How to use the Correlation and **Model Error** Spreadsheet

This Spreadsheet demonstrates how, starting with a specific input , a linear regression model with standardized variables can be used to make two different types of forecasts or estimates,

A *Point* Forecast: , and

A *Probabilistic* Forecast: .

Note that ) can also be written as , notation for “the variance of the model error.” So an equivalent way to represent the probabilistic forecast is .

Since the linear function for standardized variables can also be expressed , yet a third equivalent way to represent the probabilistic forecast is .

In other words, the probabilistic forecast takes the form of a Gaussian probability distribution function with mean = , variance = , and standard deviation = . The advantage of the probabilistic forecast is that it allows one to specify the probability that the true result will fall within a certain range around the mean – a “confidence interval.”

**Example 1.** Assume a standardized linear regression model has correlation R = .5 [Cell I38] and the point = .8 [Cell I36]. The point forecast = = .4 [Cell I40]. This value is also the mean of the Gaussian probability distribution. The standard deviation of the Gaussian is = .87 [Cell H46].

Assume one wants 50% confidence [Cell F46] that the true answer would fall within a defined range. Calculated in Excel, the range to specify would over the interval from

– (NormSInv(.75) ) to NormSInv(0.5+(%)/2)

+ (NormSInv(.75)])

This range, from -.18 [Cell I46] to .98 [Cell K46] is known as the “50% Confidence interval.”

**Example 2.** Assume the correlation is R = .353 [Cell I 38] and the value of = .09 [Cell I36].

Q: What is the mean of the probabilistic forecast?

Answer: .0315 [Cell I40].

Q: What is the 99% Confidence Interval? [Cell F50].

Answer: from -2.38 [Cell I50] to 2.44 [Cell K50].