Financial Data, Returns, and Total Returns

Financial data is often quoted as prices. Prices are non-stationary, meaning our assumptions for modeling do not hold. We have to transform prices into a series that is stationary and suitable for modeling.

Fundamentally, today's price is some function of yesterday's price, plus whatever new information arrived today. Prices are generally assumed to be positive.

3 classical ways to view this:

1. Classical Brownian Motion

$$P_t = P_{t-1} + r_t$$

2. Arithmetic Return System

$$P_t = P_{t-1}\left(1 + r_t\right)$$

3. Log Return or Geometric Brownian Motion

$$P_t = P_{t-1}e^{r_t}$$

Classical Brownian Motion is the least used of the 3. There is not a good way to assure prices are greater than 0. However this is used in commodity applications where we are often concerned with the price between two physical locations (called the basis). Basis can and often does change sign.

Arithmetic returns are often used in portfolio management. If I have a portfolio of assets $X$with returns, r, and I hold them with weight w, then the total return on my portfolio is simply the dot product of r and w.

Let h_i be the absolute holding of asset i in portfolio,
$w_{i,t}$ is the weight of asset i at time t,
PV_t is the portfolio Value at time t,
$P_{i,t}$ is the price of asset i as time t
R_t is the portfolio return at time t
$r_{i,t}$ is the asset i return at time t

The value of the portfolio at time, t is:

$$PV_t = \sum_{i=1}^{n} h_i P_{i,t} \quad and \quad w_{i,t} = \frac{h_i P_{i,t}}{PV_t} \Rightarrow \sum_{i=1}^{n} w_{i,t} = 1$$

$$PV_{t+t} = \sum_{i=1}^{n} h_i P_{i,t}\left(1 + r_{i,t}\right) = PV_t\left(1 + R_t\right)$$

$$PV_t + \sum_{i=1}^{n} h_i P_{i,t} r_{i,t} = PV_t + PV_t R_t \Rightarrow \sum_{i=1}^{n} w_{i,t} r_{i,t} = R_t$$

Prices based on arithmetic returns can still be negative if the lower bound of the return is < -1. Often we assume normal returns, but with a low enough volatility that this boundary rarely, if ever, comes into play.

You will see arithmetic returns used in the literature looking at the cross section of returns where the additive property inside the same time period makes the math easy.

Geometric Brownian Motion is often used when looking at returns through time. Here returns are additive with time

$$P_{t+1} = P_t e^{r_t} \Rightarrow P_t = P_0 e^{\sum_{t=0}^{T} r_t}$$

If we assume $r_t \sim N\left(\mu, \sigma^2\right)$ then $P_{t+t} \sim LN\left(\mu + ln\left(P_t\right), \sigma^2\right)$

$$ln\left(P_{t+1}\right) = ln\left(P_t\right) + r_t$$

$$ln\left(P_{t+1}\right) \sim N\left(\mu + ln\left(P_t\right), \sigma^2\right) \Rightarrow P_{t+t} \sim LN\left(\mu + ln\left(P_t\right), \sigma^2\right)$$

In practice, assets distribute cash or undergo changes that make the underlying price time series not indicative of total return. Some common occurrences

1. Cash distributions
    a. Dividends
    b. Coupon Payments
2. Stock Splits
3. Spin outs

This is addressed by creating a total return price series. It is assumed that any proceeds are reinvested back into the security.

The method for doing this is straight forward:

1. Calculate a price return (usually arithmetic) at each time period.
2. In the time periods where events happen adjust the return so that it reflects the actual return an investor would have experienced.
3. Starting with today's price, use the returns to calculate the price series back in time.

If you are maintaining a database of these total price returns, each day you calculate the total return for the security and if there was an event, update the previous prices in the series. Alternatively you can just maintain the total return values.

Most sources of data will provide either a price series, a total return price series, or both. When pulling data from Yahoo! Finance, the "ADJ CLOSE" column is the total return price series.

Value at Risk (VaR)
Risk is the possibility of financial loss. It is quantified in numerous ways.

