Introduction to Quantitative Trading Strategies

Harkishan Singh Baniya

Centre of Excellence for Computational
Mathematics and Data Science,
Indian Institute of Technology Madras

Disclaimer

This workshop is for educational purposes only and does not constitute an offer to sell, a solicitation to buy, or a recommendation for any security; nor does it constitute an offer to provide investment advisory or other services by the speakers. Nothing contained herein constitutes investment advice or offers any opinion with respect to the suitability of any security and any views expressed herein should not be taken as advice to buy, sell, or hold any security or as an endorsement of any security or company. The speakers are not responsible for the losses incurred due to the buying and selling of securities.

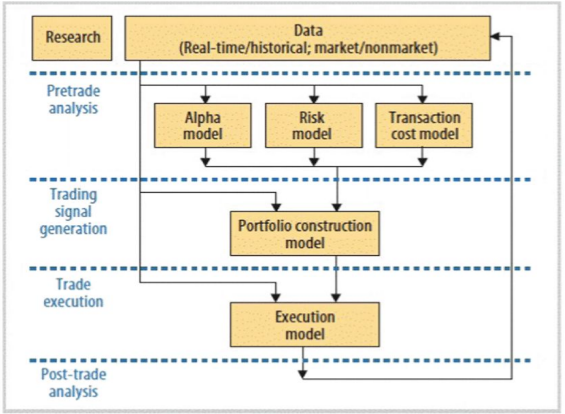
Why Quantitative Trading ?

Discretionary vs Systematic

Discretionary

Source: the balance

Systematic



Source: Research Gate

Systematic Trading

- Systematic Trading involves set of instruction and steps that is executed by an algorithm.
- It can be backtested and risks can be quantified using historical data and quantitative models.
- Avoid human cognitive biases and risk associated with human emotions.
- Complex strategies can be only executed using systematic trading as they are not possible for humans. A good example is HFT strategies.
- Systematic Trading strategies can be fully automated to run without any human intervention.

Trading Systems

Classification of Trading Strategies

Momentum

- Time series
- Cross-sectional

Source: auquan

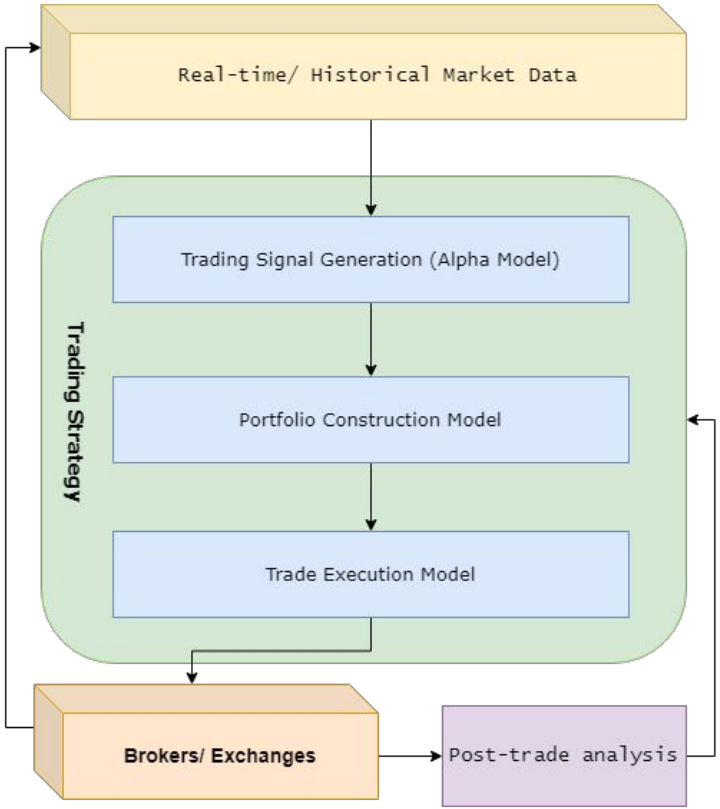
Mean-reversion

- Statistical Arbitrage
- Time series
- Pairs-trading

Market making

Source: River financials

Architecture of a Trading System



- Trading systems depends on data for trade generation from the trading strategy.
- The alpha model generates a the trading signals which are combined using a portfolio construction model for optimal risk profile.
- The orders of trades are sent to brokers/exchange by the trade execution model for optimal execution.
- Order fill data is analysed by during post-trade analysis.

Risks with Automated Trading

- **Extreme market events**: Trading systems are designed for particular market regime and condition, during extreme market conditions the algorithm may not be able to take decisions and lead to adverse effects.

