Module 1: Returns and Risk

Dan the Quant

June 6, 2025

Abstract

This module provides a rigorous mathematical and conceptual foundation for understanding the notions of **returns** and **risk** in financial markets. It begins by comparing arithmetic and logarithmic returns, highlighting their respective advantages in performance measurement and statistical modeling. The text then introduces cumulative returns and the Compound Annual Growth Rate (CAGR) as key metrics for evaluating investment growth over time. On the risk side, the module presents essential quantitative tools such as volatility, covariance, and correlation, which characterize the variability and co-movement of asset returns. Additionally, it covers the Sharpe Ratio, enabling risk-adjusted performance evaluation, and Value at Risk (VaR), a probabilistic measure of potential losses under normal market conditions. Finally, Tracking Error is discussed as a metric of portfolio deviation from a benchmark, relevant for active and passive management strategies. Together, these concepts form the core analytical toolkit for investment analysis and portfolio management.

Contents

1 Returns

In the understanding of finance and economics, the fundamental measure of profitability is the **returns**, which must be interpreted as the earnings that an investor obtains on an investment during a certain time stamp. This metric is key to evaluating whether an investment project has been profitable over time, or not. Returns, then, reflect the relative gains or losses in relation to the amount initially invested and can be expressed both in absolute values (monetary units) and in percentages. Depending on the context, time horizon and nature of the financial instrument, the concept of returns may vary. Below, we will break down the concept of returns.

1.1 Arithmetic vs. Logarithmic

There are two common ways to calculate the returns of a stock. One is the **arithmetic return**, which is simply the percentage change in price over a given period (such as a day, a week, or a year):

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{1}$$

The other is the **logarithmic return**, which is the natural logarithm of the ratio between today's price and yesterday's price. It can also be expressed as the difference of the logarithms of the two prices:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln P_t - \ln P_{t-1} \tag{2}$$

The use of arithmetic and logarithmic returns varies depending on the context. However, logarithmic returns are generally preferred when building statistical or financial models with stock data. The differences between the two methods are summarized in Table 1.

Table 1: Comparison between Arithmetic and Logarithmic Returns

Criterion	**Arithmetic Return**	**Logarithmic Return**
Formula	$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$	$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$
Interpretation	Simple percentage change between two prices	Continuous compound growth rate
Approximation	Approximates log return when returns are small	Exact representation, includes compounding effects
Time Additivity	Not additive over time periods	Additive over time: $r_{1\to n} = \sum_{t=1}^{n} r_t$
Symmetry	Not symmetric: a loss of 50% is not offset by a 50% gain	More symmetric in modeling gains and losses
Statistical Modeling	Less suitable for continuous-time modeling	Preferred for stochastic modeling and time series analysis
Distributional Properties	More skewed, heavier tails	Closer to normality under many assumptions
Practical Use	Historical performance reporting, basic portfolio evaluation	Risk modeling, econometrics, and volatility estimation
Conversion	$R_t \approx r_t$ for small R_t	$r_t = \ln(1 + R_t)$

1.2 Cumulative Returns

The concept of **cumulative return** helps us measure the growth of an investment over several periods (such as months, years, or decades), considering both capital gains and the returns we have been able to generate in the form of dividends or interest. The cumulative return can be used as an indicator that reflects the impact of the reinvestment of the returns generated over time. This rate is defined as the product of all the returns obtained in an investment strategy over several periods; It may be because this strategy is characterized by buying and selling a certain financial asset according to the fluctuations of its price:

$$CR_t = \prod_{t=1}^{T}(1 - R_t) \tag{3}$$

When using logarithmic returns, you can take advantage of their additive property and express cumulative returns as a sum. But we have to be careful because the result is going to be the logarithm of the returns:

$$\ln CR_t = \sum_{t=1}^{T} \ln(1 + R_t) = \sum_{t=1}^{T} r_t \tag{4}$$

Cumulative returns are useful for evaluating long-term investments, especially when returns are reinvested, such as dividends or interest. This calculation reflects true investment growth, making it essential for comparing different investment opportunities and measuring performance over time, as seen with retirement funds or savings. It is particularly common when analyzing investment strategies that involve active trading rather than simple buy-and-hold approaches.