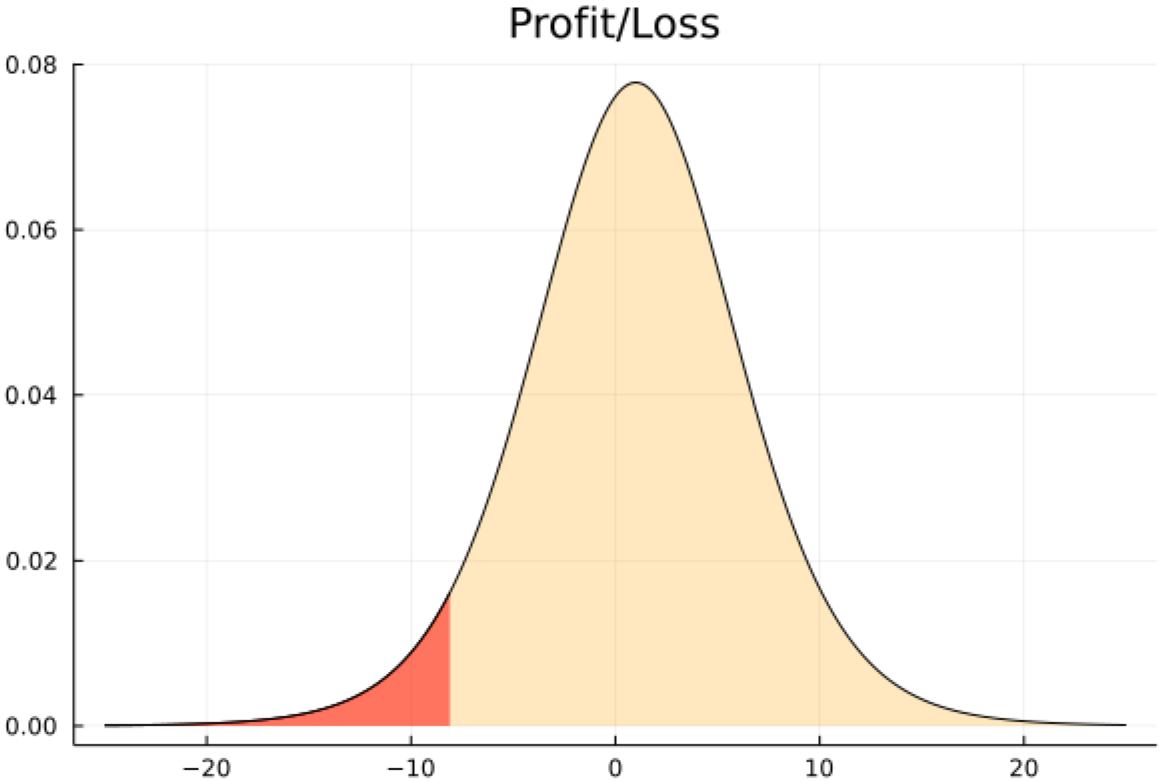
In portfolio management, we often see the standard deviation of returns as the risk measure. We will be using this when we talk about portfolio management in a few weeks.

Value at Risk, or VaR, is a widely used metric. It attempts to answer, under normal market conditions, what is the minimum we can expect to lose, on a α% bad day.

VaR is not the maximum loss. It is the minimum loss on a normal α% "normal" bad day. Most models look at recent history to calibrate market movement. Regimes can shift, and what was normal yesterday is abnormal today. Losses can in a "not normal" market can be an order of magnitude larger than VaR.

Said another way, markets are hard to model and extreme events happen with a higher probability than would be assumed from most modeling.

Mathematically, VaR is the α percentile on the profit and loss distribution.



We negate the value of VaR so that a positive number is a loss and a negative number is a gain.

$$VaR_{\alpha}(x) \;=\; -F_{x}^{-1}(\alpha)$$

VaR is reported in 2 ways in practice. It is always helpful to ask which is intended.

1. Absolute value on the profit/loss distribution. In the graph above, VaR = 8.
2. The relative difference from the mean expected value. The mean of the graph above is about 1, so VaR = 9.

In risk management, we mostly care about the loss, so the first definition is sufficient.

If we are comparing 2 portfolios and we are interested in both their expected return (profit) as well as their level of risk (VaR), then the 2nd definition is generally used.

Criticisms of VaR

The largest criticism of VaR has to do with the fact that models have a hard time describing price movements in Finance. VaR is the minimum loss under normal markets. Managing risk solely to VaR leads investors into a false sense of security.

VaR is not sub additive when the distribution is non-elliptical.

A subadditive measure is

$$q(P1)\ +\ q(P2)\ \geq\ q(P1\ +\ P2)$$

Standard deviation (and variance) are subadditive. Distributions are, subadditive. However, not all distribution percentiles are subadditive, meaning that VaR is not.

Why does that matter? Risk should be subadditive. It is hard to come up with a case where investing in 1 asset makes the 2nd asset more risky. Diversification, at worst, has no effect when things are perfectly correlated.

Example - Standard Deviation is subadditive. Imagine you have 2 assets that are independent. They each have a standard deviation of 1 and we assume multivariate normality. We hold the assets in equal weights.

$$q(0.5\ *\ P1\ +\ 0.5\ *\ P2)\ =\ \sqrt{0.5^2\ *\ 1\ +\ 0.5^2\ *\ 1} = 0.5\sqrt{2} =.707$$

$$q(0.5\ *\ P1)\ =\ q(0.5\ *\ P2)\ =\ 0.5$$

$$q(0.5\ *\ P1)\ +\ q(0.5\ *\ P2)\ =\ 1\ >\ 0.707$$

If we assume multivariate normality, then VaR is just a multiple of standard deviation

$$\alpha\ =\ 0.05\ \Rightarrow VaR\ \approx\ 1.65\sigma_P$$

Break the normality assumption – Imagine 2 bonds that are independent. Each costs \$90 and will pay \$100 in the next time period, if they do not default. Each has a 4% chance of default. You have \$180, you can buy 2 of either bond or 1 of both.