- **Lack of Transparency**: Automated trading system can get very complex and turns into a black box.

- **Bug in the algorithm**: With increasing complexity the chances of having bugs in the system increases and in turn increases the chance of catastrophe.
  https://www.bbc.com/news/magazine-19214294
  https://www.henricodolfing.com/2019/06/project-failure-case-study-knight-capital.html

Alpha Generation

Sourcing Ideas

The way to inspired for ideas is to read research papers (mostly academic and better if peer reviewed), peer reviewed journals and research articles. Some sources are listed below -

Research paper libraries

- SSRN: https://www.ssrn.com/
- arXiv: https://arxiv.org/
- RePEc: http://repec.org/

Journals

- The Journal of Finance
- Journal of Portfolio Management
- Journal of Financial Data Science

Newsletter

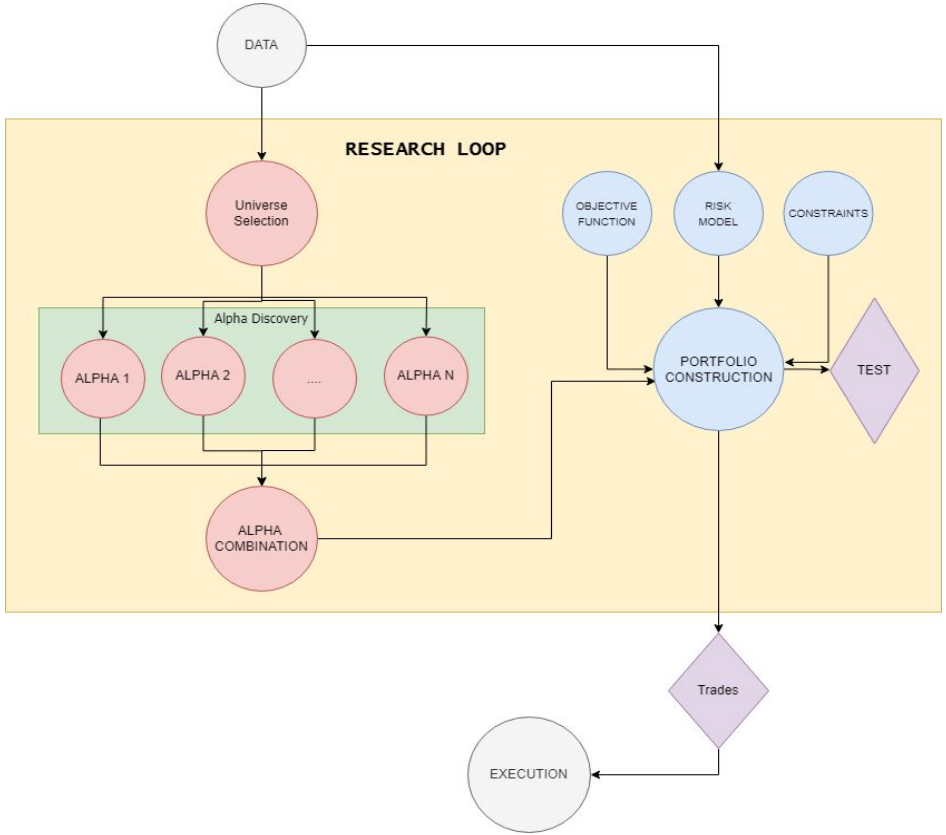
- Paper Digest: https://www.paperdigest.org/
- Quantocracy: https://quantocracy.com/

Framing a Research Question

- A research question is a question used for frame the main hypothesis of your research for the answer you are seeking for.
- A good research question is the key for doing quantitative research. Without a clear research question the focus of the research might shift and may not lead to anything useful and indeed may lead to some false discovery.
- Vague question must be avoided at all cost e.g. “Does momentum strategy work ?”.
- A research question must clear, precise and provide specific information for a reader to clearly understand it e.g.: “Did cross-sectional momentum on the top 10 equities by market cap over last 5 years out-performed the market ?”.

Research Structure

- Research is highly dependent on quality and type of data. “Garbage In Garbage Out ”
- The research loop starts with selecting an universe of securities.
- The alpha discovered from various ideas are then combined using a model.
- Then portfolios are constructed using some risk model and constraints.
- Finally, the portfolio is backtested and iterative research is conducted following some protocols to avoid overfitting.



Risk Management

Two types of market risks

Two important risks that investors must be concerned about within their investment portfolio are - *Idiosyncratic* (or specific) and *Systematic* (or market) risk.

