The Lognormal Central Limit Theorem for Positive Random Variables

by

Lilit Mazmanyan

Victor Ohanyan

and

Dan Trietsch

November 2008

Abstract

Practitioners often use the central limit theorem as justification for invoking the normal approximation for the convolution of few independent random variables. We focus on the convolution of independent nonnegative continuous random variables and advocate the use of the lognormal approximation instead of the normal. Among popular distributions, the lognormal distribution is unique in the sense that it satisfies elementary conditions that the convolution of a small number of continuous nonnegative random variables must satisfy and yet it converges to the normal when the number of random variables grows large. Therefore, one can use it as the basis of an alternative central limit theorem for nonnegative random variables.

Introduction:

Let X denote the convolution of $n \geq 2$ independent, nonnegative and continuous random variables (rv's) with positive means and finite coefficients of variation; we may refer to these rv's as the *components*. If we denote the mean of component j by μ_j and its variance by σ_j^2, then it is well known that the mean and variance of X, μ_X and σ_X^2, are given by $\mu_X = \Sigma\mu_j$ and $\sigma_X^2 = \Sigma\sigma_j^2$ (due to statistical independence). We do not require the components to be identically distributed, but they must satisfy the regularity conditions of the central limit theorem (CLT); i.e., when $n \to \infty$, no single component should dominate the convolution. Equivalently, we require that as $n \to \infty$, $\mu_j/\mu_X \to 0$ and $\sigma_j^2/\sigma_X^2 \to 0$ for all j. Denote the density function of X by $f_X(x)$ and that of component j by $f_j(x)$. Whereas $f_j(0) > 0$ is allowed (e.g., if the component is distributed exponentially), it is easy to show by a limiting argument that $f_X(0) = 0$. For small n, if the components have very high coefficients of variation, the coefficient of variation of X may also be high (although it must tend to zero when n grows large). Finally, when $n \to \infty$, the CLT applies. We list these observations as three conditions:

(*i*) $f_X(x) = 0$ for $x \leq 0$,

(*ii*) σ_X/μ_X is unbounded,

(*iii*) as $n \to \infty$, $f_X(x) \to \dfrac{1}{\sigma_X\sqrt{2\pi}} \exp\left(-\dfrac{1}{2}\left(\dfrac{x-\mu_X}{\sigma_X}\right)^2\right)$.

The vast majority of conventional distributions do not satisfy all three conditions. The normal distribution is disqualified by (*i*), as the normal rv is not nonnegative. When we limit ourselves to nonnegative rv's, distributions that satisfy condition (*i*) often violate condition (*ii*). Conversely, most distributions that satisfy condition (*ii*), such as Weibull or gamma, violate condition (*i*) because they rely on $f_X(0) > 0$ in cases with high σ_X/μ_X. The most notable exception is the lognormal. We show that it satisfies all three conditions. But first, for completeness, we provide the conversion formulas necessary to fit a lognormal distribution for X based on the parameters

μ_X and σ_X^2. Recall that the lognormal random variable is based on a *core* normal random variable with mean m and variance s^2, and the random variable is defined as the exponent of that core. The density function is,

$$f(x) = \begin{cases} \dfrac{1}{xs\sqrt{2\pi}} \exp\left[\dfrac{-(\ln x - m)^2}{2s^2}\right] & ; \quad \textit{if } x > 0 \\ 0 & ; \quad \textit{otherwise} \end{cases}$$

and the mode is equal to $\exp(m - s^2)$. To evaluate m and s, let $cv = \sigma_X/\mu_X$, and we have,

$$s^2 = \ln(1 + cv^2); \; m = \ln \mu_x - \frac{s^2}{2}$$

If we wish to approximate the sum of n lognormal rv's with parameters m_j and variance s_j^2 we must first evaluate their means and variances, each given by,

$$\mu_j = \exp(m_j + s_j^2/2); \quad \sigma_j^2 = \mu_j^2[\exp(s_j^2) - 1]$$

Given these parameters, we can proceed to calculate μ_X and σ_X^2, and then calculate m and s. Such calculations are easy to program and hence should cause no difficulty in practice.

We can show that condition (*i*) is satisfied by the lognormal distribution because as $x \to 0^+$, $\ln x \to -\infty$. Condition (*ii*) is satisfied because cv is unconstrained. To show that condition (*iii*) is satisfied, all the following claims are subject to the stipulation that $n \to \infty$. By the law of large numbers, $cv = \sigma_X/\mu_X \to 0$ (because $\mu_j > 0$ and σ_j/μ_j is finite for all j). But $s^2 = \ln(1 + cv^2)$ so as $cv \to 0$, $s^2 \to cv^2$ and, equivalently, $s \to cv$. Also, for any x in the support of the distribution, $x/\mu_X \to 1$ *almost surely*. Therefore, $xs \to \sigma_X$ and thus $xs\sqrt{2\pi} = \sigma_X\sqrt{2\pi}$. It remains to show that $(\ln x - m)/s \to (x - \mu_X)/\sigma_X$. We can write $x = \mu_X(x/\mu_X)$, so $\ln x = \ln \mu_X + \ln(x/\mu_X)$. But $x/\mu_X \to 1$ so $\ln(x/\mu_X) \to (x - \mu_X)/\mu_X$. Recall that $m = \ln \mu_X - s^2/2$ and $s^2/2 \to 0$, so $m \to \ln \mu_X$. Substituting these values for $\ln x$ and m we obtain $(\ln x - m)/s \to (x - \mu_X)/s\mu_X$. Finally, $s\mu_X \to \sigma_X$, thus completing the proof.

