

Learning Module 6: Simulation Methods

LOS 6a: explain the relationship between normal and lognormal distributions and why the lognormal distribution is used to model asset prices when using continuously compounded asset returns

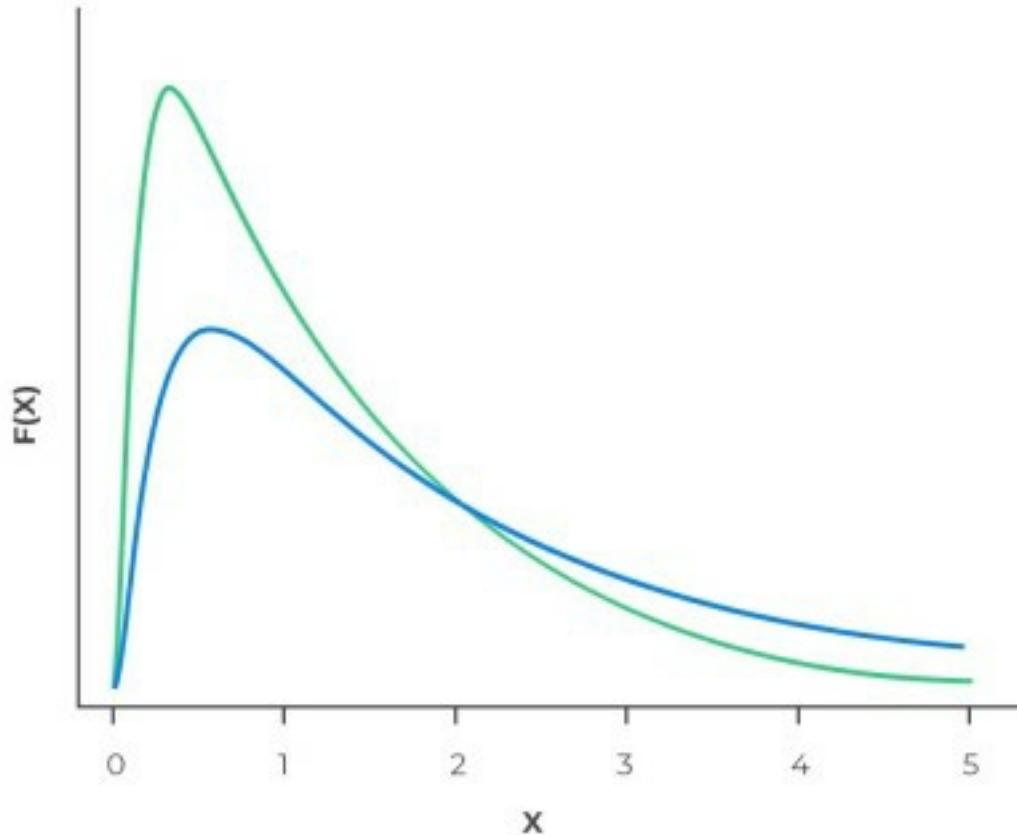
A random variable Y is lognormally distributed if its natural logarithm, $\ln Y$, is normally distributed. The opposite is true. If $\ln Y$ is normally distributed, then Y is lognormally distributed.

The lognormal distribution is positively skewed, meaning it's skewed to the right and has a long right tail. In this distribution, values are bounded by 0. Typically, the mean is greater than the mode.

Consider the following graph of two probability density functions (pdfs) of two lognormal distributions.



The lognormal distribution



Like the normal distribution, two parameters – the mean and variance of the associated normal distribution – fully describe the lognormal distribution.

Expressions for Mean and Variance of Lognormal Distribution

Assume that X is normally distributed with the mean μ and variance σ^2 . Also, define the variable $Y = e^X$.

Then $\ln Y = \ln(e^X) = X$ is lognormally distributed with the following mean and variance expressions:

$$\text{Mean} = \mu_L = e^{(\mu + \frac{1}{2}\sigma^2)}$$

$$\text{Variance} = \sigma_L^2 = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$$

Why the Lognormal Distribution is Used to Model Stock Prices

The lognormal distribution works well for modeling asset prices that cannot be negative because it has a lower bound at zero.