Idiosyncratic risk

This risk is associated with holding a particular asset like a single company stock, which is also known as company-specific risk or a group of an asset belonging to a similar category like stocks from the same sector, which is also known as sector-specific risk. It can be mitigated with diversification i.e. by reducing exposures to a particular asset or sector and transferring that to some uncorrelated assets or sectors.

Systematic risk

This is a risk associated with the market itself and it cannot be mitigated with diversification. The most popular solution is to hedge against such situations using derivatives like options. Though hedging provides good downside protection, at the same time it also limits the overall portfolio’s upsides, because it's complex and expensive to implement.

Market Risk Exposure

- Risk exposure to market also known as beta exposure is how sensitive an asset or portfolio is to the market movement i.e. how dependent it is on the market movement.

- If an asset or portfolio have high beta exposure than it will perform very well market is rising but will do very poorly when market is falling.

$$Y = \beta X + \alpha$$

$$Y_{portfolio} = \beta_1 X_{AAPL} + \beta_2 X_{SPY} + \alpha$$

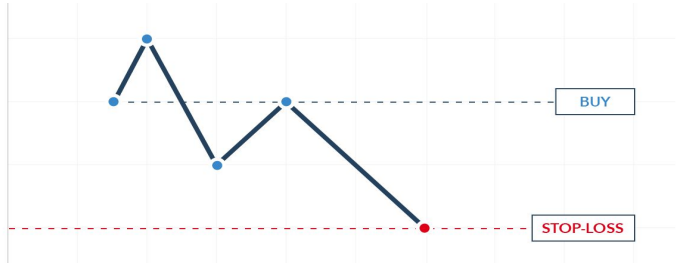
- To avoid market risk or systemic risk we want to have a portfolio with negligible beta exposure to the market that means the return of the portfolio will totally really on alpha and other independent factors (which can understood by using factor models).

Stop-loss

- Stop-loss is an order that is placed to avoid further loss when the price reaches a specific level. It is used as an risk management measure that make sure that entire account does get empty on a single bet.

- If stop-loss level is not defined properly it might lose its purpose or reduce your strategies expected return (opportunity cost). This can happen when trading a momentum strategy narrow stop-loss level will often get trigger during higher volatility and trades might get exited sooner than expected.

- Instead of a fixed stop-loss another strategy is dynamically adjust the stop-loss level depending on the strategy performance. This is also known as trailing stop-loss, as the strategies accumulate positive return the stop-loss level increase in some defined proportion.



Bet Sizing

- No matter how profitable alpha you design, if you don’t have a bet sizing strategy it will surely lose money.
- Bet sizing can be independent of the strategy i.e. it places fix bets regardless of the confidence in the signals generated by the strategy or it can be based on the probability of a trade i.e. higher probability trades gets bigger bets compared low probability trades.
- Dynamic bet sizing updates the bet size based on the market condition and volatility. Again this type of bet sizing depends on some from of predictions. The intuition is somewhat similar to volatility scaling.
- Kelly criterion is a very well know bet sizing technique developed by John Kelly at AT&T’s Bell Laboratories.

$$K = W - \frac{1 - W}{R}$$

where

$$K = \text{Kelly Percentage}$$
$$W = \text{probability of winning}$$
$$R = \frac{\text{Total Wins}}{\text{Total Loss}}$$

Hedging

- Hedging is a way to mitigate the market risks by taking a position such that it offsets the market exposure.

$$\alpha + \beta X - \beta X = \alpha$$

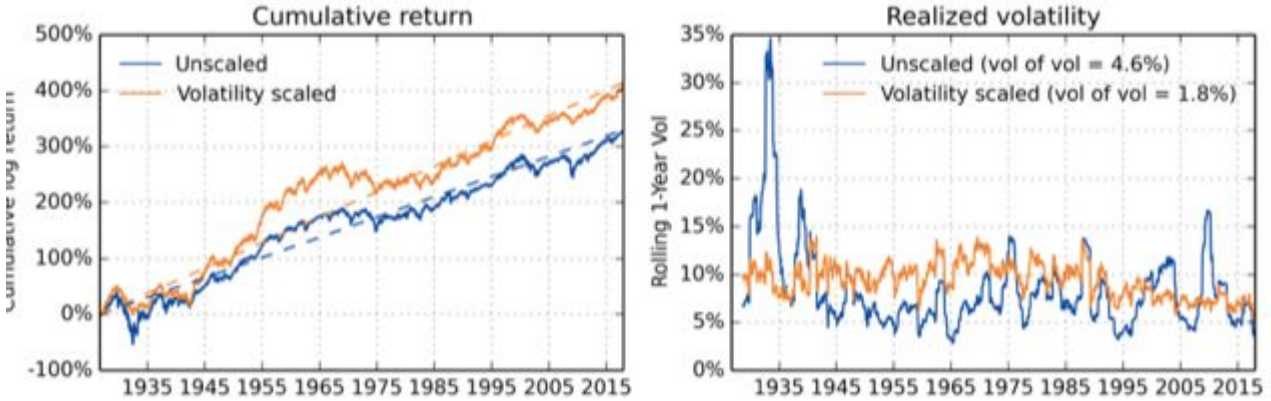
- Hedging leads to return which is purely alpha and independent of market movements i.e. it become market neutral. But it will result reduction of potential profits when the market is performing.

