Comparison of Distributions

Let's compare the various distribution functions that we've reviewed in this lesson and consider when they're best used.

Recall that the Poisson regression is useful for modeling the count or rate of rare events, whose variance increases in proportion to the mean. When the expected value increases, the Poisson distribution approaches the normal distribution. Both the gamma and lognormal distributions are appropriate for continuous positive values whose variance increases in proportion to the square of the mean. Their skewness is independent of the expected value of the dependent variable. Therefore, for skewed distributions with relatively large means, the gamma or lognormal distribution might be better choices than the Poisson distribution.

An additional consideration in selecting the correct distribution is the tail behavior. Although all of the distributions in this plot have the same expected value and variance, they have increasingly heavy tails. A few extreme outliers might indicate a lognormal distribution, whereas the absence of outliers might imply a gamma or another less extreme distribution. [Highlight the 100X view.] For many monetary-related models, the residual variance increases with the predicted value. The commonly used distributions for this type of data are Poisson, gamma, and lognormal distributions.

You can fit the gamma regression model using the GLIMMIX procedure. Remember that in Lesson 3, you learned to use PROC GLIMMIX to fit the lognormal regression model.
