Introductory Backtesting Notes for Quantitative Trading Strategies

Useful Metrics and Common Pitfalls

Leo Wong (QFIN & COSC, HKUST)

December, 2019

Abstract

This note is compiled for COMP4971C in Fall 2019 to assist the research and implementation of quantitative trading strategies. This note includes suggested evaluation metrics and common pitfalls for beginners in quantitative trading and backtest.

It is common nowadays to use mathematical and statistical models to evalute financial securities and take corresponding actions. However, attention to details is one of the crucial factors to make a strategy successful in reality, when various academic assumptions do not hold. Having proper assumptions and configuration for backtest may cover more situations and provide more robust evaluation of trading strategies. Thus, it enhances the quality of trading strategies as the backtest provides more realistic results and more accurate approximations.

"All models are wrong, but some are useful", Box 1976.

Contents

1 Introduction

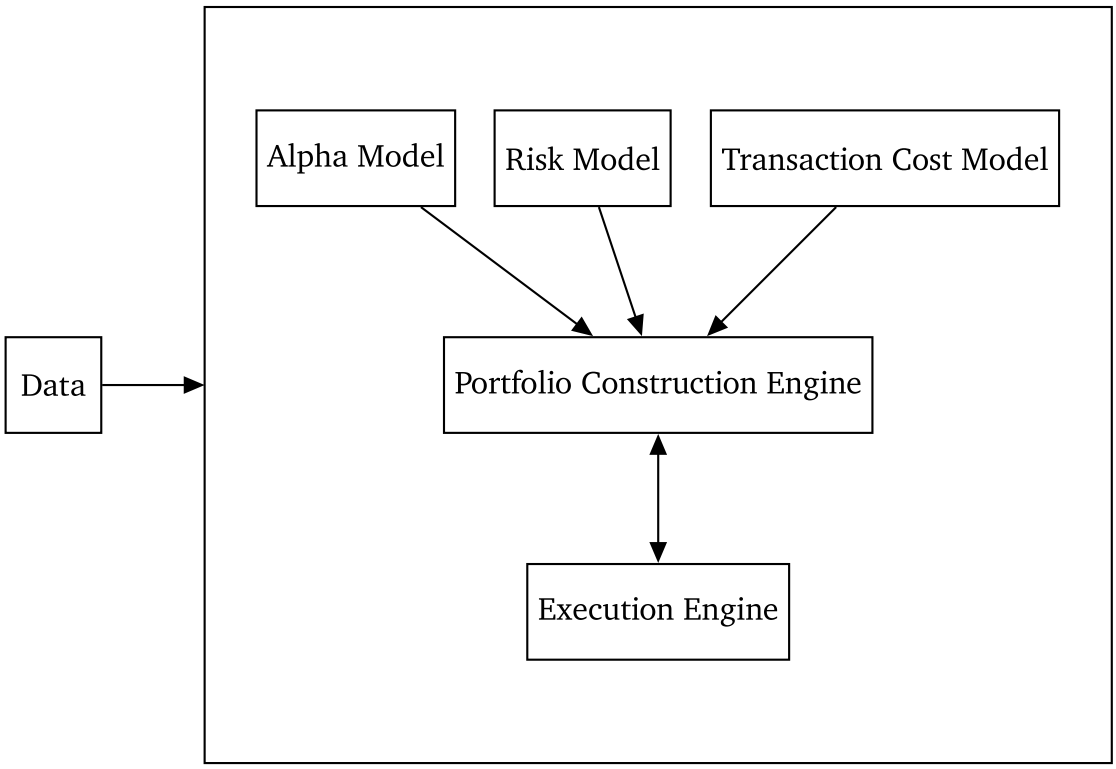
Backtesting is a compulsory stage in the development of any quantitative trading strategies. It evaluates the performance of a strategy with historical data and provides the following usages:

1. validate the effectiveness of the trading idea
2. tune the parameters of the strategy
3. predict the future performance (assuming repeating market patterns)

This note briefly introduces some industrial practices in backtesting a quantitative trading strategy for general first order securities (e.g. equity share, commodity future, etc.) along with some common mistakes. The majority of the content comes from several books and articles including but not limited to Narang 2013, QuantStart 2014, Chan 2008. All references are listed at the end of the note.

The structure diagram of a suggested backtest system is included below.