1.3 Compound Annual Growth Rate (CAGR)

To standardize the analysis, we can use the annualized rate of return, also known as the **compound annual growth rate (CAGR)**. This metric represents the average annual growth of an investment over a specific time period, assuming that profits are reinvested rather than simply holding the assets. The CAGR is calculated as the nth root of the total return over the investment's lifespan:

$$CAGR_T = \sqrt[n]{\frac{P_T}{P_0}} - 1 \tag{5}$$

Portfolio managers are very familiar with using this method to calculate average returns for long-short strategies, and it is common to apply it when comparing a portfolio's performance to a benchmark. Given that an investment strategy involves different buy and sell moments, the annualized return, or CAGR, is calculated as the nth root of the cumulative return described above:

$$CAGR_T = \sqrt[n]{\prod_{t=1}^{T}(1 - R_t)} - 1 \tag{6}$$

If we use logarithmic returns, we should apply the exponential function when converting back to a cumulative return. In this case, the compound annual growth rate (CAGR) is simply the exponential of the average logarithmic return, minus one:

$$CAGR_T = e^{\bar{r}} - 1 \tag{7}$$

$$\bar{r} = \frac{1}{n}\sum_{t=1}^{T} r_t \tag{8}$$

2 Risk Metrics

Measuring risk is crucial for understanding the degree of uncertainty and potential fluctuations in an investment's returns. As with most things, there are various mathematical and statistical methods to assess and quantify the risk associated with financial decisions. Quantifying risk leads to more informed decision-making, ultimately helping to maximize returns for any given level of risk. However, this is something we will explore in more detail later.

2.1 Volatility and Standard Deviation

Volatility is understood as the standard deviation, which is the average variability of the returns of an investment. It is one of the most popular ways to measure risk, as it reflects how dispersed the returns of a financial asset are around its average (or expected return). Consequently, we could assume that the greater the volatility, the greater the risk, since we understand that the greater the probability that returns will 'move away' from the expected returns. Mathematically it can be expressed as:

$$\sigma = \sqrt{\frac{1}{n}\sum_{t=1}^{T}(r_t - \mu)^2} \tag{9}$$

Here, σ represents the standard deviation of the returns, n is the sample size, which corresponds to the time period of the analysis, and, as we saw previously, μ represents the expected returns of our assets. It's important to note that the greater the standard deviation, the higher the probability of a significant difference between actual returns and expected returns (the mean).

2.2 Covariance

The **covariance** helps us to understand the degree of joint variation of two variables with respect to their means. In other words, it helps us to know the average dispersion of one variable compared to another. When the covariance is positive, we can understand that both variables move in the same direction, while, if it is negative, we can interpret that both variables fluctuate in opposite directions. Suppose that there are two stocks i and j:

$$\gamma_{ij} = \frac{1}{n-1}\sum_{t=1}^{T}(r_{it} - \mu_i)(r_{jt} - \mu_j) \tag{10}$$

If two stocks have positive covariance ($\gamma > 0$), it means that there is a higher probability that when the returns of one stock rise or fall, the returns of the other tend to move in the same direction. However, this does not imply a causal relationship; rather, it serves as an indication of how the two stocks are likely to move together over time.

2.3 Correlation

Another popular measure for measuring risk for a diversified portfolio is the **correlation** coefficient. Statistically speaking, this measure captures the average ratio in the movement of two financial assets. If these assets are highly correlated, they will move in a similar way (negatively or positively depending on the sign) which can initially reduce the benefits of diversifying a portfolio. The mathematical formula for calculating the Pearson Correlation Coefficient is the ratio between the covariance of both assets divided by their standard deviations:

$$\rho ij = \frac{\gamma_{ij}}{\sigma_i \sigma_j} \tag{11}$$

This correlation coefficient can vary between -1 and 1, which implies that a correlation close to -1 would indicate that both assets have returns that move in opposite directions almost identically. On the other hand, a coefficient close to one would tell us that both assets have returns that move identically, which is not very beneficial for the diversification of an investment portfolio. Finally, a coefficient of 0 would suggest a zero relationship between the two financial assets.

