Once investors have decided that they want to own a company's stock, how do investors buy shares? There are two ways to buy a stock:

- **Primary Market:** When the firm decides that it wants to issue new shares of stock, you can decide to buy directly at the time of issuance. Purchasing a share from the issuer directly occurs at an **initial public offering** or **IPO**. This is usually difficult to do for small investors. In fact, when a company decides to go public and issue stocks via an IPO, such an IPO is underwritten by one or more investment banks or broker-dealers. They purchase the shares from the company and then try to sell those shares to investors. Broker-dealers try to sell large pieces, and thus, these shares usually end up in the hands of large institutional investors. This is not because the small investors are not important, but the way things happen is faster and more efficient from the point of view of the bankers and broker-dealers involved.
- **Secondary Market:** Once the shares of stock have been issued, they can easily change hands by trading on exchanges. If you wanted to buy existing shares, then you would have to convince an owner to sell their shares at a mutually agreed-upon price. Such convincing simply means buying at the owner's offer price on a stock exchange. Similarly, if you wanted to sell your shares, then you would have to find a buyer willing to pay a price you both agree upon. Indeed, there is an extensive amount of trading on some stocks that have been issued. From the firm's point of view, this is a good thing because the company's shares are seen as marketable and liquid (terms we will formally define later in this course).

As an investor, why do we want to purchase stock? From a corporation's perspective, why do they want to issue stock? Let's tackle that next.

Value a stock:

1.4 How to Value a Stock

We mentioned that one of the reasons behind investing in stocks is capital appreciation. To assess a stock's potential for capital appreciation, investors must have a reasonable estimate of the value of a stock or need to get such value from stock analysts. While quantitative analysts may not get particularly involved in this, some knowledge is nevertheless useful. There are several techniques that can be used and they are all more or less rooted in models of corporate finance. One of these techniques, the **Discounted Cash Flow model** (DCF), is very general and requires a deep knowledge of how the company works. This type of modeling is out of reach for someone who does not know the company very well. Other models like relative valuation or the **dividend discount models** (DDM) are much more accessible. In particular, the Gordon Growth Model is relatively simple yet reasonably powerful and does not require specialized knowledge of the business of a company.

1.4.1. Gordon Growth Model (GGM).

This special dividend discount model is based on the idea that the stock's present-day price is worth the sum of all its future dividends when discounted back to its present value. The price so calculated is called **fair value**. Then, if the current price of the stock is lower than its GGM value, the stock is undervalued and we should buy it. On the other hand, if the GGM value is below its current price, then we should sell the asset or maybe even short it. This model is based on the time value of money, which we are already familiar with and requires only 2 key inputs:

- **Estimating dividends and dividend growth rate**. Estimating the value of a company's dividends going forward is not easy in general, but we can make certain assumptions. For instance, we could extrapolate a trend in past dividend payments and use that to estimate future dividends. In particular, we can assume that the company has a fixed growth rate of dividends and that this flow of dividends continues unchanged forever (the technical term is "in perpetuity"). So we need the expected next closest dividend per share in time, which we denote by D_1 and the growth rate for dividends, which we call g.
- **Estimating a discount factor**. Shareholders who invest money in stocks are aware that they are making a risky investment, and as with all risks, they want to be compensated for it. The firm cost of equity capital is usually taken to represent what the market demands as compensation for owning the asset and assuming the underlying risk. This cost of equity/compensation rate will be denoted by k. Note that k has to be greater than g so that the denominator remains positive.

The formula for the GGM model is the following

$$\text{Value of stocks} := P_0 = \frac{D_1}{k-g}$$

D1: the expected next closest dividend per share in time
g: the growth rate for dividends
k: cost of equity/compensation rate

1.4.3. Multistage Gordon Growth Model with Linearly Declining Growth Rate (H model).

This is essentially the same model as the mGGM, but it allows for the growth rate to decline linearly from high growth to stable growth over a period of H years consisting of three stages so that the assumptions are now:

- Stage 1: High growth rate is g_H; stage 1 consists of only 1 period.
- Stage 2: The growth rate transitions from g_H to g_L over a period of H periods.
- Stage 3: From time $H+1$ on, the company grows in perpetuity at the stable growth rate of g_L.