Structure of Backtest System



2 Note and Assumption

1. All "suggested" values are annualized, calculations are stated below
2. All "suggested" values are calculated after deducting transaction cost
3. Returns at different time t are assumed to be IID [1], otherwise the estimation of Sharpe ratio from sample needs to be adjusted accordingly

3 Primary Metrics

Primary metrics should be used for all types of trading strategies.

3.1 Sharpe Ratio

Metric Introduction

Sharpe ratio is first introduced by Sharpe 1966. Its original name "Reward-to-Variability Ratio" reflects its nature of balancing return and risk of a strategy. According to the

[1] Independent and identically distributed, i.e. assuming same probability distribution and mutually independent

definition in Sharpe 1994, assume R_{Pt} as a t-period return series, R_{ft} as the risk-free rate series over the same period. Then the Sharpe ratio S_h from $t = 1$ to $t = T$:

$$
\begin{aligned}
S_h &\equiv \frac{\overline{D}}{\sigma_D} \\
\text{where } D &\equiv R_{Pt} - R_{ft} \\
\overline{D} &\equiv \frac{1}{T}\sum_{t=1}^{T} D_t \\
\sigma_D &\equiv \sqrt{\frac{\sum_{t=1}^{T}(D_t - \overline{D})^2}{T-1}}
\end{aligned}
$$

This Sharpe ratio indicates the historical average differential return per unit pf historical variability of the differential return (Sharpe 1966). In simpler terms, Sharpe ratio measures the expected return gained per unit of risk taken for a zero investment strategy. The Sharpe ratio does not cover cases in which only one investment return is involved (Sharpe 1994).

Suggested Level

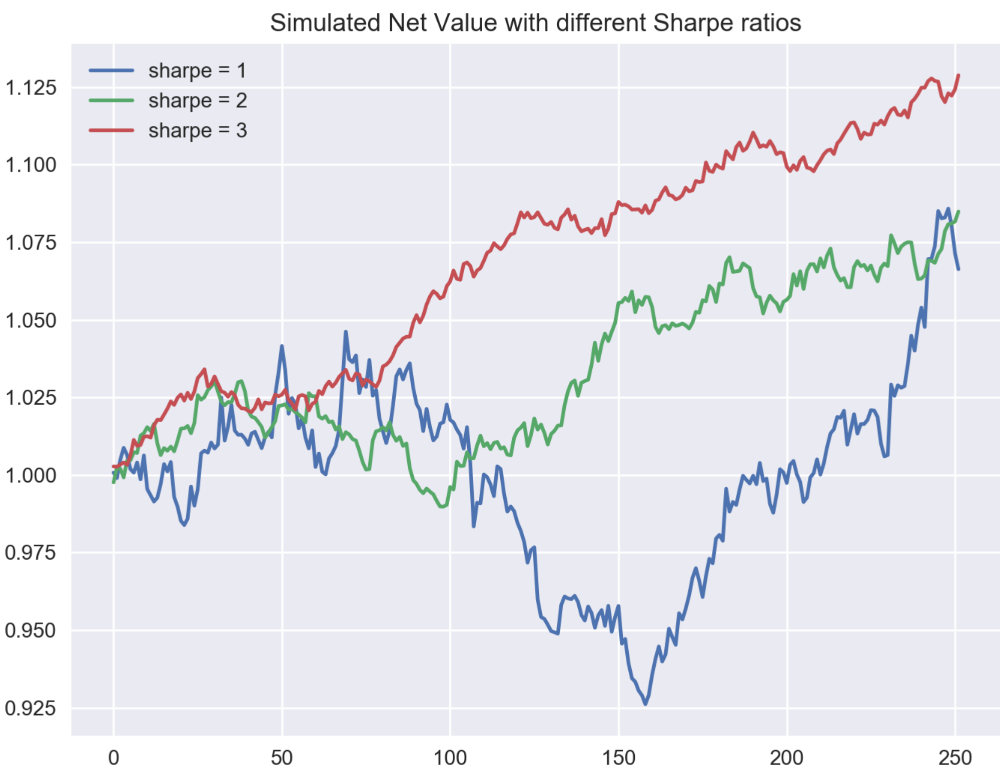


Figure 1: Equity Curve of Multiple Sharpe Ratios

The above diagram shows 3 different return series of Sharpe ratios ranging from 1 to 3, with 252 steps (simulating year-long daily returns). For a day-frequency strategy, $S_h > 1$ usually is not enough to generate consistent profits. A Sharpe value greater than 1.5 or even 2 is recommended. For longer-frequency strategies (i.e. weekly, monthly), $S_h > 0.7$ can be acceptable, $S_h > 1.2$ can be regarded as very good.