When the continuously compounded returns on a stock follow a normal distribution, the stock prices follow a lognormal distribution. Note that even if returns do not follow a normal distribution, the lognormal distribution is still the most appropriate for stock prices.

Continuously Compounded Rate of Return

Remember that given the investment horizon from time $t = 0$ to time $t = T$, the continuously compounded return of a stock is given by:

$$r_{0,T} = \ln\left(\frac{P_T}{P_0}\right)$$

If we apply the exponential function on both sides of the equation, we have the following:

$$P_T = P_0 e^{r_{0,T}}$$

Note that $\frac{P_T}{P_0}$ can be written as:

$$\frac{P_T}{P_0} = \left(\frac{P_T}{P_{T-1}}\right) \left(\frac{P_{T-1}}{P_{T-2}}\right) \dots \left(\frac{P_1}{P_0}\right)$$

If we take natural logarithm on both sides of the above equation:

$$\begin{aligned} \ln\left(\frac{P_T}{P_0}\right) &= \ln\left(\left(\frac{P_T}{P_{T-1}}\right)\left(\frac{P_{T-1}}{P_{T-2}}\right)\dots\left(\frac{P_1}{P_0}\right)\right) \\ &\Rightarrow r_{0,T} = r_{T-1,T} + r_{T-2,T-1} + \dots + r_{0,1} \end{aligned}$$

Therefore, the continuously compounded return to time T equals the sum of one-period continuously compounded returns.

Remember that a linear combination of normal random variables is also normal. Therefore, if the shorter period returns $r_{T-1,T}, r_{T-2,T-1}, \dots, r_{0,1}$ are normally distributed or approximately normal, the $r_{0,T}$ is approximately normal.

As such, if we assume that the one-period continuously compounded returns $r_{T-1,T}, r_{T-2,T-1}, \dots, r_{0,1}$ are independently and identically distributed (i.i.d) random variables, the mean of μ and variance of σ^2 , then:

- The expected value of the continuously compounded return over a holding period of T periods is given by:

$$E(r_{0,T}) = E(r_{T-1,T}) + E(r_{T-2,T-1}) + \dots + E(r_{0,1}) = \mu T$$

- The variance of the continuously compounded return over a holding period is given by:

$$\sigma^2(r_{0,T}) = \sigma^2 T$$

The standard deviation of the continuously compounded returns, also known as volatility, is given by:

$$\sigma(r_{0,T}) = \sigma \sqrt{T}$$

In other words, if $r_{T-1,T}, r_{T-2,T-1}, \dots, r_{0,1}$ are normally distributed with the mean of μ and variance of σ^2 then $r_{0,T}$ is normally distributed with the mean of μT and variance of $\sigma^2 T$.

Let us go back to the formula:

$$P_T = P_0 e^{r_{0,T}}$$

If X is normally distributed with the mean μ and variance σ^2 and that $Y = e^X$ then, $\ln Y = \ln(e^X) = X$ is lognormally distributed. Assume we apply this intuition in the above formula. In that case, it is easy to see that we can model P_T as a lognormally distributed random variable

since $r_{0,T}$ is approximately normally distributed.

Volatility and Continuously Compounded Returns

Volatility measures the standard deviation of the continuously compounded returns on the underlying asset. Conventionally, it is usually annualized.

We calculate volatility using the historical series of continuously compounded returns. Another method is converting daily holding returns into continuously compounded daily returns and then calculating annualized volatility.

We base annualizing volatility on 250 trading days in a year, which is an estimate of the business days the financial markets operate. The formula we use for annualizing volatility is:

$$\sigma(r_{0,T}) = \sigma\sqrt{T}$$

For example, if the daily volatility is 0.05, then the annual volatility is:

$$\sigma(r_{0,T}) = 0.05 \times \sqrt{250} = 0.79$$

Example: Lognormal Distribution and Continuous Compounding

Jess Kasuku is analyzing the stock of ABC Company, which is listed on the London Stock Exchange under the ABC ticker symbol. Kasuku wants to understand how the stock's price changed during a particular week when significant developments in the global economy impacted the UK stock market. To do this, she calculates the stock's volatility for that week using the closing prices shown in Table 1.