The formula in this case actually allows for a closed-form solution:

$$P_0 = \frac{D_0(1+g_L)}{k-g_L} + \frac{D_0(H/2)(g_H-g_L)}{k-g_L} = \frac{D_0(1+g_L)+D_0(H/2)(g_H-g_L)}{k-g_L} = \frac{D_0}{k-g_L} \times [(1+g_L)+(H/2)(g_H-g_L)].$$

Example. *Suppose we are given the following information about company STU:*

- *current stock price = $34*
- *cost of equity = 9.5%*
- *last dividend paid $1.85*
- *growth rate (stage 1) = 7%*
- *growth rate (stage 3) = 3%*
- *years (stage 2) of linear decline = 10*

What is the fair value of company STU? Is it overvalued, undervalued or fairly valued?

Answer. Using the following spreadsheet, we find $P_0 =$ 35.01. $Hence, the shares of$ STU $ are mildly undervalued.

1.4.4. True Multistage H Model.

The previous model was important to illustrate the technicalities of the H model. Now, we extend it to the following:

- Stage 1: High growth rate is g_H; stage 1 consists of m periods.
- Stage 2: The growth rate transitions from g_H to g_L over a period of $H = n - m$ periods.
- Stage 3: From time $n+1$ on, the company grows in perpetuity at the stable growth rate of g_L.

We could provide a formula for this case too, but it is going to become cumbersome. Therefore, we will utilize a spreadsheet and show how to apply it via an example.

Measure the performance of stocks:

1.2 Logarithmic Returns

The second way to compute returns employs the following formula:

$$\text{logarithmic return} := \tilde{r}_{t_1,t_2} = \ln\left(\frac{S_{t_2}}{S_{t_1}}\right)$$

where $\ln(\cdot)$ is the natural logarithm function.

Note the use of the tilde symbol to differentiate the way the returns are computed in the two instances. The two formulae for percent return and log return tend to yield similar numbers, but they are not the same.

The formula can be easily extended to incorporate dividends, of course:

$$\text{logarithmic return} = \ln\left(\frac{S_{t_2} + D_u}{S_{t_1}}\right).$$

Example. *Suppose we bought a stock on January 2, 2024, paying $125. On February 28, 2024, the stock is valued at $153. Compute the arithmetic return and the logarithmic return.*

Answer. This is very simple:

$$\text{arithmetic return} = 153/125 - 1 = 22.40\%;$$

and

$$\text{logarithmic return} = \ln(153/125) = 20.21\%.$$

1.3 Logarithmic Returns or Arithmetic Returns?

The question of which of the two ways to compute returns is most appropriate comes up often. To answer this question, suppose that we invest an amount of S_0 dollars into an asset and we earn the following sequence of returns at time $t_1, t_2, \ldots, t_n$. In addition, we assume that $t_0 = 0$. Let's call these returns $r_1, r_2, \ldots, r_n$. Just to be sure that everyone is clear about the meaning, r_1 is the arithmetic or percent return between time 0 and time t_1, r_2 represents the percent return over the period between t_1 and t_2, and so on.

As investors, we are interested in the **cumulative return** between time 0 and time t_n. Then:

$$\text{cumulative return} := r_{0,t_n} = \frac{S_0 \times \prod_{i=1}^{n}(1 + r_i)}{S_0} - 1$$

and we observe that, if r_{0,t_n} represents the percent return between time 0 and t_n, then it must be:

$$r_{0,t_n} = \frac{S_{t_n}}{S_0} - 1 = \prod_{i=1}^{n}(1 + r_i) - 1.$$

Essentially, we are letting the returns compound continuously, i.e., we are reinvesting any gains between the time periods considered.

