Approximating the Mean and Variance of the Sum of Lognormally-Distributed Random Variables

Gary Schurman, MBE, CFA

June, 2013

The equation for the product of $N \in \{2, 3, 4, ..., \infty\}$ lognormally-distributed random variables is well known. If X_i is the value of the ith constant, m_i is the mean of the ith θ and v_i is the variance of the ith θ, then the equation for the product of N lognormally-distributed random variables is...

$$\prod_{i=1}^{N} X_i \exp\left\{\theta_i\right\} = \exp\left\{\sum_{i=1}^{N} \ln X_i + \sum_{i=1}^{N} \theta_i\right\} \text{ ...where... } \theta_i \sim N\left[m_i, v_i\right] \tag{1}$$

Whereas the sum of N normally-distributed random variables is normally-distributed the sum of N lognormally-distributed random variables is not lognormally-distributed. We want to use a lognormally-distributed random variable to approximate the sum of N lognormally-distributed random variables. We do this because the properties of the lognormal distribution are well known. If the random variable Y is the actual sum of N lognormally-distributed random variables and M and V are the mean and variance, respectively, of the distribution of Y then the equation for Y is...

$$Y = \sum_{i=1}^{N} X_i \exp\left\{\theta_i\right\} \text{ ...where... } \theta_i \sim N\left[m_i, v_i\right] \text{ ...and... } \ln Y \nsim N\left[M, V\right] \tag{2}$$

We will define the lognormally-distributed random variable $\bar{Y}$ to be an approximation of Equation (2) above. If the variable ϵ is an error term then the equation for $\bar{Y}$ is...

$$\bar{Y} = Y + \epsilon = \sum_{i=1}^{N} X_i \exp\left\{\theta_i\right\} + \epsilon \text{ ...where... } \theta_i \sim N\left[m_i, v_i\right] \text{ ...and... } \ln \bar{Y} \sim N\left[M, V\right] \tag{3}$$

In this white paper we will find an equation for the mean (M) and variance (V) of Equation (3) above utilizing a technique known as **Moment Matching**. To develop our equations we will use the following hypothetical problem...

Our Hypothetical Problem

Imagine that we have three asset classes in our investment portfolio. Table 1 below presents portfolio composition at time zero...

Table 1 - Portfolio Composition

Asset	Dollar	Expected Return		Return Volatility	
Class	Investment	Symbol	Value	Symbol	Value
S_1	100,000	μ_1	0.20	σ_1	0.30
S_2	200,000	μ_2	0.12	σ_2	0.18
S_3	300,000	μ_3	0.08	σ_3	0.10
Total	600,000				

Note: The expected returns in Table 1 are annual returns. The return volatilities in Table 1 are the standard deviation (annualized) of asset returns.

Table 2 below presents the correlations of asset class returns...

Table 2 - Asset Class Return Correlations

Description	Correlation
Asset class S_1 returns and Asset class S_2 returns	0.42
Asset class S_1 returns and Asset class S_3 returns	0.48
Asset class S_2 returns and Asset class S_3 returns	0.56

Our go-forward assumption is that asset returns are normally distributed and therefore asset values are lognormally-distributed.

Question: Given that the variable t represents time in years, what is the probability that portfolio value, which is \$600,000 at time $t = 0$, will be less than or equal to \$700,000 at time $t = 3$?

Actual Portfolio Value At Time T

Portfolio value at time zero is known with certainty. The equation for portfolio value at time zero (P_0) where N is the number of assets in the portfolio and S_i is the value of the ith asset is...

$$P_0 = \sum_{i=1}^{N} S_i \tag{4}$$

From the vantage point of time zero we do not know the value of asset S_i at time t but we do know it's probability distribution. We will define the random variable θ_i to be the return on asset S_i over the time interval $[0, t]$ and the random variable S_i^T to be the value of asset S_i at time t. The equation for the value of asset S_i at time t is...

$$S_i^T = S_i \exp\left\{\theta_i\right\} \text{ ...where... } \theta_i \sim N\left[m_i, v_i\right] \tag{5}$$

At time zero we expect to earn a rate of return equal to $\mu_i\, t$ on asset S_i over the time interval $[0, t]$. The equation for the expected value of asset S_i at time t is therefore...