Table 1: ABC Company Daily Closing Prices

| Day | Closing Price (GBP) |
|-----------|---------------------|
| Monday | 75 |
| Tuesday | 78 |
| Wednesday | 72 |
| Thursday | 70 |
| Friday | 68 |

Using the information in Table 1, calculate the annualized volatility of ABC Company's stock for

that week, assuming 250 trading days in a year.

Solution

Step 1: Calculate the continuously compounded daily returns for each day using the formula

$$\ln\left(\frac{\text{Ending Price}}{\text{Beginning Price}}\right):$$

$$r_1 = \ln\left(\frac{78}{75}\right) = 0.03922$$

$$r_2 = \ln\left(\frac{72}{78}\right) = -0.08004$$

$$r_3 = \ln\left(\frac{70}{72}\right) = -0.02817$$

$$r_4 = \ln\left(\frac{68}{70}\right) = -0.02899$$

Step 2: Calculate the mean of the continuously compounded daily returns:

$$\begin{aligned} \mu &= \frac{r_1 + r_2 + r_3 + r_4}{4} \\ &= \frac{0.03922 + (-0.08004) + (-0.02817) + (-0.02899)}{4} \\ &= -0.024495 \end{aligned}$$

Step 3: Calculate the variance of the continuously compounded daily returns:

$$\begin{aligned} \sigma^2 &= \frac{(r_1 - \mu)^2 + (r_2 - \mu)^2 + (r_3 - \mu)^2 + (r_4 - \mu)^2}{4} \\ &= \frac{[(0.03922 - (-0.24495))^2 + (-0.08004 - (-0.024495))^2 + (-0.02817 - (-0.024495))^2 + (-0.02899 - (-0.024495))^2]}{4} \\ &= \frac{0.007179}{4} = 0.001795 \end{aligned}$$

Step 4: Calculate the standard deviation of the continuously compounded daily returns:

$$\begin{aligned} \sigma &= \sqrt{\sigma^2} \\ &= \sqrt{0.001795} = 0.042363 \end{aligned}$$

Step 5: Annualize the volatility by multiplying the daily volatility by the square root of the

number of trading days in a year.

We know that:

$$\begin{aligned}\sigma(r_{0,T}) &= \sigma\sqrt{T} \\ \therefore \sigma_{\text{annualized}} &= \sigma_{\text{daily}} \times \sqrt{250} \\ &= 0.042363 \times \sqrt{250} \\ &= 0.6698 \approx 67\%\end{aligned}$$

So, the annualized volatility of ABC Company's stock for that week was 67.23 percent.

Question

Which of the following is true about lognormal distributions compared to normal distributions?

- A. They are skewed to the right.
- B. They can take on negative values.
- C. They are less suitable for describing asset prices than asset returns.

Solution

The correct answer is A.

Lognormal distributions are continuous probability distributions that only take positive values and are often skewed to the right.

B is incorrect because lognormal distributions only take on positive values.

C is incorrect because there is no evidence to suggest that lognormal distributions are less suitable for describing asset prices than asset returns.

LOS 6b: describe Monte Carlo simulation and explain how it can be used in investment applications

Monte Carlo simulations are about producing many random variables based on specific probability distributions. This helps in estimating the probability of various results.

We will give an example to illustrate Monte Carlo Simulation implementation.

Steps Involved in Project Appraisal

Imagine an investor who wants to predict the results of a 70% stock and 30% bond portfolio over 20 years. This is how we set up a Monte Carlo simulation:

Specifying the Simulation:

Step 1: Specify the quantity of interest in terms of underlying variables.

The quantity of interest here could be the final portfolio value after 20 years, denoted as V_{iT} . In this case, this is the final portfolio value at time T resulting from ith simulation trial.

The underlying variable is the return on the portfolio. The starting portfolio value is \$100,000, with 70% invested in stocks and 30% in bonds.

