Lecture #3: Normal and Lognormal Random Variables

The purpose of this lecture is to remind you of some of the key properties of normal and lognormal random variables which are basic objects in the mathematical theory of finance. (Of course, you already know of the ubiquity of the normal distribution from your elementary probability classes since it arises in the central limit theorem, and if you have studied any actuarial science you already realize how important lognormal random variables are.)

Recall that a continuous random variable Z is said to have a *normal distribution with mean* 0 *and variance* 1 if the density function of Z is

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty.$$

If Z has such a distribution, we write $Z \sim \mathcal{N}(0,1)$.

Exercise 3.1. Show directly that if $Z \sim \mathcal{N}(0,1)$, then $\mathbb{E}(Z) = 0$ and $\mathrm{Var}(Z) = 1$. That is, calculate

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{z^2}{2}}\,\mathrm{d}z \quad \text{and} \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}}\,\mathrm{d}z$$

using only results from elementary calculus. This calculation justifies the use of the "mean 0 and variance 1" phrase in the definition above.

Let $\mu \in \mathbb{R}$ and let $\sigma > 0$. We say that a continuous random variable X has a *normal distribution with mean* μ *and variance* σ^2 if the density function of X is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

If X has such a distribution, we write $X \sim \mathcal{N}(\mu, \sigma^2)$.

Shortly, you will be asked to prove the following result which establishes the relationship between the random variables $Z \sim \mathcal{N}(0,1)$ and $X \sim \mathcal{N}(\mu, \sigma^2)$.

Theorem 3.2. *Suppose that $Z \sim \mathcal{N}(0,1)$, and let $\mu \in \mathbb{R}$, $\sigma > 0$ be constants. If the random variable X is defined by $X = \sigma Z + \mu$, then $X \sim \mathcal{N}(\mu, \sigma^2)$. Conversely, if $X \sim \mathcal{N}(\mu, \sigma^2)$, and the random variable Z is defined by*

$$Z = \frac{X - \mu}{\sigma},$$

then $Z \sim \mathcal{N}(0,1)$.

Let

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,\mathrm{d}x$$

denote the standard normal cumulative distribution function. That is, $\Phi(z) = \mathbf{P}\{Z \le z\} = F_Z(z)$ is the distribution function of a random variable $Z \sim \mathcal{N}(0,1)$.

Remark. Higham [11] writes N instead of Φ for the standard normal cumulative distribution function. The notation Φ is far more common in the literature, and so we prefer to use it instead of N.

Exercise 3.3. Show that $1 - \Phi(z) = \Phi(-z)$.

Exercise 3.4. Show that if $X \sim \mathcal{N}(\mu, \sigma^2)$, then the distribution function of X is given by

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Exercise 3.5. Use the result of Exercise 3.4 to complete the proof of Theorem 3.2.

The next two exercises are extremely important for us. In fact, these exercises ask you to prove special cases of the Black-Scholes formula.

Notation. We write $x^+ = \max\{0, x\}$ to denote the *positive part* of x.

Exercise 3.6. Suppose that $Z \sim \mathcal{N}(0, 1)$, and let $c > 0$ be a constant. Compute

$$\mathbb{E}[\,(e^Z - c)^+\,].$$

You will need to express your answer in terms of Φ.

Answer. $e^{1/2}\,\Phi(1 - \log c) - c\,\Phi(-\log c)$

Exercise 3.7. Suppose that $Z \sim \mathcal{N}(0, 1)$, and let $a > 0$, $b > 0$, and $c > 0$ be constants. Compute

$$\mathbb{E}[\,(ae^{bZ} - c)^+\,].$$

You will need to express your answer in terms of Φ.

Answer. $ae^{b^2/2}\,\Phi\left(b + \frac{1}{b}\log\frac{a}{c}\right) - c\,\Phi\left(\frac{1}{b}\log\frac{a}{c}\right)$

Recall that the characteristic function of a random variable X is the function $\varphi_X : \mathbb{R} \to \mathbb{C}$ given by $\varphi_X(t) = \mathbb{E}(e^{itX})$.

Exercise 3.8. Show that if $Z \sim \mathcal{N}(0, 1)$, then the characteristic function of Z is

$$\varphi_Z(t) = \exp\left\{-\frac{t^2}{2}\right\}.$$

Exercise 3.9. Show that if $X \sim \mathcal{N}(\mu, \sigma^2)$, then the characteristic function of X is

$$\varphi_X(t) = \exp\left\{i\mu t - \frac{\sigma^2 t^2}{2}\right\}.$$

The importance of characteristic functions is that they completely characterize the distribution of a random variable since the characteristic function always exists (unlike moment generating functions which do not always exist).