2.4 Sharpe Ratio

The **Sharpe Ratio** is a popular metric used to evaluate the risk-adjusted return on an investment. It measures how much additional return can be obtained for each unit of risk assumed. A high value of the Sharpe Ratio indicates that the investor is obtaining a higher return given a level of risk assumed. This metric turns out to be convenient and very useful when evaluating and comparing investments against each other.

$$S_P = \frac{\mu_P - r_f}{\sigma_P} \tag{12}$$

The denominator of the formula represents the volatility (σ_P) experienced by the portfolio as a measure of investment risk. On the other hand, the numerator represents the difference between the expected return of said portfolio (μ_P) and the risk-free rate; this differential is known as the market risk premium (r_f), as it explains the additional return that 'compensates' an investor for assuming a certain risk from an investment. By adjusting for risk, we can make better decisions about which investment is best for us. Furthermore, we will explore how investors select from equally efficient portfolios based on their risk aversion and risk tolerance.

- **Positive and high Sharpe ratio**: A high value indicates that the portfolio is generating a relatively high return relative to the risk assumed. For example, if a portfolio has a Sharpe Ratio of 1.5, it means that for each unit of risk, the investor gets 1.5 units of additional return.
- **Negative or low Sharpe Ratio**: If the Sharpe Ratio is close to zero or negative, it indicates that the portfolio's performance is poor compared to the risk assumed. In extreme cases, it might suggest that the investor would be better off placing their money in a risk-free asset, rather than taking on unnecessary risk in a risky-portfolio.

2.5 Value at Risk

The Value at Risk (VaR) is a statistical measure commonly used to quantify the minimum potential loss of an investment with a given confidence level. This metric is useful for assessing the risk of making losses under normal market conditions. VaR is expressed as a loss figure and a confidence level. For example, if a portfolio has a daily VaR of \$1 million with a confidence level of 95%, it means that there is a 95% chance that the loss in one day will not exceed \$1 million. VaR is defined by its three main components:

- **Confidence level**: For example, 95% or 99%, indicating the likelihood that losses will NOT exceed a certain threshold. By consensus we usually use a confidence level of 95%, however, being more rigorous we could even extend it to 99%.
- **Time Horizon**: The time stamp over which the loss is measured can vary, such as a day, a week, a month, or even a year. Note that business days are typically used to measure this time frame. Therefore, a week consists of 5 business days, a month has 21 business days, and a year typically has 252 business days.
- **Estimated Loss Value**: This is the dollar amount (or loss rate) that is expected not to be exceeded in the given time horizon. Depending on our necessities, we could use an absolute estimated loss, like a million dollars loss, or a relative measure, like -10% of returns.

Formally, the formula for calculating the VaR of an investment is mathematically described as follows:

$$VaR_{\alpha} = -Percentile_{\alpha}\left(r\right) \tag{13}$$

The Value at Risk formula indicates a negative sign in its component on the right side, which implies that we are using the left tail of the distribution of the returns on our investment. If you calculate a VaR at 95% confidence with a one-day horizon and get a VaR of $10,000, it means that there is a 95% chance of not losing more than $10,000 in a single day. In other words, there is a 5% chance that losses will exceed that value over the same time horizon.
If we assume that the returns of a stock follow a normal distribution, we could infer that they would have a mean of zero, a constant variance of one, and would also be symmetrical and mesokurtic. Thus, the VaR obtained for a significance level of 5% (or a confidence level of 95%) would be -1.645, as we are considering the left tail. This would then mean that, in 5% of the cases, the losses on our investment would represent at least 1,645 standard deviations.
The previously exemplified form is known as **Parametric VaR**, since this method assumes that the returns on our investment strictly follow a normal distribution with known mean and variance. Therefore, we can use the first and second moments of the distribution (mean and variance) to estimate the VaR.

$$VaR_{\alpha} = Z_{\alpha}\sigma_P \tag{14}$$

2.6 Tracking Error

Tracking Error is another risk measure commonly used in portfolio management because it helps us measure the difference between the returns of an investment portfolio and a benchmark index. This metric helps us understand how well or poorly a portfolio replicates its benchmark index and with it what decisions we could make in this regard. We can calculate the Tracking Error as the square root of the average of the squared differences between our portfolio and its benchmark:

$$TE_T = \sqrt{\frac{1}{n-1}\sum_{t=1}^{T}\left(r_{P,t} - r_{M,t}\right)^2} \tag{15}$$

Many also define the Tracking Error as the standard deviation of the differences between the returns of a portfolio and its reference, and computationally, it is how much software calculates it. The value of the tracking error can range from 0 to infinity. The lower, indicates that the behavior of the portfolio mimics that of the reference index, which may be suitable for indexed portfolios. On the other hand, the higher this is, it could give us indications that the portfolio manager makes more active decisions.