Distribution of Normal-Lognormal Ratios

Matthew Pollard

August 31, 2008

Abstract

The ratio distribution X/Y where X is normally distributed and Y is log-normally distributed is considered. I show there is no closed form expression for the ratio distribution, show that the distribution is shape invariant to μ_y and σ_x, and use Monte Carlo draws to observe the distribution under a set of parameters for μ_x and σ_y. I show that distribution is well approximated by a Student-t distribution when μ_x equals zero.

1 Analytic Derivation

Let X be drawn from the normal distribution and Y be drawn from the log-normal distribution. Suppose X and Y are independent, and the means and variances are given by

$$\begin{aligned} X &\sim \mathcal{N}(\mu_x, \sigma_x^2) \\ Y &\sim \mathcal{LN}(\mu_y, \sigma_y^2) \end{aligned}$$

The joint distribution $f(x, y)$ is

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\frac{1}{y}\exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(\log y - \mu_y)^2}{2\sigma_y^2}\right), \quad y > 0$$

The distribution $U = X/Y$ is has density function $p(u)$ defined by

$$\begin{aligned} p(u) &= \int_{-\infty}^{\infty} |y| f(uy, y) dy \\ &= \int_0^{\infty} y f(uy, y) dy \\ &= \frac{1}{2\pi\sigma_x\sigma_y}\int_0^{\infty} \exp\left(-\frac{(uy-\mu_x)^2}{2\sigma_x^2} - \frac{(\log y - \mu_y)^2}{2\sigma_y^2}\right) dy. \end{aligned}$$

This integral does not specify a common fat-tailed distributions, such as the Cauchy or Student-t. There is no closed form expression for this integral. To show this, I consider the case where $\mu_x = \mu_y = 0$, $\sigma_x = \sigma_y = 1$. Here,

$$p(u) = \frac{1}{2\pi}\int_0^{\infty} \exp\left(-\frac{u^2y^2 + (\log y)^2}{2}\right) dy$$

Claim: This integral has no closed form expression.

Proof: Expanding the exponential yields

$$\int_0^\infty \exp\left(-\frac{u^2y^2+(\log y)^2}{2}\right)dy = \underbrace{\int_0^\infty \exp\left(-\frac{u^2y^2}{2}\right)}_{:=p'}\underbrace{\exp\left(-\frac{(\log y)^2}{2}\right)}_{:=q}dy$$

The first term can be integrated: $\int_0^\infty \exp\left(-\frac{u^2y^2}{2}\right)dy = \frac{\sqrt{\pi/2}}{u}$. The second term has no closed integral, $\int_0^\infty \exp\left(-\frac{(\log y)^2}{2}\right)dy$, as can be confirmed in Mathematica. There is also no closed form solution to

$$\int_0^\infty pq'dx = \sqrt{\frac{\pi}{2}}\int_0^\infty \mathrm{erf}\left(\frac{uy}{\sqrt{2}}\right)\frac{1}{u}\frac{-2\log(y)}{y}\exp\left(-\frac{(\log y)^2}{2}\right)dy$$

which prevents an application of differentiation by parts.

2 Shape Invariance

$p(u)$ is shape invariant to μ_y and σ_x. Two distributions are shape invariant if they differ only up to scale and position, or $U_1(\mu_y,\sigma_x) \sim a + bU_2(\mu_y^*,\sigma_x^*)$.

Proof: For any $\mu_y \neq 0$, X/Y can be rewritten as

$$\frac{X}{Y} = \frac{1}{\exp(\mu_y+\sigma_y^2)}\frac{X}{Y^*},\ \ Y^* \sim \mathcal{LN}(0,\sigma_y^2)$$

For any $\sigma_x \neq 1$, $\frac{X}{Y}$ can be rewritten as

$$\frac{X}{Y} = \frac{\sigma X^*}{Y},\ \ X^* \sim \mathcal{N}(\mu_x,1)$$

The shape of $\frac{X}{Y}$ depends only on $\mu_x$and σ_y. Intuitively, if $\mu_x > 0$, the distribution will be positively skewed. If σ_y is large, the distribution will have high kurtosis.

3 Sledge-Hammer Approach

The distribution can be observed by Monte Carlo simulation. Figure 1 shows 10^6 Monte Carlo draws from the ratio distribution with several values for μ_x and σ_y. The plots show that larger σ_y induces greater kurtosis and $\mu_x > 0$ induces positive skewness.

The distribution is well approximated by the Student-t distribution when $\mu_x = 0$, and where σ_y is inversely proportional to the Student-t degrees of freedom. To show this, 10^6 samples were drawn for values $\sigma_y = (0.25, 0.5, 1, 1.5, 2)$, rescaled to have unit variance, and fitted against the Student-t distribution with k degrees of freedom. $k(\sigma_y)$ is estimated by maximizing the log likelihood using a Newton iteration. Goodness-of-fit tests are carried out by plotting the (log) empirical density to the estimated Student-t distribution.

Table 1 displays the fitted degrees of freedom against σ_y. As σ_y rises, the estimates appears to asymptote to $k = 1$, which corresponds to a Cauchy distribution. Figure 2 displays goodness-of-fit figures.

