# Confidence intervals around Pearson r's

**Re: Loftus, G.R. & Loftus, E.F. (1988). Essence of Statistics (2nd Edition). New York: McGraw Hill.**

## Mea Culpa

The explanation of confidence intervals in the book (p. 460) is wrong. It turns out to be more complicated. Here's how it works.

## General

First, unlike confidence intervals around means, confidence intervals around Pearson r's are not symmetrical. This is because the distribution of r is itself skewed rather than symmetrical (for example, with a high r, say r = 0.95, the actual population correlation couldn't be any higher than 1.0, but it could be substantially lower).

## Correct Formulas

The confidence interval around a Pearson r is based on Fisher's r-to-z transformation. In particular, suppose a sample of n X-Y pairs produces some value of Pearson r. Given the transformation,

$$z = 0.5\ln\left(\frac{1+r}{1-r}\right) \quad \text{(Equation 1)}$$

z is approximately normally distributed, with an expectation equal to

$$0.5\ln\left(\frac{1+\rho}{1-\rho}\right)$$

where r is the population correlation of which r is an estimate, and a standard deviation of

$$S = \sqrt{1/(n-3)}$$

Therefore, having computed an obtained z from the obtained r via Equation 1, a confidence interval can easily be constructed in z-space in more or less the usual manner as:

$$z \pm S \times (\text{criterion } z)$$

where the criterion z corresponds to the desired confidence level (e.g., 1.96 in the case of a 95% confidence interval). The upper and lower z limits of this confidence interval can then be transformed back to upper and lower r limits.

**An Example**

Suppose that a sample of n = 20 X-Y pairs produces a Pearson r of 0.80, and a 95% confidence interval is desired. The obtained z is thus

$$0.5 \text{ x } \ln [(1+.80)/(1-.80)] = 0.5 \text{ x } \ln(1.80/.20) = 1.099$$

which is distributed with a standard deviation of

$$\sqrt{1/(20-3)} = 0.243$$

The upper and lower confidence interval limits in z-space are therefore

$$1.099+(.243)(1.96) = 1.574$$

and

$$1.099-(.213)(1.96) = 0.624.$$

To translate from z-space back to r-space, it is necessary to invert Equation 1, It is easily shown that such inversion produces,

$$r = \frac{e^{2z} - 1}{e^{2z} + 1} \quad \text{(Equation 2)}$$

The upper and lower confidence-interval limits may then be computed from Equation 2:

$$\textit{upper} \;\; \text{limit}: \; r = \frac{e^{2 \text{ x } 1.574} - 1}{e^{2 \text{ x } 1.574} + 1} = 0.918$$

and

$$\text{lower} \;\; \text{limit}: \; r = \frac{e^{2 \text{ x } 0.624} - 1}{e^{2 \text{ x } 0.624} + 1} = 0.554$$

Thus, the 95% confidence interval around the original obtained r of 0.90 ranges from 0.554 to 0.918.

### Picture

The situation described in the example above is depicted in the figure below.