On the other hand, if we removed the profit in between periods so that at the beginning of each subperiod we always have an investment of S_0, then the cumulative return would be calculated as:

$$R_{0,t_n} = \frac{S_0 + S_0\left(\sum_{i=1}^{n} r_i\right)}{S_0} - 1 = \sum_{i=1}^{n} r_i.$$

Note the use of R instead of r to distinguish the two situations. R_{0,t_n} is the formula for total return without compounding and it is called the formula for total returns with **rebalancing**, meaning that if the return is positive, we take out the additional amount earned and spend it on something else (even use it to make another investment), and if the return is negative, we add money to our investment to bring it back to its original level of S_0. That is where the notion of re-balancing comes into play. Students have probably noticed the use of different symbols for total returns in the two instances as we are computing total returns differently.

6. Volatility Divided by Return: The Coefficient of Variation

We have discussed returns. We have discussed volatility. How can we use these to make investment decisions? Ideally, we would like to have a lot of return for as little volatility as possible. More expected return means higher profits. Less volatility means we are more likely to achieve those high profits. Lowering returns reduces our profits. Raising volatility means we increase the likelihood of being far below those profits to the point where they become losses!

So let's combine the two. First, we'll start by dividing one by the other. We'll start by dividing volatility by return. Volatility divided by return is known as the coefficient of variation. Volatility is the numerator and return is the denominator. This is a classical metric from statistics. In finance, when we compute the coefficient of variation, we are basically asking the following question: how much volatility does an investment encounter to earn a given amount of return? That is, you are effectively seeing how much volatility costs per unit of return. Think of it as an exchange rate. You have to spend x amount of volatility dollars to earn y amount of return units.

Suppose you use this ratio to compare the relative risk/return relationship of two investments, such as two stocks. Stock A has a 50% volatility per unit of return. Stock B has 25% volatility per unit of return. Which of these seems like a better investment? We've intentionally not provided the specific returns and volatilities, just this ratio. This is important because the returns of the two stocks can be very different, as well as their volatility. This ratio can be thought of as using a volatility currency to purchase a return currency. How much volatility do you have to spend to earn a given amount of return? For Stock A, you have to spend 50 volatility units. For Stock B, you only have to spend 25 volatility units. Relatively speaking, Stock B is a bargain. You do not have to spend, that is, encounter, as much volatility to earn the same amount of return. The lower the coefficient of variation, the more preferable the stock is because you don't expect to face as much volatility to earn the same amount of return that you expect for the other.

Note that this assumes you had good samples of the data, that these are representative of the future price changes, and that you did your calculations correctly. All these assumptions can be questioned and perhaps even validated.

The lower the coefficient of variation, the more preferable the stock is because you don't expect to face as much volatility to earn the same amount of return that you expect for the other.

7. Return Divided by Volatility: Sharpe Ratio

Suppose instead of dividing volatility by return, we change the order and divide return by volatility. This is the more classical way to approach the problem. In this case, now, we have return as the numerator and volatility as the denominator. Suppose we change the numerator to reflect the market premium:

$$\text{Market Premium} = \text{Stock Return–Risk-Free Rate}$$

The idea combines not only return and volatility but also the fixed-income market. The fixed-income market is important because it shows what any risk-adverse investor would earn by simply being in the most conservative investment: the risk-free rate. The risk-free rate is the rate earned simply for investing in a risk-free bond. Since stocks are risky, they should earn returns in excess of that return. If a stock were to earn less than the risk-free return, then a rational investor would forgo the volatility of the stock market and stick with risk-free investments, which protect the principal while still earning some return.