If we let $\alpha\ =\ 0.05$, then in the 5% worst case scenario, the each bond will not default – default happens in the 4% worst case scenario.

If I buy 2 of a single bond then, VaR = -\$20. It's negative because the 5% loss is still a profit.

If I buy 1 of each bond, then the probability of one of the 2 defaulting is

$$P(nDefault\ \geq\ 1)\ =\left(1\ -\ (1\ -\ 0.04)^2\right) = 0.078$$

In the 5% worst case, we will see 1 bond default.

VaR = 80 = (-90+10)

$$q(P1) = q(P2) = -10, \quad q(p1 + P2) = 80$$

$$q(P1) + q(P2) \geq q(P1 + P2)$$

Does not hold.

This makes VaR non-coherent as a risk measure.

But we still use it.

Delta Normal VaR

Sometimes called parametric VaR, assumes payoffs are linear and returns are distributed multivariate normal.

The assumptions make this a closed form solution.

$$\alpha = 0.05 \Rightarrow VaR_{ret} \approx 1.65\sigma_P$$

Where VaR is expressed as a return. You transform that into a price to get a $ value.

Assuming arithmetic returns

$$VaR_{\$} = -\left(P_0 - P_0\left(1 + \mu - VaR_{ret}\right)\right) = P_0 VaR_{ret} - P_0\mu$$

Assuming geometric returns

$$VaR_{\$} = -\left(P_0 - P_0 e^{\mu - VaR_{ret}}\right)$$

These numbers will be close but not exact. It is important to understand how your returns were calculated and use the proper version.

Most times we assume $\mu = 0$. We will assume that from here out.

Break down change in portfolio value as

$$\Delta PV = \sum_{i=1}^{n} d_i \Delta P_i$$

Assuming arithmetic returns and the Δ function as the return, then

$$R = \sum_{i=1}^{n} d_i r_i$$

We assume arithmetic returns because they are summable across the portfolio.

What's left is to find d_i. If we think of the equation above as Taylor Polynomial, then d_i is just the first derivative of portfolio value with respect to price i.

We don't hold prices, we hold assets (also called positions). So by the chain rule we need 2 derivatives.

$$\frac{dR}{dr_i} = \frac{dR}{dRA_i} \frac{dRA_i}{dr_i}$$

The derivative of the portfolio return, R, wrt the return on the asset, RA_i, is easy, that is the weight, w_i, from before (assuming arithmetic returns). If I hold 50% of the portfolio in Asset 1 and asset 1 returns 5%, then the change in the portfolio value is 2.5%.

The derivative of the asset return wrt the price return requires us to have a pricing function. Calculate the first derivative of the asset with respect to the current price.

$$\Delta A_i \approx \delta \, \Delta P_i$$

$$\delta = \frac{dA_i}{dP_i}$$

The change in the price in the asset is approximately the derivative of the asset with respect to price times the change in price divided by the asset price..

$$r_i = \frac{\Delta P_i}{P_i},\ RA_i = \frac{\Delta A_i}{A_i} \Rightarrow RA_i \approx \frac{\delta P_i r_i}{A_i}$$

$$\frac{dRA_i}{dr_i} = \frac{\delta P_i}{A_i}$$

Linear instruments like stocks are easy. The pricing function is just the price of the stock. The derivative is 1. The price of the asset and the price of the underlying are the same. A 1% change in value of the price causes a 1% change in the value of the position.

Derivatives are additive. If we have m assets that are affected by price i

$$\frac{dR}{dr_i} = \sum_{j=1}^{m} w_j \frac{dRA_j}{dr_i} = P_i \sum_{j=1}^{m} \frac{w_j \delta_j}{A_j}$$

If you recognize that $w_j = \frac{h_j A_j}{PV}$, where h_j is the units of the asset held, we can also express this in terms of holdings instead of weights. Depending on how you set up the problem, use the one that makes the most sense.

$$\frac{dR}{dr_i} = \frac{P_i}{PV} \sum_{j=1}^{m} h_j \delta_j$$

We now have a gradient, ∇R, of the portfolio return with respect to our underlying price returns. With the assumption of multivariate normality, it follows:

$$R \sim N\left(\nabla R^T \mu, \nabla R^T \Sigma \, \nabla R \right)$$

We assumed that $\mu_i = 0 \; \forall \; i$, so VaR is simply

$$VaR(\alpha) = - PV * F_X^{-1}(\alpha) * \sqrt{\nabla R^T \Sigma \, \nabla R}$$

Where $F_X^{-1}(\alpha)$ is the quantile function for the standard normal.