Step 2: Specify a time horizon.

Assume we're interested in yearly returns, so the time horizon is 20 years. Divide the calendar time into sub-periods. In this case, we will assume yearly returns so that the number of sub-periods is K = 20, and the time increment Δt is, therefore, one year.

Step 3: Specify the method for generating the data used in the simulation.

Here, we need to make distributional assumptions. We might assume that the annual portfolio return follows a normal distribution. Let's say we estimate an average return μ of 7% for stocks, 3% for bonds, a standard deviation σ of 15% for stocks, and 5% for bonds. We can model changes in the portfolio value using the formula below:

$$\begin{aligned}\Delta \text{Portfolio value} = & 0.7 * (\mu_{\text{stock}} \times \text{Prior portfolio value} \times \Delta t \\ & + \sigma_{\text{stock}} \times \text{Prior portfolio value} \times Z_k) \\ & + 0.3 * (\mu_{\text{bond}} \times \text{Prior portfolio value} \times \Delta t \\ & + \sigma_{\text{bond}} \times \text{Prior portfolio value} \times Z_k)\end{aligned}$$

Here, Z_k is a standard normal random variable representing the uncertainty in the portfolio return (risk factor). We can use a computer program to draw 20 random values of Z_k .

Running the Simulation Over a Given Number of Trials:

Step 4: Use the simulated values to produce portfolio values.

This step involves converting the standard normal random numbers (Z_k) generated in step 3 into yearly changes in portfolio value ($\Delta \text{Portfolio value}$) using our model from step 3. This gives us 20 observations of possible changes in portfolio value over the 20-year period. From these observations, we create a sequence of 20 portfolio values, starting with the initial value of \$100,000.

Step 5: Calculate the final portfolio value.

The average portfolio value at the end of 20 years (V_{iT}) is calculated by summing up the portfolio values at the end of each year and dividing by 20. We then calculate the present value (V_{i0}) of this average value by discounting it to the present using an appropriate interest rate. The subscript i in V_{iT} and V_{i0} indicates that these values are from the i th simulation trial. This completes one simulation trial.

Step 6: Repeat steps 4 and 5 over the required number of trials.

Finally, we repeat steps 4 and 5 multiple times, say, 1,000 times. We then calculate summary statistics, such as the mean, median, and percentiles of the distribution of V_{i0} values. These summary statistics provide a range of potential outcomes for the portfolio value after 20 years, helping the investor understand the risks and rewards of the investment strategy.

Major Applications of Monte Carlo Simulations

- It can also be used to value complex securities such as American or European options.

Limitations of Monte Carlo Simulations

- It only provides us with statistical estimates of results, not exact figures.
- It is fairly complex and can only be carried out using specially designed software that may be expensive.
- The complexity of the process may cause errors, leading to wrong results that can be potentially misleading.

Question

Which of the following is a correct statement about the use of Monte Carlo simulations in finance and investment?

- A. They provide exact valuations of call options.
- B. They estimate a portfolio's potential returns by simulating its performance.
- C. They assess how changes in assumptions, such as interest rates or market volatility, affect a financial model.

Solution

The correct answer is C.

Monte Carlo simulations can assess how changes in assumptions, such as interest rates or market volatility, affect a financial model. This allows analysts to understand the impact of these changes on the model's results.

A is incorrect because Monte Carlo simulations do not provide exact valuations of call options. Instead, they can estimate the value of these options by simulating their potential outcomes.

B is incorrect because while Monte Carlo simulations can estimate a portfolio's potential returns, they do not simply simulate its performance. Instead, they use probability distributions to model the uncertainty in the portfolio's returns.