Theorem 3.10. *Suppose that X and Y are random variables. The characteristic functions φ_X and φ_Y are equal if and only if X and Y are equal in distribution (that is, $F_X = F_Y$).*

Proof. For a proof, see Theorem 4.1.2 on page 160 of [9]. □

Exercise 3.11. One consequence of this theorem is that it allows for an alternative solution to Exercise 3.5. That is, use characteristic functions to complete the proof of Theorem 3.2.

We will have occasion to analyze sums of normal random variables. The purpose of the next several exercises and results is to collect all of the facts that we will need. The first exercise shows that a linear combination of independent normals is again normal.

Exercise 3.12. Suppose that $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are independent. Show that for any a, $b \in \mathbb{R}$,

$$aX_1 + bX_2 \sim \mathcal{N}\left(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2\right).$$

Of course, whenever two random variables are independent, they are necessarily uncorrelated. However, the converse is not true in general, even in the case of normal random variables. As the following example shows, uncorrelated normal random variables need not be independent.

Example 3.13. Suppose that $X_1 \sim \mathcal{N}(0,1)$ and suppose further that Y is independent of X_1 with $\mathbf{P}\{Y = 1\} = \mathbf{P}\{Y = -1\} = 1/2$. If we set $X_2 = YX_1$, then it follows that $X_2 \sim \mathcal{N}(0,1)$. (Verify this fact.) Furthermore, X_1 and X_2 are uncorrelated since

$$\mathrm{Cov}(X_1, X_2) = \mathbb{E}(X_1X_2) = \mathbb{E}(X_1^2Y) = \mathbb{E}(X_1^2)\mathbb{E}(Y) = 1 \cdot 0 = 0$$

using the fact that X_1 and Y are independent. However, X_1 and X_2 are not independent since

$$\mathbf{P}\{X_1 \ge 1, X_2 \ge 1\} = \mathbf{P}\{X_1 \ge 1, Y = 1\} = \mathbf{P}\{X_1 \ge 1\}\mathbf{P}\{Y = 1\} = \frac{1}{2}\mathbf{P}\{X_1 \ge 1\}$$

whereas

$$\mathbf{P}\{X_1 \ge 1\}\mathbf{P}\{X_2 \ge 1\} = [\mathbf{P}\{X_1 \ge 1\}]^2.$$

Since $\mathbf{P}\{X_1 \ge 1\}$ does not equal either 0 or 1/2 (it actually equals $\doteq 0.1587$) we see that

$$\frac{1}{2}\mathbf{P}\{X_1 \ge 1\} \ne [\mathbf{P}\{X_1 \ge 1\}]^2.$$

An extension of this same example also shows that the sum of uncorrelated normal random variables need not be normal.

Example 3.13 (continued). We will now show that $X_1 + X_2$ is not normally distributed. If $X_1 + X_2$ were normally distributed, then it would necessarily be the case that for any $x \in \mathbb{R}$, we would have $\mathbf{P}\{X_1 + X_2 = x\} = 0$. Indeed, this is true for any continuous random variable. But we see that $\mathbf{P}\{X_1 + X_2 = 0\} = \mathbf{P}\{Y = -1\} = 1/2$ which shows that $X_1 + X_2$ cannot be a normal random variable (let alone a continuous random variable).

However, if we have a bivariate normal random vector $\mathbf{X} = (X_1, X_2)'$, then independence of the components and no correlation between them are equivalent.

Theorem 3.14. *Suppose that* $\mathbf{X} = (X_1, X_2)'$ *has a bivariate normal distribution so that the components of* $\mathbf{X}$*, namely* X_1 *and* X_2*, are each normally distributed. Furthermore,* X_1 *and* X_2 *are uncorrelated if and only if they are independent.*

Proof. For a proof, see Theorem V.7.1 on page 133 of Gut [8]. □

Two important variations on the previous results are worth mentioning.

Theorem 3.15 (Cramér)**.** *If* X *and* Y *are independent random variables such that* $X+Y$ *is normally distributed, then* X *and* Y *themselves are each normally distributed.*

Proof. For a proof of this result, see Theorem 19 on page 53 of [6]. □

In the special case when X and Y are also identically distributed, Cramér's theorem is easy to prove.

Exercise 3.16. Suppose that X and Y are independent and identically distributed random variables such that $X + Y \sim \mathcal{N}(2\mu, 2\sigma^2)$. Prove that $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y \sim \mathcal{N}(\mu, \sigma^2)$.

Example 3.13 showed that uncorrelated normal random variables need not be independent and need not have a normal sum. However, if uncorrelated normal random variables are known to have a normal sum, then it must be the case that they are independent.