The numerator in our example is the market premium. So, suppose for stock C, the market premium and standard deviation are 7% and 10%, respectively. Suppose for stock D, the market premium and standard deviation are 8% and 12%, respectively. Which is the preferred investment?

$$\text{Stock D:} = \frac{\text{Market Premium}}{\text{Volatility}} = 8/12 = 0.67 \text{ or } 67\%.$$

In this case, stock C has a higher return per unit of volatility (risk) than stock D does. We would prefer the investment that has a higher amount of return per unit of risk. Now, this is the opposite of our coefficient of variation. When we express volatility units per return, we would like the lower number. When we express return units per volatility, we prefer the larger number. Here, the larger number means we expect to earn more return per unit of volatility.

In this course, we'll name this ratio (keep a *sharp* lookout for it) and use it extensively to compare investments, but for now, the important thing is that we've seen how return and volatility can be combined to compare investments. This can just as easily work if these securities are from different asset classes. Return and risk help encapsulate the key statistical properties of a financial asset's performance. In the next lesson, we'll dive deeper into the details of these statistics.

When we express return units per volatility, we prefer the larger number. Here, the larger number means we expect to earn more return per unit of volatility.

Model the performance of stocks:

A comparison between the distribution of closing prices and log returns can be seen in the following spreadsheet that plots the histogram of daily returns for the Apple stock and which you can easily manipulate (just be sure that you change the data interval in the charts if you redefine the time interval). You should notice that the histograms seem to confirm something we said in Lesson 3 of this module: if stock prices are log-normally distributed, returns are normally distributed.

Example. *Why is it not possible for stock prices to be normally distributed?*

Answer. A normal distribution allows for negative values and prices cannot become negative.

When we are reading the plot in Figure 1, there are a lot of questions we can ask. For example,

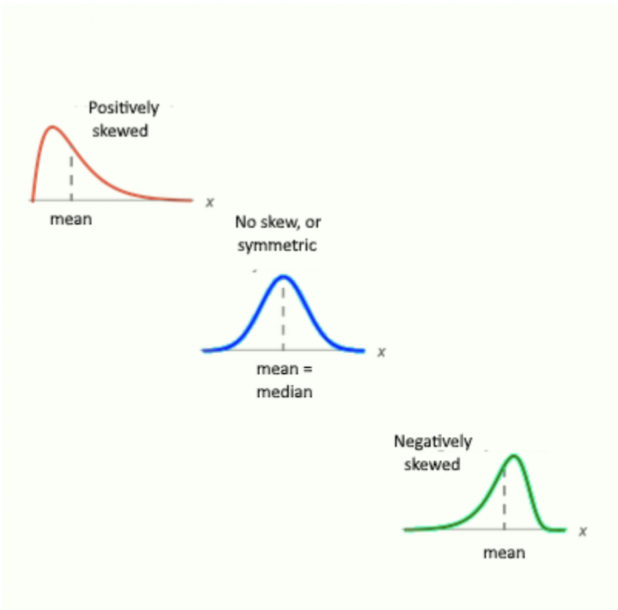
- How high does the series go?
- How low does the series go?
- Does it oscillate between these two points frequently?
- Does the series tend to rise or fall over time, creating a trend?
- Does there tend to be more fluctuation at one time than there is at another time?
- Does the series rise more abruptly than it falls?

if stock prices are log-normally distributed, returns are normally distributed.

We cannot merely assume that data comes from a normal distribution. In fact, much of science does assume this because of a property known as the **Central Limit Theorem**. What does the Central Limit Theorem say? If you have random variables coming from independent and identically distributed distributions, then the sum of those variables approaches a normal (or Gaussian) distribution. Is that the case for our financial series? In other words, will economic, financial, and individual decision makers combine their results in such a way that causes the cumulative distributions of their decisions to be independent and identically distributed? Stated practically, does a stock's behavior on Monday come from the same type of random distribution that it does on Tuesday? And again on Wednesday? Thursday? Friday? Unfortunately, this is shown not to be true. In other words, because of the dynamic nature of finance, because of the ability of market participants to change their behavior, we cannot merely assume that each trading day is both independent and identically distributed to other trading days. As a result, we cannot just assume that data is normally distributed. So, what helps is to assess whether the distribution is Gaussian by checking on its basic features.