LOS 6c: describe the use of bootstrap resampling in conducting a simulation based on observed data in investment applications

Resampling

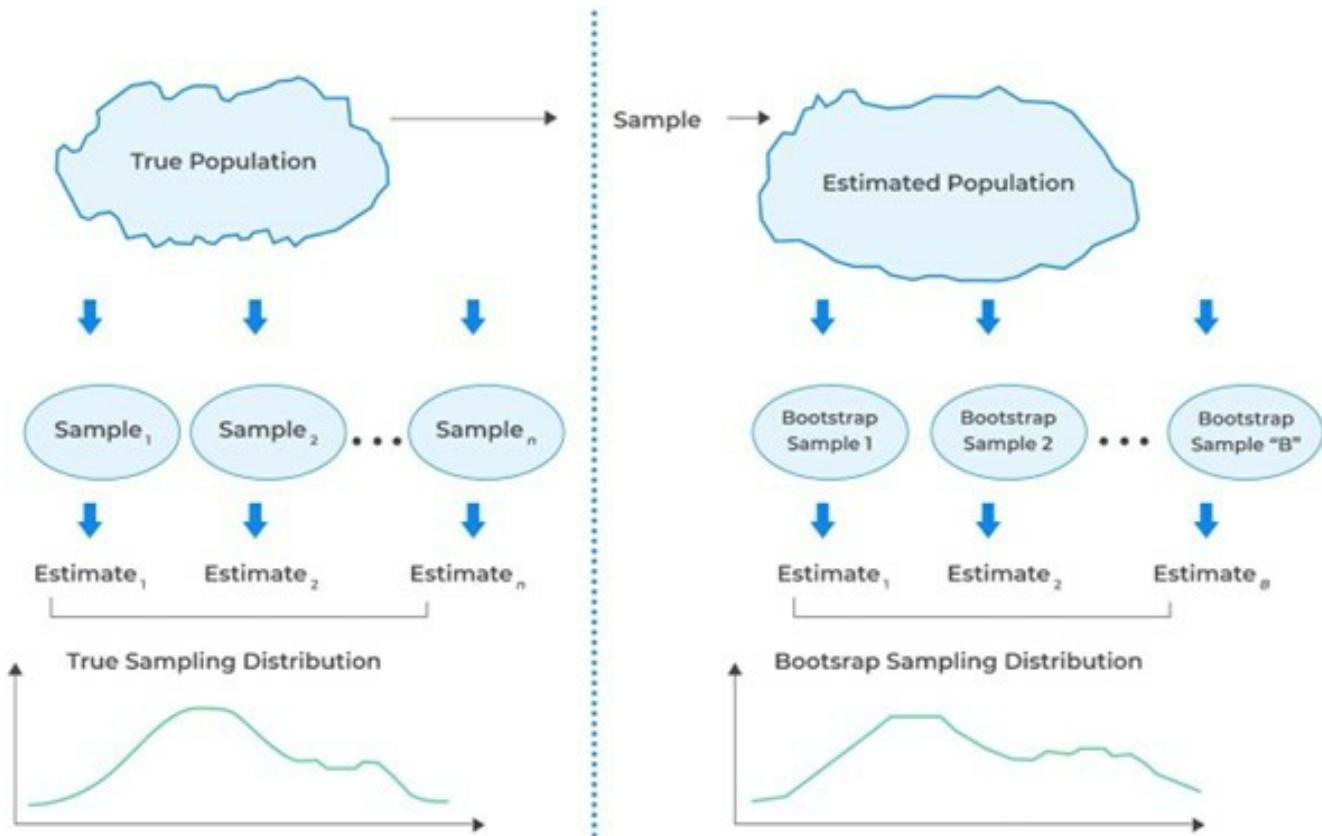
Resampling means repeatedly drawing samples from the original observed sample to make statistical inferences about population parameters. There are two common methods: Bootstrap and jackknife. Here, we'll focus on the Bootstrap method.

Bootstrap Resampling

Bootstrap resampling relies on computer simulations for statistical inferences, bypassing the need for conventional analytical formulas like z-statistics. The bootstrap technique is underpinned by a strategy that mirrors the random sampling process from a population to create a sampling distribution.



Bootstrap Resampling Method



Note that in bootstrapping, we do not have information about the population. Our only insight comes from a sample of size n drawn from this “unknown population.”

The core concept is that a random sample can effectively stand in for the entire population. So, we can mimic drawing samples from the population by repeatedly resampling from the initial sample. Essentially, the bootstrap method treats the initially obtained sample as a stand-in for the entire population.