Theorem 3.17. *If* $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ *and* $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ *are normally distributed random variables with* $\operatorname{Cov}(X_1, X_2) = 0$*, and if* $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$*, then* X_1 *and* X_2 *are independent.*

Proof. In order to prove that X_1 and X_2 are independent, it is sufficient to prove that the characteristic function of $X_1 + X_2$ equals the product of the characteristic functions of X_1 and X_2. Since $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ we see using Exercise 3.9 that

$$\varphi_{X_1+X_2}(t) = \exp\left\{ i(\mu_1 + \mu_2)t - \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2} \right\}.$$

Furthermore, since $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ we see that

$$\varphi_{X_1}(t)\varphi_{X_2}(t) = \exp\left\{ i\mu_1 t - \frac{\sigma_1^2 t^2}{2} \right\} \cdot \exp\left\{ i\mu_2 t - \frac{\sigma_2^2 t^2}{2} \right\} = \exp\left\{ i(\mu_1 + \mu_2)t - \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2} \right\}.$$

In other words,

$$\varphi_{X_1}(t)\varphi_{X_2}(t) = \varphi_{X_1+X_2}(t)$$

which establishes the result. □

Remark. Actually, the assumption that $\operatorname{Cov}(X_1, X_2) = 0$ is unnecessary in the previous theorem. The same proof shows that if $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are normally distributed random variables, and if $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$, then X_1 and X_2 are independent. It is now a consequence that $\operatorname{Cov}(X_1, X_2) = 0$.

A variation of the previous result can be proved simply by equating variances.

Exercise 3.18. If $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are normally distributed random variables, and if $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)$, then $\operatorname{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2$ and $\operatorname{Corr}(X_1, X_2) = \rho$.

Our final result gives conditions under which normality is preserved for limits in distribution. Before stating this theorem, we need to recall the definition of *convergence in distribution.*

Definition 3.19. Suppose that $X_1, X_2, \ldots$ and X are random variables with distribution functions F_n, $n = 1, 2, \ldots$, and F, respectively. We say that X_n *converges in distribution* to X as $n \to \infty$ if

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for all $x \in \mathbb{R}$ at which F is continuous.

The relationship between convergence in distribution and characteristic functions is extremely important for us.

Theorem 3.20. *Suppose that $X_1, X_2, \ldots$ are random variables with characteristic functions φ_{X_n}, $n = 1, 2, \ldots$. It then follows that $\varphi_{X_n}(t) \to \varphi_X(t)$ as $n \to \infty$ for all $t \in \mathbb{R}$ if and only if X_n converges in distribution to X.*

Proof. For a proof of this result, see Theorem 5.9.1 on page 238 of [9]. □

It is worth noting that in order to apply the result of the previous theorem we must know *a priori* what the limiting random variable X is. In the case when we only know that the characteristic functions converge to something, we must be a bit more careful.

Theorem 3.21. *Suppose that $X_1, X_2, \ldots$ are random variables with characteristic functions φ_{X_n}, $n = 1, 2, \ldots$. If $\varphi_{X_n}(t)$ converges to some function $\varphi(t)$ as $n \to \infty$ for all $t \in \mathbb{R}$ and $\varphi(t)$ is continuous at 0, then there exists a random variable X with characteristic function φ such that X_n converges in distribution to X.*

Proof. For a proof of this result, see Theorem 5.9.2 on page 238 of [9]. □

Remark. The statement of the central limit theorem is really a statement about convergence in distribution, and its proof follows after a careful analysis of characteristic functions from Theorems 3.10 and 3.21.

We are now ready to prove that normality is preserved under convergence in distribution. The proof uses a result known as Slutsky's theorem, and so we will state and prove this first.

Theorem 3.22 (Slutsky)**.** *Suppose that the random variables X_n, $n = 1, 2, \ldots$, converge in distribution to X and that the sequence of real numbers a_n, $n = 1, 2, \ldots$, converges to the finite real number a. It then follows that $X_n + a_n$ converges in distribution to $X + a$ and that $a_n X_n$ converges in distribution to aX.*

Proof. We begin by observing that for $\varepsilon > 0$ fixed, we have

$$\begin{aligned}\mathbf{P}\{X_n + a_n \le x\} &= \mathbf{P}\{X_n + a_n \le x, |a_n - a| < \varepsilon\} + \mathbf{P}\{X_n + a_n \le x, |a_n - a| > \varepsilon\} \\ &\le \mathbf{P}\{X_n + a_n \le x, |a_n - a| < \varepsilon\} + \mathbf{P}\{|a_n - a| > \varepsilon\} \\ &\le \mathbf{P}\{X_n \le x - a + \varepsilon\} + \mathbf{P}\{|a_n - a| > \varepsilon\}\end{aligned}$$