Some distributions are not symmetric. In general, the most typical case of lack of symmetry occurs when we can have distribution with longer tails on the left (negatively skewed) or the right (positively skewed). If we determined that the distribution of returns of a certain investment is negatively skewed, this is loaded with practical consequences. In fact, that there returns that can be from zero on the left side, that means there are potentially large losses. If we insisted in using a symmetric distribution like a normal for example, we could end up grossly underestimating the loss potential for the investment because we are ignoring the more extreme values in the left tail. Indeed, this is a common trend in finance: in equity markets, there tend to be days of panic where there are large negative drops but not corresponding large positive gains, which would represent days of great optimism. Later in the program, we will study behavioral finance and learn that humans are hard-wired to remember painful events (e.g., large stock drops) more than pleasant events (e.g., large stock gains), so we tend to avoid painful events at greater lengths.

Figure 2: The 3 types of ideal possible skewness: positive (long right tail), symmetric, and negative (long left tail)



Skewness:

Now, finance is not physics. It is more like a combination of evolutionary biology, cognitive biases, and socially influenced decision making. That means that financial returns cannot always be explained as if they came from physical systems. They represent the millions of actions and interactions of both humans and computers running programs, whose values are in some way still decided by humans.

The result is that if returns are no longer symmetric, then we may not be able to assume that the normal distribution is sufficient. Symmetry is an easy way to check because the absence of symmetry is readily seen in a graph.

There are two ways to do this visually:

1. Draw a histogram. Are the bars equally populated on both sides?
2. Draw a density plot. Think of dropping a mound of sand on each data point. If the aggregate shape looks like a normal distribution, then the overall distribution can be thought of as normal. Departures from normality will be easy to spot.

If we don't want to rely on simple visuals, there are several statistics to measure skewness. An easy one is that known as Pearson's first skewness coefficient and defined as

$$\text{skewness} = 3 \times \frac{\text{mean} - \text{median}}{\text{standard deviation}}.$$

It is easy to see from this formula that when the distribution is symmetric (which one can show necessarily implies that the mean and the median of the distribution are the same), then skewness is zero. When we are not given a distribution, but instead are only given a certain number of observations from a population (almost always the case in applications), $x_1, x_2, \ldots, x_n$, then we can use the following formula:

$$\text{skewness} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\frac{1}{n-1}\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{3/2}}.$$

However, in this course, we are not going to do any math or use the formulae in any meaningful manner. What is important for students to remember are the notions of symmetric, positive and negative skewed, distributions and their possible implications for risk.

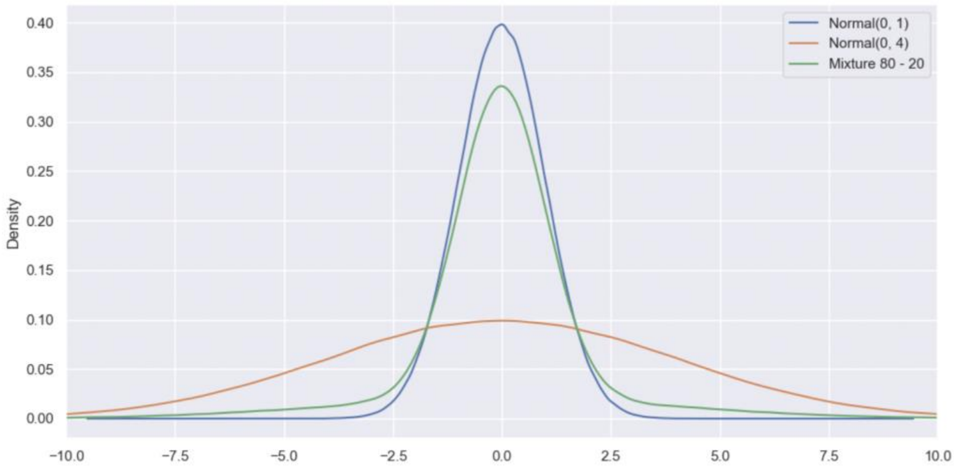
For a symmetric distribution, the standard deviation could be an acceptable measure of risk. For skewed distributions, in general, it is not.

excess kurtosis:

20% of the mixture is a normal distribution with a mean 0 and a standard deviation of 4.