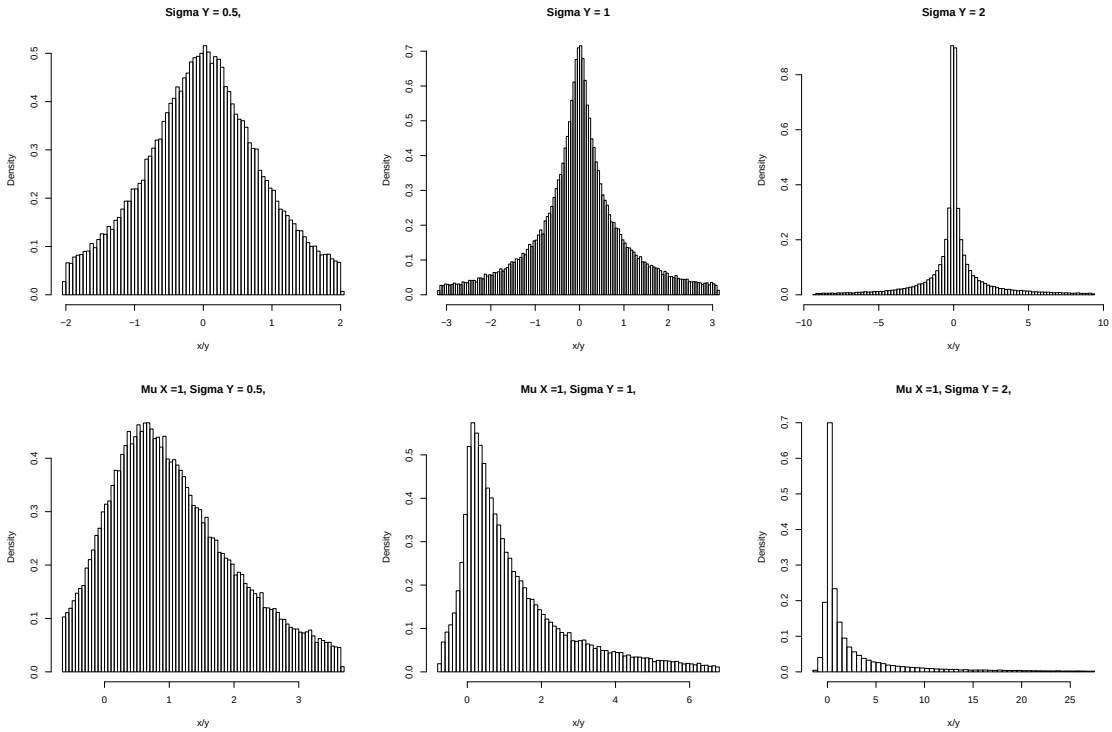

Figure 1: Distribution of X/Y with varying σ_y and μ_x.

σ_y	0.25	0.5	1	1.5	1.75	2
$\hat{k}(\sigma_y)$	15.0	4.58	1.82	1.41	1.008	1.003

Table 1: Estimated Student-t degrees of freedom.

The two distributions match one-another very well. For values of $\sigma_y > 1.5$, the ratio distribution and Student-t comparison approaches Cauchy-like behaviour, and the empirical density estimates are affected by extreme outliers. Nonetheless, the goodness-of-fit figures suggest that the ratio distribution is very well approximated by the Student-t.

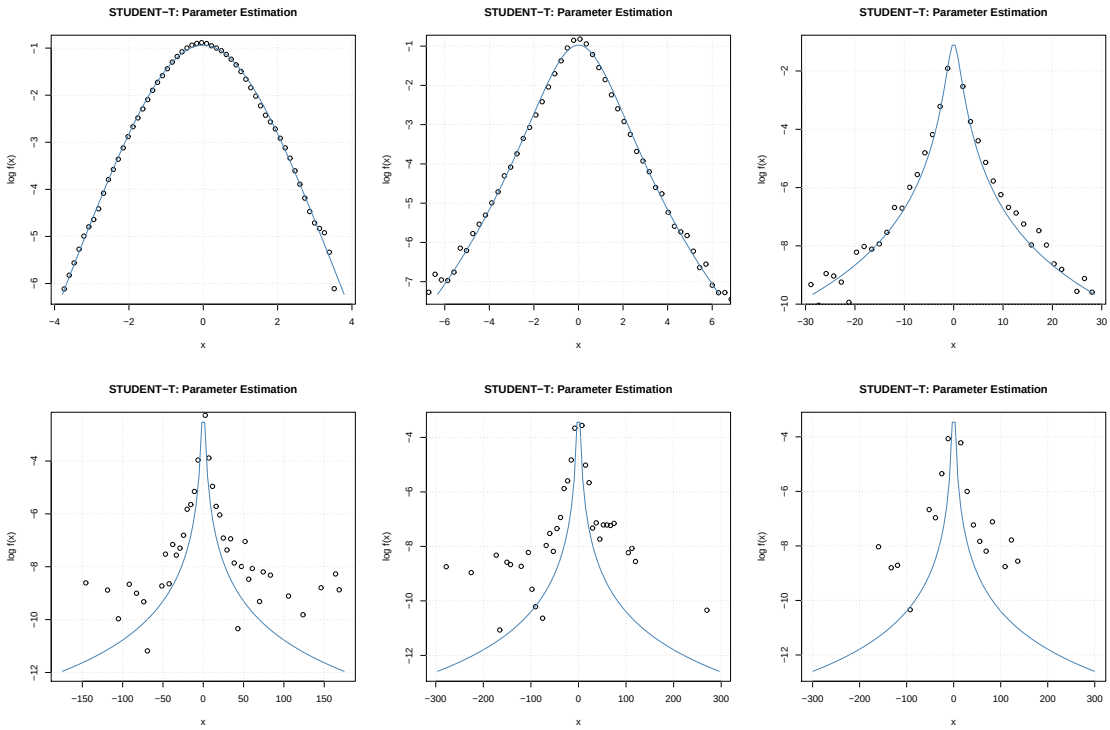

Figure 2: Goodness-of-fit against Student-t. Logarithm of empirical distribution plotted against fitted Student-t (blue line).