$$\mathbb{E}\left[S_i^T\right] = S_i \exp\left\{\mu_i\, t\right\} \tag{6}$$

It can be shown via Stochastic Calculus that in order to get the result in Equation (6) the mean (m_i) and variance (v_i), respectively, of the random variable θ_i in Equation (5) must be...

$$m_i = \left(\mu_i - \frac{1}{2}\sigma_i^2\right) t \text{ ...and... } v_i = \sigma_i^2\, t \tag{7}$$

Portfolio value at time t is unknown at time zero and therefore is a random variable. Using Equation (5) above the equation for portfolio value at time t (P_t) is...

$$P_t = \sum_{i=1}^{N} S_i^T = \sum_{i=1}^{N} S_i \exp\left\{\theta_i\right\} \text{ ...where... } \theta_i \sim N\left[m_i, v_i\right] \tag{8}$$

Using Equation (8) above the first moment of the distribution of portfolio value at time t is...

$$\mathbb{E}\left[P_t\right] = \mathbb{E}\left[\sum_{i=1}^{N} S_i \exp\left\{\theta_i\right\}\right] = \sum_{i=1}^{N} \mathbb{E}\left[S_i \exp\left\{\theta_i\right\}\right] \tag{9}$$

Using Appendix Equation (37) and Equation (7) we can rewrite Equation (9) as...

$$\begin{aligned} \mathbb{E}\left[P_t\right] &= \sum_{i=1}^{N} \mathbb{E}\left[S_i \exp\left\{\theta_i\right\}\right] \\ &= \sum_{i=1}^{N} S_i \exp\left\{m_i + \frac{1}{2} v_i\right\} \\ &= \sum_{i=1}^{N} S_i \exp\left\{\mu_i\, t\right\} \end{aligned} \tag{10}$$

Using Equation (8) above the second moment of the distribution of portfolio value at time t is...

$$\mathbb{E}\bigg[P_t^2\bigg] = \mathbb{E}\bigg[\bigg(\sum_{i=1}^{N} S_i \exp\bigg\{\theta_i\bigg\}\bigg)^2\bigg] = \sum_{i=1}^{N}\sum_{j=1}^{N} \mathbb{E}\bigg[S_i\, S_j \exp\bigg\{\theta_i\bigg\} \exp\bigg\{\theta_j\bigg\}\bigg] \tag{11}$$

Using Appendix Equation (39) and Equation (7) we can rewrite Equation (9) as...

$$\begin{aligned}
\mathbb{E}\bigg[P_t^2\bigg] &= \sum_{i=1}^{N}\sum_{j=1}^{N} \mathbb{E}\bigg[S_i\, S_j \exp\bigg\{\theta_i\bigg\} \exp\bigg\{\theta_j\bigg\}\bigg] \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} S_i\, S_j \exp\bigg\{m_i + m_j + \frac{1}{2}\bigg(v_i + v_j + 2\sqrt{v_i}\,\sqrt{v_j}\,\rho_{i,j}\bigg)\bigg\} \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} S_i\, S_j \exp\bigg\{\mu_i\, t + \mu_j\, t + \sqrt{\sigma_i^2\, \sigma_j^2\, t^2}\,\rho_{i,j}\bigg)\bigg\}
\end{aligned} \tag{12}$$

Approximating Portfolio Value At Time T

We will define the random variable $\bar{P}_t$ to be the lognormally-distributed approximation to actual portfolio value at time t. Given that M is mean and V is variance the approximation $\bar{P}_t$ has the following distribution...

$$\ln \bar{P}_t \sim N\bigg[M, V\bigg] \tag{13}$$

Given the definition above and Appendix Equation (37), the equation for the first moment of the distribution of $\bar{P}_t$ is...

$$\mathbb{E}\bigg[\bar{P}_t\bigg] = P_0 \exp\bigg\{M + \frac{1}{2}\,V\bigg\} \tag{14}$$

Given the definition above and Appendix Equation (38), the equation for the second moment of the distribution of $\bar{P}_t$ is...

$$\mathbb{E}\bigg[\bar{P}_t^2\bigg] = P_0^2 \exp\bigg\{2\,M + 2\,V\bigg\} \tag{15}$$

To estimate the mean M and variance V of $\bar{P}_t$ we will employ a moment matching technique and match Equations (10) and (12), which are the first and second moments, respectively, of the distribution of actual portfolio value P_t, with Equations (14) and (15), which are the first and second moments, respectively, of the distribution of portfolio value approximation $\bar{P}_t$. Let's begin...