When you compute the standard deviation of this mixture of normals, you get a value around 2.

Figure 3: Comparison of Gaussian (Normal) and Non-Gaussian Distributions



Is this value a reliable measure of volatility?

The answer cannot be understood by merely looking at the volatility. Think what the second moment of a distribution measures: the variance. It gives the squared variation around the mean. What we'd like to know is how accurate that squared variation is. We can get an idea of this from the excess kurtosis. Think of the excess kurtosis as a quality check on the variance. For a normal distribution, the variance is known with certainty (since the standard deviation is known). Therefore, there is no uncertainty about the variance, and there is no excess kurtosis. But suppose we take an empirical distribution and find that there is excess kurtosis. Then, we can likely conclude that this distribution did not come from a normal distribution. Indeed, when we measure the excess kurtosis of the individual normal distributions, they are no different from 0. But when we measure excess kurtosis of the mixture of normal distributions, we get a value around 7. We have a non-Gaussian distribution.

But when we measure excess kurtosis of the mixture of normal distributions, we get a value around 7. We have a non-Gaussian distribution.

6. Normal or Not?

If you wanted to introduce skewness, there are distributions that allow us to do so. We will examine these in the Financial Data course. These include:

- the F distribution,
- the skew normal distribution,
- the log-normal distribution, and many others.

There are formal tests for determining if a distribution has skewness. There are even robust measures that can handle skewness by focusing on either ranks instead of values or even by choosing to use only one side of returns (usually the negative ones) and ignoring positive returns (e.g., semi-variance).

Another important concept to consider in the shape of the distribution is kurtosis. In its essence, the notion of kurtosis investigates whether a distribution is peaked or flat. The relevance is that a peaked distribution exhibits lighter tails, while a flat distribution tends to have heavier tails and this is relevant when we want to assess the risk of certain investment.

For a sample of observations $x_1, x_2, \ldots, x_n$ from a population, one formula to compute the sample kurtosis is the following:

$$\text{kurtosis} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\frac{1}{n-1}[\sum_{i=1}^{n}(x_i - \bar{x})^2]^2}.$$

It can be shown that for a standard normal distribution, the kurtosis is approximately 3. So, instead of using the kurtosis, we use the excess kurtosis, which is just the kurtosis minus 3. In this way, if the excess kurtosis

- is positive, we know right away that the tails of the distribution are heavier than for a standard normal;
- is negative, then the distribution has thinner tails than a normal distribution.

There are distributions that exhibit kurtosis. One example would be the Student's t-distribution. The Student's t-distribution is symmetric (but also comes in an asymmetric version). Unlike the normal distribution, the Student's t has excess kurtosis. Instead of being described by a mean and standard deviation, the Student's t-distribution is described by a single number: the degrees of freedom. When the number of degrees of freedom is relatively small, for example 4, then there is indeed a high amount of excess kurtosis in the distribution. As the number of degrees of freedom increases, the distribution asymptotically approaches the normal.

In its essence, the notion of kurtosis investigates whether a distribution is peaked or flat. The relevance is that a peaked distribution exhibits lighter tails, while a flat distribution tends to have heavier tails and this is relevant when we want to assess the risk of certain investment.

It can be shown that for a standard normal distribution, the kurtosis is approximately 3. So, instead of using the kurtosis, we use the excess kurtosis, which is just the kurtosis minus 3. In this way, if the excess kurtosis.

When the number of degrees of freedom is relatively small, for example 4, then there is indeed a high amount of excess kurtosis in the distribution. As the number of degrees of freedom increases, the distribution asymptotically approaches the normal