Requirements

1. The paired data are a simple random sample.
2. The data are ranks or can be converted to ranks.

Note: Unlike the parametric methods of Section 10-2, there is *no* requirement that the sample pairs of data have a bivariate normal distribution (as described in Section 10-2). There is *no* requirement of a normal distribution for any population.

Test Statistic

Within each sample, first convert the data to *ranks*, then find the exact value of the rank correlation coefficient r_s by using Formula 10-1:

$$r_s = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

No ties: After converting the data in each sample to ranks, if there are no ties among ranks for the first variable and there are no ties among ranks for the second variable, the exact value of the test statistic can be calculated using Formula 10-1 (at the left) or with this relatively simple formula:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$

P-Values

P-values are sometimes provided by technology. (*Caution:* When finding r_s, you can convert the sample data to ranks and then use the same technology that was used for the linear correlation coefficient described in Section 10.2, but *do not use P-values from linear correlation for the methods of rank correlation.* Use the *P*-value from technology only if the technology has a procedure designed specifically for rank correlation. See the "Using Technology" instructions given at the end of this section.)

Critical Values

1. If $n \le 30$, critical values are found in Table A-9.
2. If $n > 30$, critical values of r_s are found using Formula 13-1.

Formula 13-1

$$r_s = \frac{\pm z}{\sqrt{n - 1}} \quad \text{(critical values when } n > 30\text{)}$$

where the value of z corresponds to the significance level. (For example, if $\alpha = 0.05$, $z = 1.96$.)

CAUTION When working with data having ties among ranks, the rank correlation coefficient r_s can be calculated using Formula 10-1. Technology can be used instead of manual calculations with Formula 10-1, but the displayed *P*-values for linear correlation do not apply to the methods of rank correlation. *Do not use P-values from linear correlation for methods of rank correlation.*

Advantages: Rank correlation has these advantages over the parametric methods discussed in Chapter 10:

1. Rank correlation can be used with paired data that are ranks or can be converted to ranks. Unlike the parametric methods of Chapter 10, the method of rank correlation does *not* require a normal distribution for any population.
2. Rank correlation can be used to detect some (not all) relationships that are not linear.