Equation (10) gives us the equation for the first moment of the distribution of P_t. We can rewrite that equation in vector product notation as...

$$\text{if...}\ \vec{\mathbf{u}} = \begin{bmatrix} S_1 \\ S_2 \\ * \\ * \\ S_N \end{bmatrix} \ \text{...and...}\ \vec{\mathbf{v}} = \begin{bmatrix} \exp\{\mu_1 t\} \\ \exp\{\mu_2 t\} \\ * \\ * \\ \exp\{\mu_N t\} \end{bmatrix} \ \text{...then...}\ \mathbb{E}\bigg[P_t\bigg] = \vec{\mathbf{u}}^T \vec{\mathbf{v}} \tag{16}$$

Equation (12) gives us the equation for the second moment of the distribution of P_t. We can rewrite that equation in matrix:vector product notation as...

$$\text{if...}\ \vec{\mathbf{u}} = \begin{bmatrix} S_1 \\ S_2 \\ * \\ * \\ S_N \end{bmatrix} \ \text{...and...}\ \mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & * & a_{1,N} \\ a_{2,1} & a_{2,2} & * & a_{2,N} \\ * & * & * & \\ * & * & * & * \\ a_{N,1} & a_{N,2} & * & a_{N,N} \end{bmatrix} \ \text{...then...}\ \mathbb{E}\bigg[P_t^2\bigg] = \vec{\mathbf{u}}^T \mathbf{A}\, \vec{\mathbf{u}} \tag{17}$$

...where the equation for each element of Matrix $\mathbf{A}$ is...

$$a_{i,j} = \exp\bigg\{\mu_i\, t + \mu_j\, t + \sqrt{\sigma_i^2\, \sigma_j^2\, t^2}\,\rho_{i,j}\bigg)\bigg\} \tag{18}$$

If we match Equation (10), which is the first moment of the distribution of P_t, with Equation (14), which is the first moment of the distribution of $\bar{P}_t$, then we get the following linear equation...

$$
\begin{aligned}
\mathbb{E}\left[\bar{P}_t\right] &= \mathbb{E}\left[P_t\right] \\
P_0 \exp\left\{M + \frac{1}{2}\,V\right\} &= \vec{\mathbf{u}}^T\vec{\mathbf{v}} \\
\ln\left\{P_0\right\} + M + \frac{1}{2}\,V &= \ln\left\{\vec{\mathbf{u}}^T\vec{\mathbf{v}}\right\} \\
M + \frac{1}{2}\,V &= \ln\left\{\vec{\mathbf{u}}^T\vec{\mathbf{v}}\right\} - \ln\left\{P_0\right\} \qquad (19)
\end{aligned}
$$

If we match Equation (12), which is the second moment of the distribution of P_t, with Equation (15), which is the second moment of the distribution of $\bar{P}_t$, then we get the following linear equation...

$$
\begin{aligned}
\mathbb{E}\left[\bar{P}_t^2\right] &= \mathbb{E}\left[P_t^2\right] \\
P_0^2 \exp\left\{2\,M + 2\,V\right\} &= \vec{\mathbf{u}}^T\,\mathbf{A}\,\vec{\mathbf{u}} \\
2\,\ln\left\{P_0\right\} + 2\,M + 2\,V &= \ln\left\{\vec{\mathbf{u}}^T\,\mathbf{A}\,\vec{\mathbf{u}}\right\} \\
2\,M + 2\,V &= \ln\left\{\vec{\mathbf{u}}^T\,\mathbf{A}\,\vec{\mathbf{u}}\right\} - 2\,\ln\left\{P_0\right\} \qquad (20)
\end{aligned}
$$

Note that we now have two linear equations with the two unknowns being M and V. Using Equations (19) and (20) the system of linear equations that we must solve is...