- The caveat is that beta will not remain same over time and it becomes difficult to stay perfectly hedged at any given point in time.

- Strategies like pairs-trading are automatically hedged on its own as we take two opposite offsetting positions.

Volatility Scaling

- Realized volatility of a asset is calculated as the standard deviation of the returns. It is basically measure the fluctuation in the returns, higher fluctuation will result in higher volatility and uncertainty.
- The idea of volatility scaling or targeting is maintain a constant volatility exposure by increasing or decreasing the amount of *leverage* on that asset. If the volatility goes up, he/she has to scale down the portfolio. On the other hand, if the volatility goes down, he/she should take more leverage.
- Volatility scaling depends on the future expectation of realized volatility which is some form of a prediction of the future volatility.



Portfolio Construction

- Portfolio construction is a way to mitigate Idiosyncratic risk by diversifying the specific risk among uncorrelated investment.

- After we deciding the instruments to make a portfolio, next we need to decide how much money to allocation each of the instrument such that the risk is minimized or profit is maximized or both. This is an optimization problem which is know as portfolio optimization.

- A classic portfolio optimization method is called mean-variance optimization where the problem is to *minimise risk* subject to a *return constraint* (i.e the portfolio must return more than a certain amount).

- Another very famous method is the Efficient frontier method which was published by Harry Markowitz [1952]. It finds the portfolio with maximum expected return at a specific risk level, which is represented by the efficient frontier curve.

Market Execution

Trade Execution Strategies

- Trade execution strategies are methods used by traders to place orders in the market with the goal of achieving optimal results, such as minimizing costs, reducing market impact, and improving execution speed.
  - Key Objectives of Trade Execution:
    - **Minimize Market Impact**: Avoid moving the price unfavorably when placing large orders.
    - **Reduce Transaction Costs**: Minimize commissions, slippage, and spread costs.
    - **Maximize Fill Probability**: Ensure orders are executed efficiently at desired prices.

  - Types of Trade Execution:
    - **Passive Execution**: Trades that do not seek to move the market (e.g., limit orders).
    - **Aggressive Execution**: Trades that take immediate action to secure a fill (e.g., market orders).

Some Trade Execution Strategies

- **Time-Weighted Average Price (TWAP)**: A strategy that divides the order into smaller parts and executes them over a set period, aiming to match the average price of the asset over that time.
  - Pros: Reduces market impact by avoiding large price fluctuations.
  - Cons: Execution time can result in slight missed opportunities if the market moves sharply.
  - Ideal for traders with large orders in illiquid markets.

- **Volume-Weighted Average Price (VWAP):** A strategy that breaks the order into smaller pieces and executes them in proportion to the market's trading volume throughout the day.
  - Pros: Attempts to minimize impact and achieve an average price based on volume.
  - Cons: Can be slow in execution.
  - Suitable for traders with institutional-level order volumes aiming to minimize market impact.

Advanced Execution Strategies and Algorithmic Trading

- **Implementation Shortfall**: A strategy designed to minimize the difference between the decision price (the price when the order was placed) and the actual execution price.
  - Pros: Designed for efficiency in volatile or illiquid markets.
  - Cons: Risk of missing favorable opportunities if the market moves before execution.
  - Used by institutional traders to avoid significant market impact.

- **Smart Order Routing (SOR):** A technology-driven strategy where an order is automatically routed to the most optimal exchange or venue, depending on factors like liquidity, price, and cost.
  - Pros: Ensures the best possible execution, reduces the impact of spread costs.
  - Cons: Dependency on technological infrastructure, higher complexity.
  - Used in high-frequency trading (HFT) and for large-scale institutional trades.