All values in the above section should be treated as reference instead of absolute limit/standard to judge a strategy.

3.2 Maximum Drawdown

Metric Introduction

Maximum drawdown is a specific measure of drawdown (the peak-to-trough decline during a specified timespan) that measures the greatest decline from a peak, before a new peak is reached.

$$MDD = \min DD_i \qquad \text{where } i \in \{0, ..., T\}$$
$$DD_t = \frac{V_t}{\max\{V_0, V_1, ..., V_t\}} - 1 \qquad \text{for } t \in \{0, ..., T\}$$

Note that it only measures the size of the largest loss, not the frequency of large losses. MDD does not indicate how long it took an investor to recover from the loss, or if the investment even recovered at all.

Suggested Level

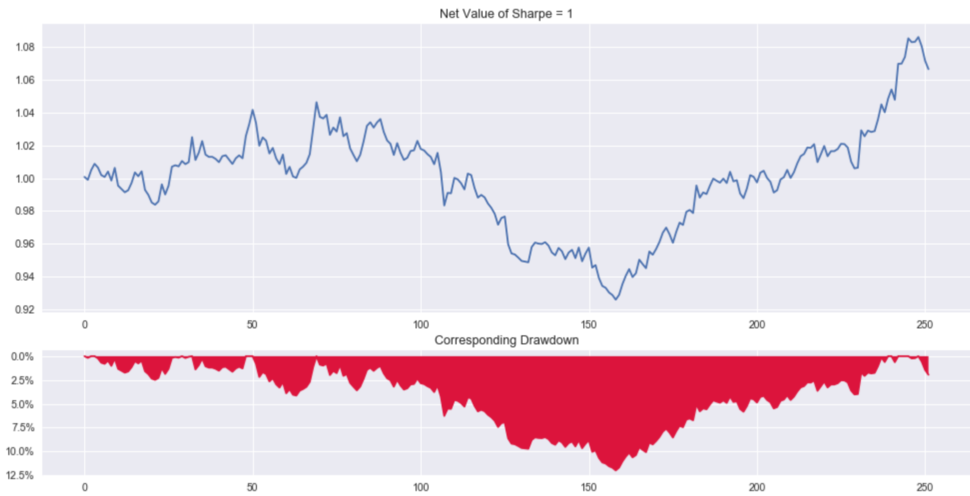


Figure 2: Equity Curve and Drawdown Graph

The above diagram shows the net value and corresponding drawdown throughout the investment timespan. As shown in the diagram, 0% drawdowns refer to new peaks, consistent drawdowns below 0 refer to continuous loss. The maximum drawdown is the largest loss relative to the most recent peak in the investment timespan.

3.3 Win Rate, Profit Factor and Payoff Ratio

Metric Introduction

Let π_t be the profit/loss of a strategy at time t, T be the total number of steps (timespan). Assume the profit/loss is non-zero at every time t, i.e. $n_{\pi=0} = 0$, then $T = n_{\pi<0} + n_{\pi>0}$. Let w be the win rate, pf be the profit factor, pr be the payoff ratio.

$$
\begin{aligned}
w &= \frac{pf}{pf + pr} \\
\text{where } w &\equiv \frac{n_{\pi>0}}{n_{\pi<0} + n_{\pi>0}} \\
pf &\equiv \frac{\sum_{t,\pi_t>0} \pi_t}{\sum_{t,\pi_t<0} \pi_t} \\
pr &\equiv \frac{\sum_{t,\pi_t>0} \pi_t}{\sum_{t,\pi_t<0} \pi_t} \cdot \frac{n_{\pi<0}}{n_{\pi>0}}
\end{aligned}
$$

Win rate is expressed as the ratio of profiting time to the total investment timespan. Profit factor is the ratio of the sum of winning trades and losing trades. Payoff ratio is the ratio of winning trades average to losing trades average.

Suggested Level

Win rate and payoff ratio are determined by the strategy characteristics, for example bare momentum-based strategies are usually related with relatively lower win rates due to the price reversion nature in multiple short term periods. Profit factor is more deterministic due to its definition. Therefore, we should focus on the former two metrics while win rate being the more dominant factor.