$$
\begin{aligned}
M + \frac{1}{2}\,V &= \ln\left\{\vec{\mathbf{u}}^T\vec{\mathbf{v}}\right\} - \ln\left\{P_0\right\} \\
2\,M + 2\,V &= \ln\left\{\vec{\mathbf{u}}^T\,\mathbf{A}\,\vec{\mathbf{u}}\right\} - 2\,\ln\left\{P_0\right\} \qquad (21)
\end{aligned}
$$

We will make the following definitions...

$$
\mathbf{B} = \begin{bmatrix} 1.00 & 0.50 \\ 2.00 & 2.00 \end{bmatrix} \text{ ...and... } \vec{\mathbf{x}} = \begin{bmatrix} M \\ V \end{bmatrix} \text{ ...and... } \vec{\mathbf{y}} = \begin{bmatrix} \ln\left\{\vec{\mathbf{u}}^T\vec{\mathbf{v}}\right\} - \ln\left\{P_0\right\} \\ \ln\left\{\vec{\mathbf{u}}^T\,\mathbf{A}\,\vec{\mathbf{u}}\right\} - 2\,\ln\left\{P_0\right\} \end{bmatrix} \qquad (22)
$$

Given the definitions in Equation (22) above we can write the system of linear equations as represented by Equation (21) as...

$$
\mathbf{B}\,\vec{\mathbf{x}} = \vec{\mathbf{y}} \qquad (23)
$$

Our goal is to match moments such that the mean and variance of the lognormally-distributed approximation of portfolio value $\bar{P}_t$ can be determined. Using Equation (23) we can solve for vector x, whose first element is the mean of $\bar{P}_t$ and the second element is the variance of $\bar{P}_t$, via the following equation...

$$
\vec{\mathbf{x}} = \mathbf{B}^{-1}\vec{\mathbf{y}} \qquad (24)
$$

Our goal has been accomplished!

The Answer To Our Hypothetical Problem

We will define the random variable θ_P to be the return on our portfolio over the time interval $[0, t]$. If portfolio value is lognormally-distributed then the equation for portfolio value at time t is...

$$
P_t = P_0 \exp\left\{\theta_P\right\} \text{ ...where... } \theta_P \sim N\left[M, V\right] \qquad (25)
$$

We will normalize θ_P as follows...

$$\theta_P = M + \sqrt{V}\,Z \quad \text{...where...} \quad Z \sim N\left[0,1\right] \tag{26}$$

Using the definition in Equation (26) above we can rewrite Equation (25) as...

$$P_t = P_0 \exp\left\{M + \sqrt{V}\,Z\right\} \quad \text{...where...} \quad Z \sim N\left[0,1\right] \tag{27}$$

After taking the log of both sides of Equation (27) and solving for Z...

$$Z = \frac{\ln P_t - \ln P_0 - M}{\sqrt{V}} \quad \text{...where...} \quad Z \sim N\left[0,1\right] \tag{28}$$

Using Equation (28) above and noting that CDF is the cumulative normal disribution function, the probability that actual portfolio value at time t will be less than or equal to P_t is...

$$\text{Prob}\left[Z\right] = \int_{-\infty}^{Z} \sqrt{2\,\pi}\,\exp\left\{-\frac{1}{2}\,x^2\right\}\delta\,x = \text{CDF}\left[Z\right] \tag{29}$$

Using Tables 1 and 2 and Equations (16) and (17) above we will make the following matrix and vector definitions (see Appendix Equations (40) and (41) for example calculations)...

$$\vec{\mathbf{u}} = \begin{bmatrix} 100 \\ 200 \\ 300 \end{bmatrix} \quad \text{...and...} \quad \vec{\mathbf{v}} = \begin{bmatrix} 1.82212 \\ 1.43333 \\ 1.27125 \end{bmatrix} \quad \text{...and...} \quad \mathbf{A} = \begin{bmatrix} 4.34924 & 2.79558 & 2.41863 \\ 2.79558 & 2.26415 & 1.87806 \\ 2.41863 & 1.87806 & 1.66529 \end{bmatrix} \tag{30}$$