To illustrate the efficacy of using the lognormal approximation we use it to represent the

k-Erlang distribution for various *k* values. We tested the mean absolute deviation (MAD) of the lognormal approximation, as a fraction of the mean, and compared it with that of the normal distribution. The results are given in the first four columns of Table 1.

Table 1: Comparing the relative MAD of *k*-Erlang and Chi-Square approximations with *k* d.f.

Erlang Case				Chi-Square Case			
k=	Normal	Lognormal	Ratio	d.f.	Normal	Lognormal	Ratio
1	0.314351	0.123926	0.394229	1	0.599782	0.202787	0.338102
2	0.159869	0.070044	0.438132	2	0.314351	0.123926	0.394229
3	0.107003	0.048841	0.456447	3	0.212118	0.089476	0.421823
4	0.08038-	0.037492	0.466434	4	0.159869	0.070044	0.438132
5	0.064357	0.030422	0.472708	5	0.128215	0.057551	0.448865
6	0.053658	0.025595	0.477011	6	0.107003	0.048841	0.456447
7	0.046008	0.022090	0.480143	7	0.091803	0.042421	0.462083
8	0.040266	0.019429	0.482526	8	0.080380	0.037492	0.466434
9	0.035798	0.017341	0.484399	9	0.071483	0.033589	0.469893
10	0.032223	0.015657	0.485909	10	0.064357	0.030422	0.472708
12	0.026857	0.013112	0.488196	12	0.053658	0.025595	0.477011
14	0.023024	0.011278	0.489844	14	0.046008	0.022090	0.480143
16	0.020147	0.009894	0.491089	16	0.040266	0.019429	0.482526
18	0.017910	0.008813	0.492062	18	0.035798	0.017341	0.484399
20	0.016120	0.007945	0.492843	20	0.032223	0.015657	0.485909
25	0.012897	0.006375	0.494257	25	0.025784	0.012599	0.488656
30	0.010748	0.005323	0.495205	30	0.021490	0.010541	0.490507
40	0.008062	0.004002	0.496395	40	0.016120	0.007945	0.492843
50	0.006450	0.003206	0.497112	50	0.012897	0.006375	0.494257
75	0.004300	0.002142	0.498073	75	0.008599	0.004267	0.496156
100	0.003225	0.001608	0.498555	100	0.006450	0.003206	0.497112
150	0.002150	0.001073	0.499040	150	0.004300	0.002142	0.498073
200	0.001613	0.000805	0.499284	200	0.003225	0.001608	0.498555

By the table, it is evident that the relative MAD of the lognormal approximation is at most 50% of the normal's in this case (i.e., the approximation is at least two times better). We obtained very similar results for the chi-square distribution, as shown in the subsequent columns of the table. Notice that, due to the larger variance of the chi-square distribution, convergence seems to be exactly twice as fast for the Erlang case.

Our results for these two cases are very encouraging: the lognormal approximation not only avoids violating conditions (*i*) and (*ii*) but also outperforms the normal approximation for

high k values for which the normal is highly unlikely to violate the conditions. On the one hand, when components are symmetric, if the normal approximation is highly unlikely to yield a negative result then the advantage goes to the normal (because a symmetric result is advantageous in that case). On the other hand, however, typical nonnegative rv's in practice are usually skewed to the right, in which case the lognormal tends to outperform the normal.

Conclusion:

We showed that the lognormal distribution satisfies three necessary conditions that a convolution of nonnegative continuous rv's must satisfy. It must have zero density at any nonpositive argument, it must support any coefficient of variation, and it must not violate the central limit theorem (CLT). Whereas most popular distributions that we might consider fail in one or more of these conditions—e.g., the normal fails the first two—the lognormal satisfies all of them. Therefore, we can formulate an alternative CLT for such rv's by replacing the normal distribution as the limiting case by the lognormal. That is, as $n \to \infty$, the distribution of the convolution of n independent nonnegative continuous random variables with positive means and finite coefficients of variation tends to lognormal. When we use this version of the CLT as the basis of an approximation for the convolution of few rv's, the results do not share the main weaknesses of the analogous normal approximation, namely that negative realizations are possible and the density at 0 is positive. Nonetheless, because the lognormal distribution always has a positive skew, the normal approximation may be better in the practical sense when the probability of nonpositive realizations is negligible and the components are not skewed, or skewed to the left. We believe that practical nonnegative random variables do tend to be positively skewed, and hence the use of the lognormal distribution is usually warranted.