As win rate reflects the profiting consistency of a strategy, a higher win rate is almost always preferred. Win rate below the 50% mark requires higher payoff ratio to compensate, the opposite portion is fine with smaller profit margin. Many profitable strategies made by individuals have win rates around 45% to 55%. Professional knowledge and optimisation is often required to go beyond the common band.

4 Secondary Metrics

Secondary metrics provide easy explanation for non-finance-heavy personnel.

4.1 Compound Annual Growth Rate (CAGR)

Metric Introduction

Compound annual growth rate (CAGR) is the annualized, required rate of return for an investment to grow in timespan T (in years), assuming the intermediate profits are reinvested.

$$CAGR = \left(\frac{V_T}{V_0}\right)^{\frac{1}{T}} - 1$$

CAGR is not the true rate of return, but rather a smoothed, representational figure, usually used for easier explanation and comparison.

Suggested Level

The desired CAGR depends on the nature of the security and even its sector. Different types of securities (e.g. equity, fixed income, index, derivative) have different return characteristics. Equity-type securities generally have a higher CAGR while corresponding derivatives are even more volatile. The performance of fixed income products are more consistent over time.

Note that CAGR can be affected by the level of leverage. Let I_0 and L_0 be the amount of private capital and leverage (borrowed capital) at time $t = 0$, r be the interest rate. Assume a leverage ratio of 0.2, i.e. $\frac{L_0}{I_0+L_0} = 0.2$. Then

$$\begin{aligned} CAGR &= \left(\frac{V_T - L_0 \times (1+r)^T}{I_0}\right)^{\frac{1}{T}} - 1 \\ &= \left(\frac{V_T - 0.25 I_0 \times (1+r)^T}{I_0}\right)^{\frac{1}{T}} - 1 \\ &= \left(\frac{V_T}{I_0} - 0.25(1+r)^T\right)^{\frac{1}{T}} - 1 \end{aligned}$$

Therefore, when comparing the CAGR of different strategies, the leverage level and cost should be kept the same.

4.2 Volatility of Return

Metric Introduction

Volatility is a statistical measure of the dispersion of returns. Volatility is often measured as either the standard deviation or variance of returns. Given $r_1, r_2, \ldots, r_t$ be the return series of a t-step timespan.

$$\sigma_r = \sqrt{\frac{\sum_{i=1}^{t}(r_i - \mu_r)^2}{t-1}}$$

In most cases, higher volatility reflects higher risk of the strategy or security. So investors should monitor the metric as a proxy of investment risk.

Suggested Level

Investors will have their corresponding level of tolerance. One may not prefer frequent, large fluctuations of their asset value such as a retired person who just aim to keep his/her savings with minimal investment income to combat inflation. On the other hand, younger investor may be more aggressive as they are able to endure more risks. Thus, the level of risk tolerance should be tailored based on personal situation independently.

4.3 Maximum Drawdown Duration

Metric Introduction

lorem

$$y = f(x)$$

lorem

Suggested Level

lorem

5 Common Pitfall

This section introduces multiple common mistakes made by quants in backtest.

5.1 Survivorship Bias

lorem

5.2 Transaction Costs

lorem

5.3 Market Nature/Pattern

lorem

5.4 Look Ahead Bias

lorem

5.5 Overfitting

lorem

Conclusion

lorem

Reference

[1] George E. P. Box. “Science and Statistics”. In: *Journal of the American Statistical Association* 71.356 (1976), pp. 791–799.

[2] Rishi K. Narang. *Inside the Black Box. A Simple Guide to Quantitative and High Frequency Trading*. 2nd ed. Wiley, Mar. 20, 2013.

[3] QuantStart. *Successful Algorithmic Trading. A Step-by-step Guide to Quantitative Strategies*. 2014.

[4] Ernest Chan. *Quantitative Trading. How to Build Your Own Algorithmic Trading Business*. Wiley, Nov. 17, 2008.

[5] William F. Sharpe. “Mutual Fund Performance”. In: *The Journal of Business* 39.1 (1966), pp. 119–138.

[6] William F. Sharpe. “The Sharpe Ratio”. In: *The Journal of Portfolio Management* 21.1 (1994), pp. 49–58.