Using Equations (16) and (30) the first moment of the distribution of actual portfolio value P_t is...

$$\mathbb{E}\left[P_t\right] = \vec{\mathbf{u}}^T\vec{\mathbf{v}} = 850 \tag{31}$$

Using Equation (17) and (30) the second moment of the distribution of actual portfolio value P_t is...

$$\mathbb{E}\left[P_t^2\right] = \vec{\mathbf{u}}^T\mathbf{A}\,\vec{\mathbf{u}} = 766\,243 \tag{32}$$

Using Equation (22) vector $\vec{\mathbf{y}}$ is...

$$\vec{\mathbf{y}} = \begin{bmatrix} \ln 850 - \ln 600 \\ \ln 766243 - 2\ln 600 \end{bmatrix} = \begin{bmatrix} 0.3486 \\ 0.7554 \end{bmatrix} \tag{33}$$

Using Equations (24) and (33) the solution to vector $\vec{\mathbf{x}}$ is...

$$\vec{\mathbf{x}} = \mathbf{B}^{-1}\vec{\mathbf{y}} = \begin{bmatrix} 2.00 & (0.50) \\ (2.00) & 1.00 \end{bmatrix} \begin{bmatrix} 0.3486 \\ 0.7554 \end{bmatrix} = \begin{bmatrix} 0.3195 \\ 0.0582 \end{bmatrix} \tag{34}$$

We want to find the probability that portfolio value at time t will be less than \$700,000. Using Equation (28) and setting the random variable $P_t = 700$ the value of the random variable Z is...

$$Z = \frac{\ln 700 - \ln 600 - 0.3195}{\sqrt{0.0582}} = -0.6855 \tag{35}$$

Using Equation (29) the probability that the value of our portfolio will be less than or equal to \$700,000 at the end of year three is...

$$\text{Prob}\left[P_t \le 700\right] = 0.2465 \tag{36}$$

We have solved our hypothetical problem! The probability that portfolio value at time t will be less than \$700,000 is approximately 0.25.

Appendix

A. The equation for the first moment of the distribution of the lognormally-distributed random variable $C\exp\{\theta\}$ (see The Lognormal Distribution, Schurman April, 2012) is...

$$\mathbb{E}\left[C\exp\left\{\theta\right\}\right] \text{ ...where... } \theta \sim N\left[m, v\right] = C\exp\left\{m + \frac{1}{2}v\right\} \tag{37}$$

B. The equation for the second moment of the distribution of the lognormally-distributed random variable $C\exp\{\theta\}$ (see The Lognormal Distribution, Schurman April, 2012) is...

$$\mathbb{E}\left[\left(C\exp\left\{\theta\right\}\right)^2\right] = C^2\exp\left\{2\,m + 2\,v\right\} \tag{38}$$

C. The equation for the expected value of the product of two lognormally-distributed random variables (see The Mean and Variance of the Product of Two Lognormally-Distributed Random Variables, Schurman September, 2012) is...

$$\mathbb{E}\left[A\exp\left\{\theta_a\right\}B\exp\left\{\theta_b\right\}\right] = A\,B\exp\left\{m_a + m_b + \frac{1}{2}\left(v_a + v_b + 2\sqrt{v_a}\sqrt{v_b}\,\rho_{a,b}\right)\right\} \tag{39}$$

D. Calculation example: Equation (30), Vector $\vec{\mathbf{v}}$, first element...

$$\vec{\mathbf{v}}_1 = \exp\left\{\mu_1\,t\right\} = \exp\left\{(0.20)(3.00)\right\} = 1.82212 \tag{40}$$

E. Calculation example: Equation (30), Matrix $\mathbf{A}$, row one column 2...

$$\begin{aligned}\mathbf{A}_{1,2} &= \exp\left\{\mu_1\,t + \mu_2\,t + \sqrt{\sigma_1^2\,\sigma_2^2\,t^2}\,\rho_{1,2}\Big)\right\}\\ &= \exp\left\{(0.20)(3.00) + (0.12)(3.00) + \sqrt{(0.30^2)(0.18^2)(3.00^2)} \times 0.42\Big)\right\}\\ &= 2.79558\end{aligned} \tag{41}$$