There are 2 big issues with the Delta Normal VaR

1. Assumption of Normality of returns. See the GFC example in the code with this lecture. The largest 1 day return during the financial crisis was so large that it should not have happened in the lifetime of the universe.
2. The linear assumption of portfolio value with returns. Options, and bonds are common portfolio holdings and their return profiles are not linear wrt underlying prices. The sign of the second derivative will decide if we over or under estimate VaR using this method.

Normal Monte Carlo VaR
The second issue with Delta Normal VaR is that asset pricing functions can be, and often are, non-linear. This makes the return distribution of the asset not well defined, meaning we need to estimate it with a Monte Carlo Simulation.

We are still assuming returns are multivariate normal.

1. Calculate Current Portfolio Value
2. Simulate N draws from the assumed distribution.
3. Calculate new prices from the returns (using chosen methodology)
4. Price each asset for all N draws.
5. Calculate portfolio Value for each N Draws
6. Using the simulated portfolio values, find the α% of the distribution.
7. Calculate VaR.

We find the α% by sorting the portfolio values and picking the $N * \alpha$ value from the vector. For a large number of draws, this is sufficient. If the number of draws is small, then using something like a Kernel Density Estimate or another smoothing technique might provide better results.

See https://stats.stackexchange.com/questions/341023/quantile-of-kernel-density-estimator for an example of calculating a quantile from a KDE.

Historical VaR
Both issues with Delta Normal VaR, returns are not normal and asset prices are not always linear, are sometimes addressed with Historical VaR.

This looks much like Monte Carlo VaR, but instead of using a model for returns, we simply assume observed past returns mimic the distribution. Historical VaR is a non-parametric statistic – we make no distributional assumptions on returns.

Historical VaR works like MC VaR, we do not simulate returns, we sample them from history.

1. Calculate Current Portfolio Value
2. Simulate N draws from history with replacement.
3. Calculate new prices from the returns (using chosen methodology)
4. Price each asset for all N draws.
5. Calculate portfolio Value for each N Draws
6. Using the simulated portfolio values, find the α% of the distribution.
7. Calculate VaR.

Historical VaR still suffers from recency bias. Often there is not enough history to show all regimes, so a limited sample is selected.

Using a KDE to smooth the VaR estimation is highly recommended.

We can add weights to historical values, much like we added weights for the Exponentially Weighted Covariance estimate. Without weights each day has an equal chance of being selected. Weights alter that probability.

Assume you have m days of history and each day, i, is given weight w_i

$$\sum_{i=1}^{m} w_i = 1$$

Calculate the cumulative weight vector, cw, as

$$cw_i = \sum_{j=1}^{i} w_j \; \forall\, i \in 1:m$$

The algorithm for selection a day, i, is

$$r = random\, uniform(0, 1)$$

$$max(i)\; where\; r \le cw_i$$

Time Periods Larger than 1

Sometimes we are interested in holding periods greater than a single day. This can be an issue as the data required to estimate multi-day ahead dynamics are more than single day ahead.

Often you see the "$\sqrt{t}$" rule. This comes from an assumption that returns are summable across time (geometric returns).

$$R_T = \sum_{t=1}^{T} R_t \;,\; R_t \sim N\left(\mu,\, \sigma^2\right) iid \Rightarrow R_T \sim N\left(T\mu,\, T\sigma^2\right)$$

(assuming $\mu = 0$)

$$VaR_t(\alpha) = -\, PV * F_X^{-1}(\alpha) * \sigma$$

$$VaR_T(\alpha) = -\, PV * F_X^{-1}(\alpha) * \sqrt{T}\sigma = \sqrt{T} * VaR_t(\alpha)$$

This estimation is often good enough. The options for each type of VaR other than the "$\sqrt{t}$" rule

1. Delta Normal – none. Use the "$\sqrt{t}$" rule. There is no other closed form solution and you are often using this approach when you need something that is quick to estimate.
2. Monte Carlo Normal – Two options
    a. If you are using arithmetic returns – simulate T*N random normals instead of N. Compound your returns across time, and then price the portfolio.
    b. If you are using geometric returns – Geometric returns are summable across time, and we have assumed normality, so we know the T period ahead distribution. Scale the covariance matrix by T (and the mean vector if you are using one). That is, multiply all elements of the covariance matrix by T. Simulate N values (which now have properly scaled variance), convert to prices, and price the portfolio.
3. Historical Simulation – Two options
    a. If you have a lot of historical data, segregate the data into M non-overlapping time periods. Sample from those M periods.
    b. If you do not have a lot of historical data, segregate the data into M overlapping time periods. This can cause a bias to your estimate. The samples are no longer independent.