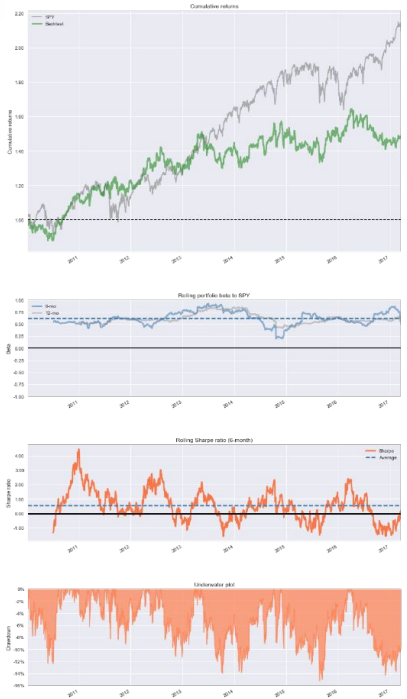
Backtesting

Motivation behind Backtesting

- Does the trading strategy you have hypothesised even perform ?
- Backtest is a process of testing the historical performance of a trading strategy. It is used as to check the if a strategy works on past and what performance we can expect in future.
- A backtest is a simulation which can be used for sanity check on the hypothesis behind the strategy and various factors under a given scenario of the market.
- Backtest is not an experiment and shouldn't aimed to prove profitability of a strategy as it is only historical performance.

```
Entire data start date: 2010-01-04
Entire data end date: 2017-03-21
Backtest months: 86
```

	Backtest
Annual return	5.7%
Cumulative returns	49.1%
Annual volatility	15.0%
Sharpe ratio	0.44
Calmar ratio	0.37
Stability	0.79
Max drawdown	-15.3%
Omega ratio	1.08
Sortino ratio	0.63
Skew	-0.28
Kurtosis	3.15
Tail ratio	1.03
Common sense ratio	1.09
Daily value at risk	-1.9%
Gross leverage	2.00
Daily turnover	16.2%
Alpha	-0.00
Beta	0.59



Preparing the Data

- Data is the fuel to quantitative models and the high quality data is crucial for success and doing quantitative research.
- Data can have missing values, error in recording, fault in cleaning or issues with aggregation.
- There might also be other issues with the dataset altogether i.e. survivorship bias in the universe selection or having low frequency data which might restrict the backtest to have a narrow view.
- Handling outliers in the data is nuisance, removing dataset will introduce bias in research and keeping it may lead to skewness in analysis. Some methods that scientists prefer are - Winsorization, Mean/Median Imputation etc.
- Check these videos for more details:
  https://www.youtube.com/watch?v=bN5aZLp0gvw&ab_channel=Quantopian
  https://www.youtube.com/watch?v=ZME_61rPNVY&ab_channel=Hudson%26Thames

Types of Backtest algorithms

Vectorized

- Vectorized the backtesting loop using Numpy/Pandas
- Fastest
- Not a reliable backtest. Is used as a sanity check.
- Easy to implement.

For Loop

- The backtest is executed using a for loop.
- Slower than vectorized.
- Not a reliable backtest. Is used as a sanity check.
- Easy to implement.

Event Driven

- Uses market simulation and runs the backtest as events takes place.
- Slowest
- The most reliable backtest.
- Difficult to implement.

Analysing a Backtest

Risk and Performance Metrics (1)

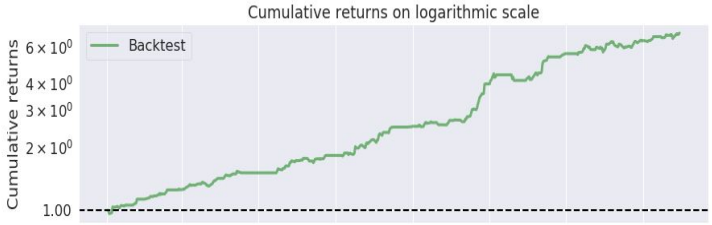
- **Cumulative Returns**

$$R_c = \prod_{t=0}^{n}(1 + r_t)$$

where

r_t = return at t

n = number of observations



- **Annualized Return**

$$R_A = R_c^{\frac{m}{n}} - 1$$

where

R_c = Cumulative returns

m = periods per year

n = number of observations

Risk and Performance Metrics (2)

