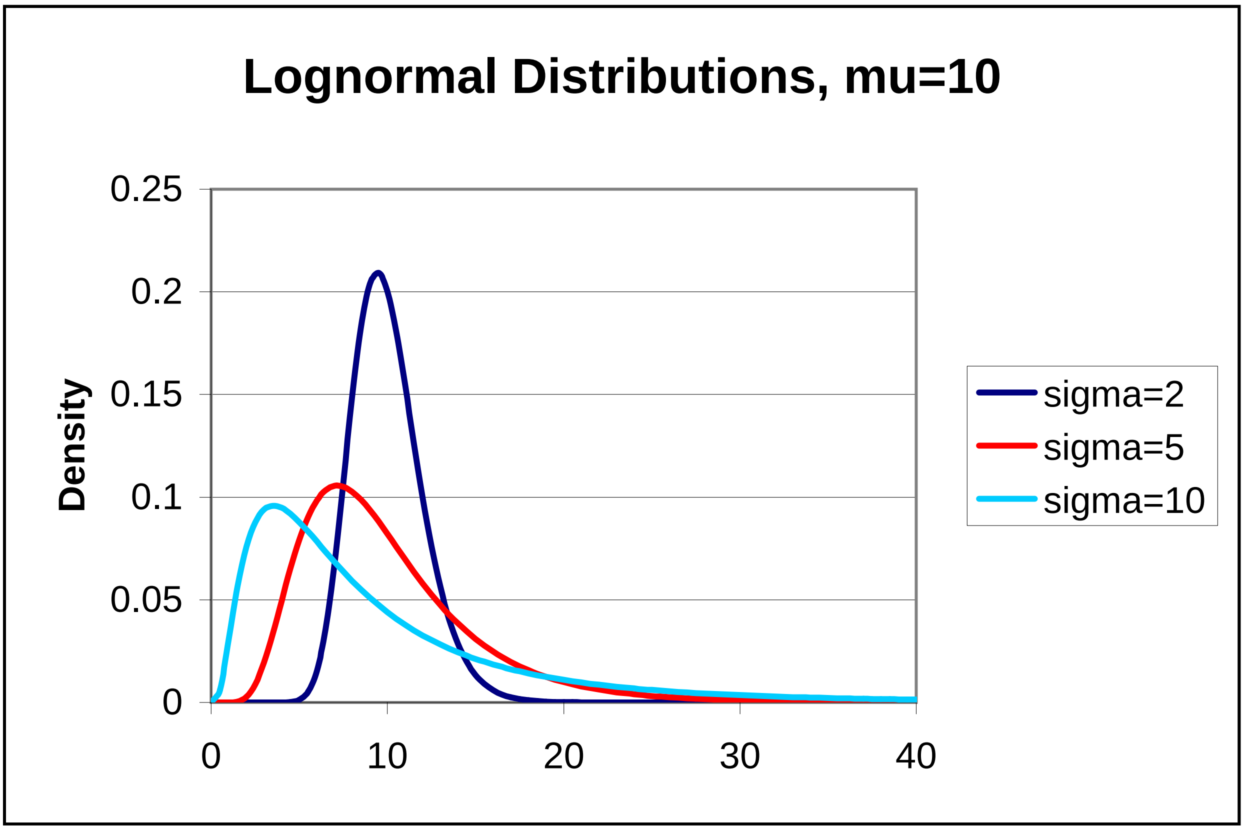
# The Lognormal Distribution

Kurt M. Anstreicher

The *lognormal* distribution is a commonly used probability distribution in finance and engineering. A random variable $Y$ is said to be lognormally distributed if the logarithm of $Y$ has a normal distribution. In other words, if a new random variable $X$ is defined by $X = \ln(Y)$, then $Y$ is lognormally distributed exactly when $X$ has a normal distribution. Here ln denotes the natural logarithm. Equivalently, one can think of starting with a normally distributed random variable $X$, and then defining $Y$ by exponentiating $X$; $Y = \exp(X) = e^X$, where $e$ is about 2.7182818. The distribution of $Y$ is approximately normal if $\sigma_Y$ is small relative to $\mu_Y$, but becomes more and more skewed as $\sigma_Y$ increases. The chart below shows the densities of three lognormal distributions, with $\mu_Y = 10$ and $\sigma_Y = 2, 5, 10$.




Let $Y$ be a lognormal random variable, and let $X$ be the normal random variable $X = \ln(Y)$. Let $\mu_X$, $\sigma_X$ denote the mean and standard deviation of $X$. The quantities $\mu_X$, $\sigma_X$, $\mu_Y$ and $\sigma_Y$ are related by relatively simple formulas which are very useful in practice. In one typical application one is given $\mu_Y$ and $\sigma_Y$, and it is desired to generate observations of the random variable $Y$. Typically this is done by generating observations $X_1, X_2, \ldots, X_n$ of the normal random variable $X$, and then letting $Y_i = e^{X_i}$. To generate the observations of $X$ one obviously needs $\mu_X$ and $\sigma_X$. These are given in terms of $\mu_Y$ and $\sigma_Y$ by the formulas:

$$\begin{aligned} \mu_X &= 2\ln(\mu_Y) - \frac{1}{2}\ln(\mu_Y^2 + \sigma_Y^2), \\ \sigma_X^2 &= \ln\left(\frac{\sigma_Y^2}{\mu_Y^2} + 1\right). \end{aligned}$$

For example, suppose that you want to generate observations of a lognormal random variable $Y$ having mean 10, and standard deviation 5. Using the formulas, one gets

$$\begin{aligned} \mu_X &= 2\ln(10) - .5\ln(125) = 2.1910133, \\ \sigma_X^2 &= \ln\left((25/100) + 1\right) = 0.2231436, \end{aligned}$$

so $\mu_X = 2.1910133$, $\sigma_X = \sqrt{0.2231436} = 0.4723808$ . (Because of the use of exponentiation to generate $Y$ from $X$, the distribution of $Y$ is very sensitive to the values of $\mu_X$ and $\sigma_X$. As a result, it is a good idea to keep lots of digits in the computed values of $\mu_X$ and $\sigma_X$.)

In other applications one is given $\mu_X$ and $\sigma_X$, and needs $\mu_Y$ and $\sigma_Y$. This situation arises, for example, when making forecasts using a regression where the dependent variable is "logged." The formulas giving $\mu_Y$ and $\sigma_Y$ in terms of $\mu_X$ and $\sigma_X$ are:

$$\begin{aligned} \mu_Y &= e^{(\mu_X + \sigma_X^2/2)}, \\ \sigma_Y^2 &= e^{(2\mu_X + \sigma_X^2)}[e^{\sigma_X^2} - 1]. \end{aligned}$$

For example, suppose that $\mu_X = 2.1910133$, $\sigma_X = 0.4723808$. Using the formulas one gets

$$\begin{aligned} \mu_Y &= e^{(2.1910137 + .1115718)} = 10, \\ \sigma_Y^2 &= e^{(4.3820274 + .2231436)}[e^{.2231436} - 1] = 25, \end{aligned}$$

so $\mu_Y = 10$, $\sigma_Y = \sqrt{25} = 5$, as expected.