medians, and $SA_{m_x} = \sum |x_i - m_x|$, $SA_{m_y} = \sum |x_i - m_y|$. Unlike r_{av} this correlation, r_{avm}, has not been shown to be bounded between -1 and +1.

Because these ideas were motivated by the Greatest Deviation correlation coefficient, we list it to show it also looks like a dpnc-dppc correlation coefficient. Let $I(x) = 1$ if x is a true statement, otherwise 0. Then the Greatest Deviation correlation coefficient is

$$r_{gd} = \frac{\max_{1 \le i \le n} \sum_{j=1}^{i} I(n + 1 - p_j > i) - \max_{1 \le i \le n} \sum_{j=1}^{i} I(p_j > i)}{\left\lceil \frac{n}{2} \right\rfloor}. \quad (5)$$

In Kendall' s book on Rank Correlation Methods, Chapter Two is entitled, Introduction to the general theory of correlation methods. He defines an inner product type correlation coefficient that he calls a "generalized correlation coefficient" when it really should be called a generalized correlation coefficient of the inner product type. The last four correlation coefficients defined above do not fit into his definition. Also for the four, dpnc + dppc does not equal a constant. For Kendall's t, discordance relates to dppc and concordance to dpnc. At the end of Chapter Two Kendall notes that Spearman's footrule does not fit his general type, but he does not make it into a real correlation coefficient as was done here. See Gideon(1987, section 3) or Schweizer and Wolfe(1981) for what properties a "real" nonparametric corelation coefficient should have.This next section will develop some additional median methods for correlation coefficients using the general framework of dpnc and dppc.

4. MEDIAN ABSOLUTE DEVIATION CORRELATION COEFFICIENTS

In this section a correlation analog of the MAD, median absolute deviation estimate of variation, is given and denoted by r_{mad}. Let X,Y have a bivariate normal distribution with parameters, $m_x, m_y, s_x^2, s_y^2, r$. Now define for a random sample $MAD_x = med|x_i - med(x_i)|$ and similarly for the data from Y. A median-type correlation coefficient is defined as

$$r_{mad} = \frac{1}{2}\left(med\left|\frac{x_i - med(x_i)}{MAD_x} + \frac{y_i - med(y_i)}{MAD_y}\right| - med\left|\frac{x_i - med(x_i)}{MAD_x} - \frac{y_i - med(y_i)}{MAD_y}\right|\right). \quad (6)$$

It is evidently not true that $|r_{mad}| \le 1$, let $x_i^* = \frac{x_i - med(x)}{MAD_x}$ and similarly for y_i^*. Now $med|x_i^*| = med|y_i^*| = 1$. The proof that $|r_{mad}| \le 1$ breaks down because the median of the sum of two sets of nonegative numbers is not always less than the sum of the medians. It would be true if the following equation held r_{mad}

$$med|x_i^* + y_i^*| \le med(|x_i^*| + |y_i^*|) \le med|x_i^*| + med|y_i^*| = 2,$$

however, the last inequality does not hold. The computer language S or Splus has been used to examine r_{mad} and values greater slightly greater than one were occasionally obtained. Simulation studies of r_{mad} show it to behave very much like other correlation coefficients even with the anomaly of being greater than one. The spread of the distribution is very close to other correlations, and only when the population correlation is near one, does r_{mad} become only slightly greater than one on rare occasions.

4.1 **Population Value of** r_{mad}

In order to find the population value of r_{mad} the following results are needed for random variables. Let $Z_1 = (X - m_x)/s_x, Z_2 = (Y - m_y)/s_y$. Since $P(|Z_1| \le 0.6745) = 0.5000$, the median absolute deviation for random variable Z_1 is $MAD_{Z_1} = med|Z_1| = 0.6745$. Also, $V(Z_1 + Z_2) = 2(1 + r)$ and $V(Z_1 - Z_2) = 2(1 - r)$, and

$MAD_{Z_1} = med|Z_1| = med|(X - m_x)/s_x| = 0.6745$ so that

$MAD_X = med|X - m_x| = 0.6745 s_x$.

In a similar fashion

$med|Z_1 + Z_2| = 0.6745 s_{z_1+z_2} = 0.6745(2(1 + r))^{1/2}$ and

$med|Z_1 - Z_2| = 0.6745 s_{z_1-z_2} = 0.6745(2(1 - r))^{1/2}$.