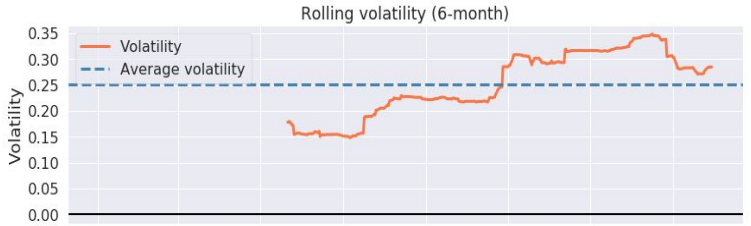
- **Annualized Volatility**

$$\sigma_A = \sigma_r \sqrt{m}$$

where

σ_r = std. dev. of returns

m = periods per year



- **Annualized Sharpe Ratio**

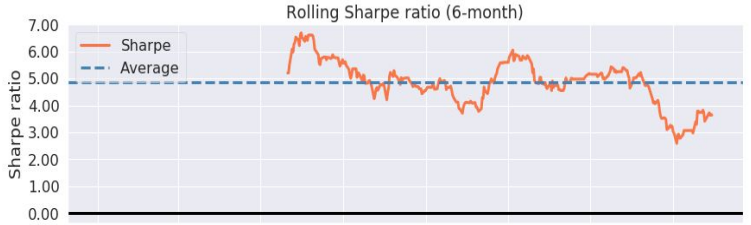
$$SR = \frac{R_A - R_f}{\sigma_A}$$

where

R_A = Annualized return

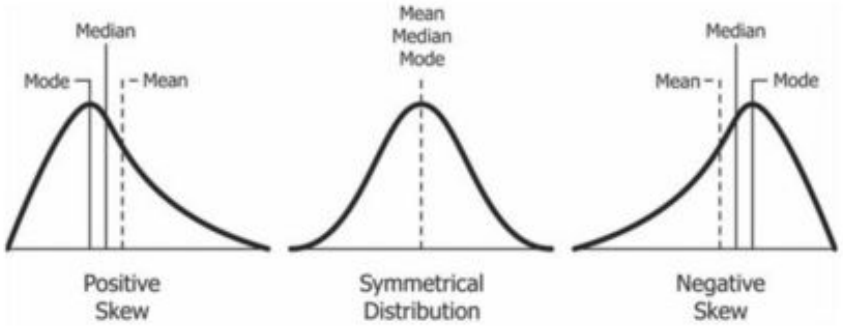
R_f = risk-free rate

σ_A = Annualized volatility



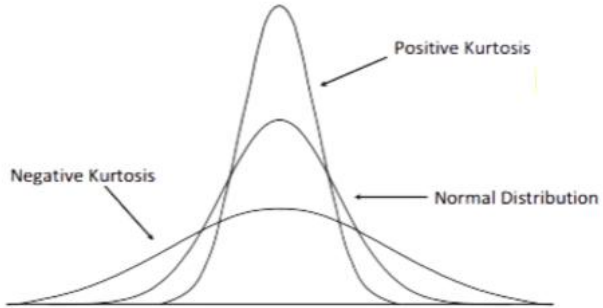
Risk and Performance Metrics (3)

- **Skewness**



Source: wikipedia

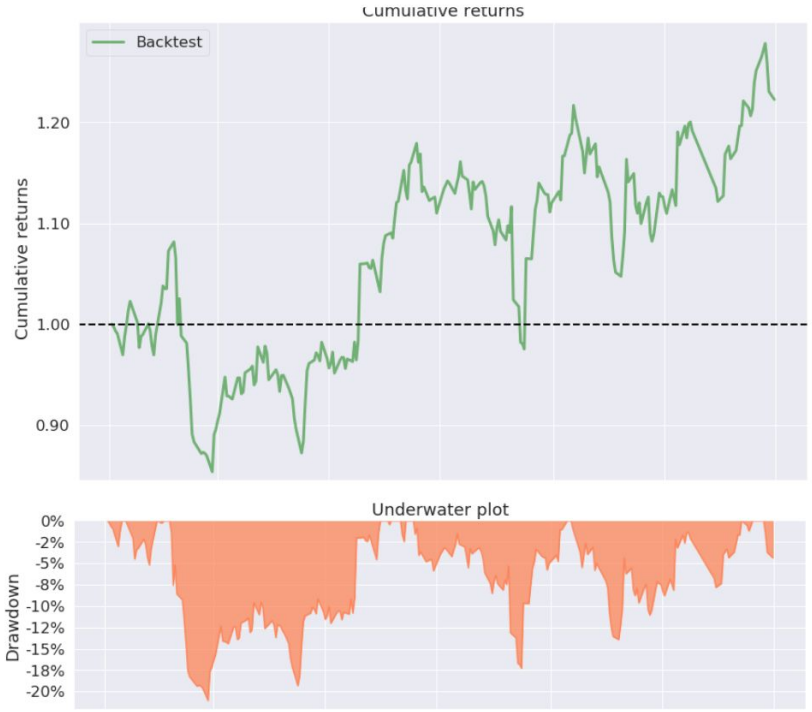
- **Kurtosis**



Risk and Performance Metrics (4)