That is,

$$F_{X_n+a_n}(x) \le F_{X_n}(x - a + \varepsilon) + \mathbf{P}\{|a_n - a| > \varepsilon\}.$$

Since $a_n \to a$ as $n \to \infty$ we see that $\mathbf{P}\{|a_n - a| > \varepsilon\} \to 0$ as $n \to \infty$ and so

$$\limsup_{n\to\infty} F_{X_n+a_n}(x) \le F_X(x - a + \varepsilon)$$

for all points $x - a + \varepsilon$ at which F is continuous. Similarly,

$$\liminf_{n\to\infty} F_{X_n+a_n}(x) \ge F_X(x - a - \varepsilon)$$

for all points $x - a - \varepsilon$ at which F is continuous. Since $\varepsilon > 0$ can be made arbitrarily small and since F_X has at most countably many points of discontinuity, we conclude that

$$\lim_{n\to\infty} F_{X_n+a_n}(x) = F_X(x - a) = F_{X+a}(x)$$

for all $x \in \mathbb{R}$ at which F_{X+a} is continuous. The proof that $a_n X_n$ converges in distribution to aX is similar. □

Exercise 3.23. Complete the details to show that $a_n X_n$ converges in distribution to aX.

Theorem 3.24. *Suppose that $X_1, X_2, \ldots$ is a sequence of random variables with $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots$. If the limits*

$$\lim_{n\to\infty} \mu_n \quad \textit{and} \quad \lim_{n\to\infty} \sigma_n^2$$

each exist and are finite, then the sequence $\{X_n, n = 0, 1, 2, \ldots\}$ converges in distribution to a random variable X. Furthermore, $X \sim \mathcal{N}(\mu, \sigma^2)$ where

$$\mu = \lim_{n\to\infty} \mu_n \quad \textit{and} \quad \sigma^2 = \lim_{n\to\infty} \sigma_n^2.$$

Proof. For each n, let

$$Z_n = \frac{X_n - \mu_n}{\sigma_n}$$

so that $Z_n \sim \mathcal{N}(0, 1)$ by Theorem 3.2. Clearly, Z_n converges in distribution to some random variable Z with $Z \sim \mathcal{N}(0, 1)$. By Slutsky's theorem, since Z_n converges in distribution to Z, it follows that $X_n = \sigma_n Z_n + \mu_n$ converges in distribution to $\sigma Z + \mu$. If we now define $X = \sigma Z + \mu$, then X_n converges in distribution to X and it follows from Theorem 3.2 that $X \sim \mathcal{N}(\mu, \sigma^2)$. □

We end this lecture with a brief discussion of lognormal random variables. Recall that if $X \sim \mathcal{N}(\mu, \sigma^2)$, then the *moment generating function* of X is

$$m_X(t) = \mathbb{E}(e^{tX}) = \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}.$$

Exercise 3.25. Suppose that $X \sim \mathcal{N}(\mu, \sigma^2)$ and let $Y = e^X$.

(a) Determine the density function for Y

(b) Determine the distribution function for Y. *You will need to express your answer in terms of* Φ.

(c) Compute $\mathbb{E}(Y)$ and $\text{Var}(Y)$. *Hint: Use the moment generating function of* X.

Answer. (c) $\mathbb{E}(Y) = \exp\{\mu + \frac{\sigma^2}{2}\}$ and $\text{Var}(Y) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$.

Definition 3.26. We say that a random variable Y has a *lognormal distribution with parameters* μ *and* σ^2, written

$$Y \sim \mathcal{LN}(\mu, \sigma^2),$$

if $\log(Y)$ is normally distributed with mean μ and variance σ^2. That is, $Y \sim \mathcal{LN}(\mu, \sigma^2)$ iff $\log(Y) \sim \mathcal{N}(\mu, \sigma^2)$. Equivalently, $Y \sim \mathcal{LN}(\mu, \sigma^2)$ iff $Y = e^X$ with $X \sim \mathcal{N}(\mu, \sigma^2)$.

Exercise 3.27. Suppose that $Y_1 \sim \mathcal{LN}(\mu_1, \sigma_1^2)$ and $Y_2 \sim \mathcal{LN}(\mu_2, \sigma_2^2)$ are independent lognormal random variables. Prove that $Z = Y_1 \cdot Y_2$ is lognormally distributed and determine the parameters of Z.

Remark. As shown in STAT 351, if a random variable Y has a lognormal distribution, then the moment generating function of Y does not exist.