Bootstrapping vs Monte Carlo Simulation

Both bootstrap and Monte Carlo simulation techniques lean heavily on the concept of repetitive sampling. Bootstrap considers the resampled dataset as a proxy for the true population and infers the population parameters such as mean, variance, skewness, and kurtosis from the statistical distribution of these samples.

On the other hand, Monte Carlo simulation is centered on the generation of random data with pre-determined statistical distribution of parameter values.

Simulation Using Bootstrapping

Simulation using bootstrapping is similar to Monte Carlo Simulation except for the source of random variables. In bootstrapping, the random variables are taken from a bootstrap sample instead of a probability distribution.

Consider the previous example:

Let's say an investor wants to understand the potential outcomes of investing in a portfolio with a 70-30 split between stocks and bonds over a 20-year period. Here's how a Monte Carlo simulation could be set up:

The simulation steps using the bootstrap sampling distribution are as follows:

Specifying the Simulation:

Step 1: Specify the quantity of interest in terms of underlying variables.

The quantity of interest here could be the final portfolio value after 20 years, denoted as V_{iT} . The underlying variable is the return on the portfolio. The starting portfolio value is \$100,000, with 70% invested in stocks and 30% in bonds.

Step 2: Specify a time horizon.

Assume we're interested in yearly returns, so the time horizon is 20 years. Divide the calendar time into sub-periods. In this case, we will assume yearly returns so that the number of subperiods is $K = 20$, and time increment Δt is, therefore, one year.

Step 3: Generate bootstrap samples from the empirical distribution of portfolio returns.

Here, we use the historical return data as our empirical distribution. Instead of assuming that the annual portfolio return follows a specific theoretical distribution, we will use the bootstrap procedure to draw the $K = 20$ yearly returns from the observed empirical distribution.

Running the Simulation Over a Given Number of Trials

Step 4: Use the bootstrap samples to produce portfolio values used to value the contingent claim.

This step involves using the bootstrap samples drawn in Step 3 to compute the yearly changes in portfolio value. From there, we create a sequence of 20 portfolio values, starting with the initial value of \$100,000.

Step 5: Calculate the final portfolio value:

The average portfolio value at the end of 20 years (V_{iT}) is calculated by summing up the portfolio values at the end of each year and dividing by 20. We then calculate the present value (V_{i0}) of this average value by discounting it to the present using an appropriate interest rate. The subscript i in V_{iT} and V_{i0} indicates that these values are from the i th bootstrap sample. This completes one bootstrap sample.

Step 6: Repeat steps 4 and 5 over the required number of trials.

Finally, we repeat steps 4 and 5 multiple times, say, 1,000 times. We then calculate summary statistics, such as the mean, median, and percentiles of the distribution of V_{i0} values. These summary statistics provide a range of potential outcomes for the portfolio value after 20 years, helping the investor understand the risks and rewards of the investment strategy based on the observed empirical distribution of returns.

Question

Which of the following statements is *most likely* accurate in relation to bootstrap analysis?

- A. Bootstrap analysis aims to deduce statistics about population parameters from a singular sample.
- B. Bootstrap analysis involves the repeated extraction of samples of equal size, with replacement, from the initial population.
- C. During bootstrap analysis, it is necessary for analysts to determine probability distributions for primary risk factors that govern the underlying random variables.

Solution

The correct answer is A.

The bootstrap analysis employs random sampling to generate an observed variable from a set of unknown population parameters. Although the actual distribution of the population is unknown to the analyst, the parameters of the population can be inferred through the sample produced via random sampling.

B is incorrect. In bootstrap analysis, the analyst repeatedly samples from the initial sample, not the entire population. Each resample has the same size as the original sample, and for each new draw, selected items go back into the sample.

C is incorrect. During bootstrap analysis, analysts simply utilize the empirical distribution of the observed underlying variables. In contrast, the analyst must establish probability distributions for the key risk factors that govern the underlying variables in a Monte Carlo simulation.