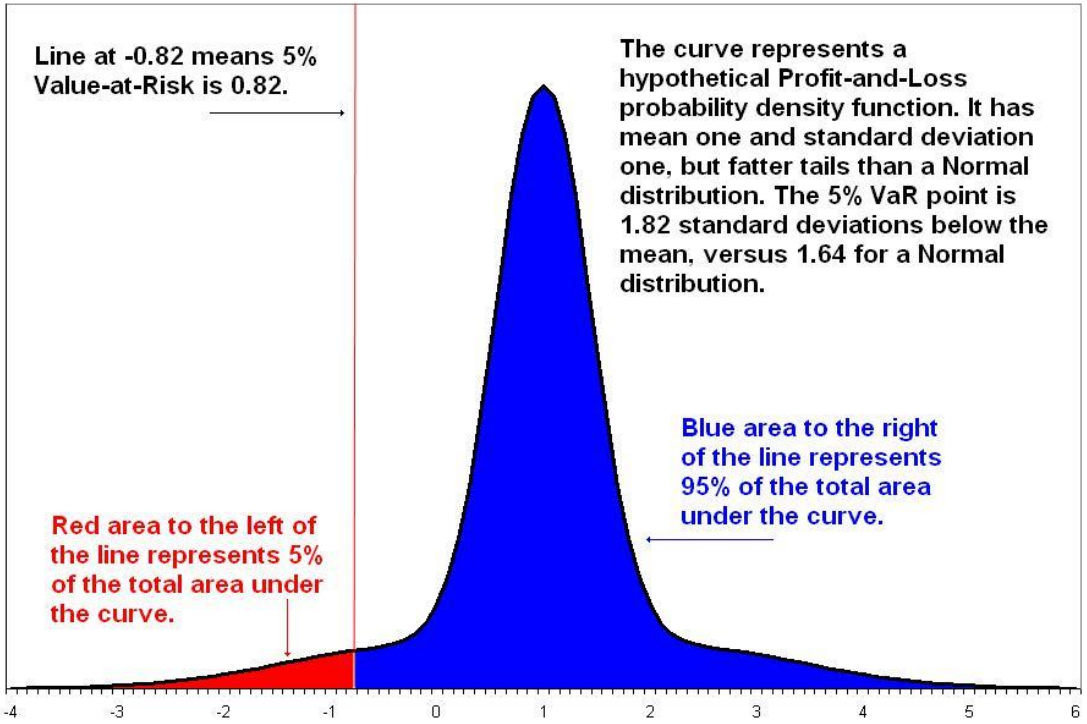
- **Maximum Drawdown**

$$MDD = \frac{\text{Trough Value} - \text{Peak Value}}{\text{Peak Value}}$$



Risk and Performance Metrics (5)

- **VaR (Value at Risk)**



Source: wikipedia

Backtest Examples

Backtest of a Momentum Strategy

Notebook Link: https://bit.ly/42X7YRN

Backtest of a Constant Proportion Portfolio Insurance

Notebook Link: https://bit.ly/4kfcW2z

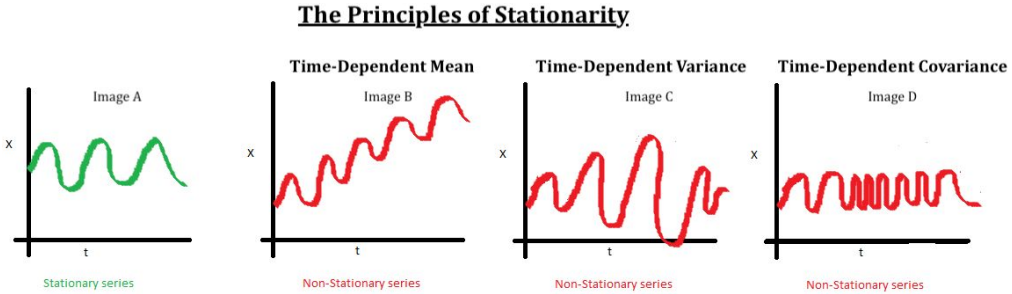
Common Pitfalls

Philosophical Misconceptions

- **Forecasting Price**: Market is a dynamic system that changes its states depending on the supply and demand. Most of the statistical methods and machine learning techniques has the base assumption that the data is stationary, samples are IID and from a normal distribution. Market data does qualify even a single requirement, although there are techniques to solve each of these problem the dynamic nature of the market doesn't seem to be fit for forecasting.