We are now in a position to obtain the population value of r_{mad}. The median of the X(Y) sample is converging in probabiliy to m_x (m_y) and MAD_x is converging to $0.6745 s_x$. So for the population value we make changes from the sample to the population and obtain

$$r_{mad} = \frac{1}{2}\left(med\left|\frac{X-m_x}{0.6745 s_x} + \frac{Y-m_y}{0.6745 s_y}\right| - med\left|\frac{X-m_x}{0.6745 s_x} - \frac{Y-m_y}{0.6745 s_y}\right|\right)$$

$$= \frac{1}{2}\left(med\frac{|Z_1+Z_2|}{0.6745} - med\frac{|Z_1-Z_2|}{0.6745}\right) = \frac{1}{2}\left(\sqrt{2(1+r}\, - \sqrt{2(1-r}\right).$$

Finally,

$$r_{mad} = \sqrt{\frac{1+r}{2}} - \sqrt{\frac{1-r}{2}}\,.$$

4.2 **A Direct Estimate of the Covariance Function of the Bivariate Normal Distribution Using the Median**

Let $t^+ = \{x_i - med(x_i) + y_i - med(y_i), i = 1,2,\cdots,n\}$ and similarly $t^- = \{x_i - med(x_i) - (y_i - med(y_i)), i = 1,2,\cdots,n\}$, the sets of sample values. The random variable equivalents of these sample quantities are defined to be $T^- = X - med(X) - (Y - med(Y))$ and $T^+ = X - med(X) + (Y - med(Y))$. It is clear that $med|t^+| \xrightarrow{P} med|T^+|$ and $med|t^-| \xrightarrow{P} med|T^-|$. It is now easy to show that $(med^2|t^+| - med^2|t^-|)/(4(0.6745)^2) \xrightarrow{P} (med^2|T^+| - med^2|T^-|)/(4(0.6745)^2) = \mathrm{cov}(X,Y)$. Thus, $(med^2|t^+| - med^2|t^-|)/(4(0.6745)^2)$ is a robust estimate of the covariance function. It follows that a direct robust estimate of r would be

$$\hat{r} = \frac{(med^2|t^+| - med^2|t^-|)}{4\,MAD_x MAD_y}. \qquad (7)$$

If $x_i = y_i, i = 1,2,\cdots,n$ the estimate of the covariance function becomes the estimate of the variance; i.e.,

$$\hat{s}^2 = (med^2|t^+| - med^2|t^-|)/(4(0.6745)^2) = med^2|2(x_i - med(x_i)|/(4(0.6745)^2)$$
$$= MAD_x^2/(0.6745)^2).$$

Very little developmental work has been done on the ideas in this section. The language S was used to show that for the estimate of the covariance function the result $\mathrm{cov}(ax,by) = ab\,\mathrm{cov}(x,y)$ does not hold, but it is approximately true. In the limit this result does hold. For this result, let $T^+(a,b) = T^+$ applied to aX and bY instead of X and Y. Now $med|T^+| = med|X - m_x + Y - m_y| \xrightarrow{P} 0.6745\sqrt{s_x^2 + s_y^2 + 2\,\mathrm{cov}(X,Y)}$. A similar result hold for $med|T^-|$ but with a minus sign by the covariance. Then

$$\frac{med^2|T^+(a,b)| - med^2|T^-(a,b)|}{4(0.6745)^2} \xrightarrow{P} \mathrm{cov}(aX,bY) = ab\,\mathrm{cov}(X,Y).$$

References to MAD can be found in the Rousseeuw and Croux (1993) article, although there are no generalizations to correlations.

5. GENERALIZATIONS AND ESTIMATING FUNCTIONS

Most people who use statistics seem to think of correlation coefficients as descriptive statistics. The only point of this section is to refute that belief. First, consider simple linear regression; $y = a1 + bx + e1$, where 1 is a vector of 1's. Let r be any correlation coefficient, and define the following function of b: $f(b) = r(x, y - xb)$. Under the usual simple linear regression assumptions the values of $f(b)$ for random samples will produce a random variable with a null distribution, symmetric about zero. For a particular random sample the solution b to f(b)=0 will give an estimate of b. The monotonicity (sometimes as a step function) of f(b) as a function of b will give a Hodges-Lehmann type confidence interval for b.

The review article by Liang and Zeger (1995) in Section 2.1 gives a review of "estimating functions". They referred to Durbin's (1960) study of the single parameter time series model $y_t = qy_{t-1} + e_t$ as one of the starting points for "estimating functions". Here the e_t are taken as i.i.d. with zero means and variance s^2,and the unbiased linear estimating equation is $g(y, \hat{q}) = \sum_{t=1}^{k} y_{t-1} y_t - \hat{q} \sum_{t=1}^{k} y_{t-1}^2 = 0$ with $\hat{q}$ being the estimate of q. Now if we center the data , say $\sum_{t=1}^{k} y_{t-1} = 0$ and let $x = y_{t-1}, y = y_t$ (appropriately lagged vectors of time series data) then the estimating equation is $r(y_{t-1}, y_t - \hat{q} y_{t-1}) = 0$, where r is Pearson's correlation coefficient. Thus, in Durbin's estimating equation, the function g is really Pearson's correlation coefficient.