- **Backtest for validation:** Backtest shouldn't considered as a research tool rather a sanity check of the theory. It is must to develop a theory (or hypothesis) first than work on proving that or contradicting using feature importance analysis which is robust to overfitting.

- **All weather strategies:** Try to avoid researching for that holy grail strategy which you think should work on all market regimes because likelihood of the existence of such strategy is rather slim and discovery is most likely be a false positive. Rather focus on strategies that can perform optimally under a specific regime like mean-reversion or market sell-off.



Seven Sins of Quantitative Investing

1. *Survivorship bias*: Ignoring the stocks that have gone bust or delisted in past and only considering the existing stocks in the sample. This can lead to overestimation of historical performance. E.g.- Only considering the current composition of SP500.

2. *Look-ahead bias*: Using information that is not available at the current moment.

3. *Storytelling*: Making up a story ex-post to justify some random pattern in the result.

4. *Data snooping*: Using the test sample for tuning the strategy and improving the backtest.

5. *Transaction cost*: Ignoring the transaction cost results inaccurate backtest results. While including transaction cost some strategies that were profitable earlier may fail.

6. *Outliers*: A backtest is severely skewed if it performance (profit or loss) is heavily dependent on few extreme outcomes that is observed in past. E.g.: Financial crash, Tech Boom.

7. *Shorting*: Shorting involves finding lender who can lend securities, which is dependent on inventory, relative demand and markets.

References

- Michael Halls-Moore. *Successful Algorithmic Trading*
- Yves Hilpisch. *Python for Finance*
- Chan, 2013. Algorithmic Trading: Winning Strategies and Their Rationale
- Lopez de Prado, 2018. Advances in Financial Machine Learning
- Luo et al. [2014]. “Seven Sins of Quantitative Investing”
- Arnott, Robert D. and Harvey, Campbell R. and Markowitz, Harry, A Backtesting Protocol in the Era of Machine Learning (November 21, 2018). Available at SSRN: https://ssrn.com/abstract=3275654 or http://dx.doi.org/10.2139/ssrn.3275654
- Bailey, David H. and Borwein, Jonathan and Borwein, Jonathan and López de Prado, Marcos and López de Prado, Marcos and Zhu, Qiji Jim, Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-of-Sample Performance (April 1, 2014). Notices of the American Mathematical Society, 61(5), May 2014, pp.458-471, Available at SSRN: https://ssrn.com/abstract=2308659 or http://dx.doi.org/10.2139/ssrn.2308659

References

- Chan, 2013. Algorithmic Trading: Winning Strategies and Their Rationale.
- Chan, 2008. Quantitative Trading: How to Build Your Own Algorithmic Trading Business.
- Carver Robert, 2015. Systematic Trading: A unique new method for designing trading and investing systems
- Moskowitz, Tobias J. and Moskowitz, Tobias J. and Ooi, Yao Hua and Pedersen, Lasse Heje, Time Series Momentum (September 1, 2011). Chicago Booth Research Paper No. 12-21, Fama-Miller Working Paper, Available at SSRN: https://ssrn.com/abstract=2089463 or http://dx.doi.org/10.2139/ssrn.2089463
- Gatev, Evan and Goetzmann, William N. and Rouwenhorst, K. Geert, Pairs Trading: Performance of a Relative Value Arbitrage Rule (February 2006). Yale ICF Working Paper No. 08-03, Available at SSRN: https://ssrn.com/abstract=141615 or http://dx.doi.org/10.2139/ssrn.141615
- Jegadeesh and Titman. Journal of Finance 2001. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations

References

- Lopez de Prado. 2018. Advances in Financial Machine Learning

- Cerniglia, J. A., Fabozzi, F. J., & Kolm, P. N. (2016). Best Practices in Research for Quantitative Equity Strategies. The Journal of Portfolio Management, 42(5), 135–143. doi:10.3905/jpm.2016.42.5.135

- López de Prado, Marcos and López de Prado, Marcos, The 7 Reasons Most Machine Learning Funds Fail (Presentation Slides) (September 2, 2017). Available at SSRN: https://ssrn.com/abstract=3031282 or http://dx.doi.org/10.2139/ssrn.3031282

- López de Prado, Marcos and López de Prado, Marcos and Lipton, Alex, Three Quant Lessons from COVID-19 (Presentation Slides) (March 27, 2020). Available at SSRN: https://ssrn.com/abstract=3562025 or http://dx.doi.org/10.2139/ssrn.3562025

- López de Prado, Marcos and López de Prado, Marcos, Backtesting (May 14, 2015). Available at SSRN: https://ssrn.com/abstract=2606462 or http://dx.doi.org/10.2139/ssrn.2606462