Thus, any correlation coefficient can be considered an estimating function, with some being nonlinear (e.g. r_{gd}). The objective function for any correlation coefficient could be considered the equalization of the distance from perfect positive and negative correlation; i.e., dpnc = dppc. Let us alter Durbin's definition 1 of an estimating equation for the simple time series model.

Altered Definition 1: Suppose that an estimator $\hat{q}$ of a parameter q is given by $r(y_{t-1}, y_t - \hat{q} y_{t-1}) = 0$ where r is any legitimate correlation coefficient such that $r(y_{t-1}, y_t - q y_{t-1}) = r(y_{t-1}, e_t)$ has a distribution that is symmetric and centered at zero. Then $r(y_{t-1}, y_t - \hat{q} y_{t-1}) = 0$ is called a correlation estimating equation.

If Pearson's r is substituted in the above equation one gets an unbiased estimating equation, but if either r_{gd} or r_{av} are substituted one does not get a linear equation.

The correlation coefficient r_{gd} has been examined in a number of situations as an estimating function for location, scale, linear, nonlinear, and generalized linear models. It works very well. The modified footrule correlation coefficient r_{mf} has been studied only as a correlation coefficient , but it has the potential to be a valuable objective function or estimating equation. The median correlations should prove useful in a number of situations where robust methods are necessary. These methods give rise to a number of interesting questions. For example, in simple linear regression minimizing the absolute value of the errors does not give the

same results(the estimate of slope and intercept) as does equalizing the distances from perfect postive and negative correlation, whereas in least squares procedures these two methods can give the same estimates.

It should be pointed out that the idea of using a correlation coefficient to do simple linear regression is not new. Sen (1968) examined simple linear regression using Kendall's Tau. However, he did not use the general notation of this section which allows the posing of the equations with any correlation coefficient for use in multiple and generalized regression situations. One also needs a tied value procedure that works in these complex estimation problems. The idea of using the average of the maximum and minimum for tied values to determine the value of a correlation coefficient based on ranks is illustrated in Gideon (1987), and has proven effective in using the Greatest Deviation correlation coefficient as an estimating tool, and hence, should work with other correlation coefficients.

For the many users of statistics who do not have deep and broad training, this notation and the use of correlation coefficients other than the nonrobust Pearson's r (which is equivalent to least squares) should be appealing, because it does not rely on a educated selective choice of the data.

References

Betro, B.(1993), "On the Distribution of Gini's Rank Correlation Association Coefficient", Communications in Statistics: Simulation and Computation, M. Dekker, 22,No. 2, 497-505.

Durbin,J. (1960), "Estimation of parameters in time-series regression models," Journal of the Royal Statistical Society, Series B, 22 , 139-153.

Gideon, R.A. and Hollister, R.A. (1987), "A Rank Correlation Coefficient Resistant to Outliers," Journal of the American Statistical Association 82,no.398, 656-666.

Gideon, R.A. and Li, H. (1996), "A Correlation Coefficient Based on Spearman's Measure of Disarray," submitted to Applied Statistics (Series C, Royal Statistical Society).

Gini, C. (1914), "L'Ammontare c la Composizione della Ricchezza della Nazioni", Bocca, Torino.

Gnanadesikan, R. (1977), *Methods for Statistical Data Analysis of Multivariate Observations,* John Wiley and Sons.

Gnanadesikan, R. and Kettenring, J.R. (1972), "Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data, Biometrics, 28, 81-124.

Huber, P.J. (1981), *Robust Statistics* , John Wiley and Sons.

Kendall, M.G. and Gibbons, J.D. (1990), *Rank Correlation Methods,* 5th ed. Oxford University Press, or also Kendall, M.G. (1962), *Rank Correlation Methods* , 3rd ed. Hafner Publ. Co.

Liang, K.-Y. and Zeger, S.L. (1995), "Inference Based on Estimating Functions in the Presence of Nuisance Parameters," Statistical Science 10, no. 2, 158-172.

Rodgers, J.L. and Nicewater, W.A. (1988), "Thirteen Ways to Look at the Correlation Coefficient," The American Statistician, 42, no. 1, 59-66.

Rousseeuw, P.J. and Croux, C. (1993), "Alternatives to the Median Absolute Deviation," Journal of the American Statistical Association , 88, 1273-1283.

Scarsini, M. (1984), "On Measures of Condordance", Stochastica, 8, No.3, 201-218.

Schweizer, B. and Wolfe, E.F. (1981), "On Nonparametric Measures of Dependence for Random Variables," The Annals of Statistics , 9, 879-885.

Sen, P.K. (1968), "Estimates of the Regression Coefficient Based on Kendall's Tau," Journal of the American Statistical Association, 63, 1379-1389.

Spearman, C. (1906), " 'Footrule' for Measuring Correlations," British Journal of Psychology, 2, 89-108.

figure title

Pearson's r = $\cos^2(a/2) - \sin^2(a/2)$