# Computational Investing Part I – 24/02/2013

## Week 1 – 24/02/2013

How Portfolio Managers get compensated?

1) Expense ratio – Take a fix percentage of the total assets under management, usually less than 1%.
2) Two and Twenty – Take a fix 2% of total assets under management + 20% of profits.

How to attract investors?

1) Who's your audience?
   a. Individuals – Wealthy individuals who looks for portfolio management services.
   b. Institutions – Large institutions like Endowments and Pension funds are looking to diversify by outsourcing some of the capital to outside managers.
   c. Funds of Funds – There are funds that specialized in investing in other funds.
2) You must have a positive track record or
3) Have a compelling story of a proven back-tested strategy.
4) Fit a "pigeon hole" – Some funds and institution are looking to diversify by investing in different asset classes, if your fund fit what they are looking for they might invest in your fund.

What do investors want to see?

1) Reference to a benchmark – In the Pigeon Hole example you are compared to the "hole" you should have fitted (for example Emerging Markets specialization).
2) Absolute return – "don't lose", willing to expect lower return but demanding lower risk/volatility.

Common Metrics

1) Annual Return
2) Risk – Standard deviation of return = Volatility.
3) Risk – Maximum Draw down.
4) Reward / Risk – Sharpe Ration / Sortino Ratio (Sharpe ratio "panelized" for both upward and downward volatility, where Sortino Ratio panelize for downward volatility only).
5) Jensen's Alpha.

Example:

| | Return | Sharpe | STDEV | Draw-Down | Correlation |
|---|---|---|---|---|---|
| **Fund** | 33% | 0.94 | 0.58% | -8.67% | 0.89 |
| **$DJI** | 43% | 0.63 | 1.23% | -27.38% | 1.00 |

In the above example, while the Dow Jones index (Benchmark) had overall better Return, it also had a greater volatility and a bigger maximum draw-down, meaning it had more risk in it compared to the fund.

Calculating Return

Period Return = (Value[End] / Value[Start]) – 1
Example: if the fund goes from 100$ to 110$ → 110/100 – 1 = 0.1 == 10%.

Daily Return = (Value[t] / Value[t-1]) – 1
Standard Deviation of Daily Return = stdev(Daily Return)

Sharpe Ratio

Sharpe Ratio is a Reward/Risk measurement, how much reward are you getting for your risk, it is the ratio of the Average Daily Return and the Standard Deviation of the Daily Returns.

Sharpe Ratio = $\frac{E(R-R_f)}{\sqrt{var[R-R_f}} * K$

Where Rf is the Risk Free rate of return (treasure bill etc), and K is the Square Root of trading days.
Rf it is usually emitted and we are left with:
Average(Daily Returns) / stdev(Daily Returns) * sqrt(Trading days)
Where Trading Days is 250 per year.
If we are calculating a monthly Sharpe Ratio with use the Monthly Return instead of daily return and 12 as Trading Days etc…

What does it mean?

1) Sharpe Ratio is the most "important" or commonly used measure of asset performance.
2) How well does the return of an asset compensate the investor for risk taken.
3) The higher Sharpe Ratio the better.
4) When comparing two assets each with the same return, higher Sharpe Ratio gives more return for the same risk (volatility).

**Week 2** – 03/03/2013

The Value of a company (Fundamental approach)

1) Market Cap – The value of all shares offered to the public on the market, this is value is determined by the market. The underlying assumption is that the market is very efficient in processing information (Efficient Market Hypothesis).
   Value = #shares outstanding * Price.
2) Future Returns / Intrinsic Value – The present value of future return, like dividend payments.
3) Book Value – The net value of the company's assets as described in their Annual/Quarterly reports. Total assets minus intangible assets and liabilities.

Intrinsic Value

The value of a company can be calculated by the Present Value of all its Future Dividends payments:

$$PV = \frac{D}{1+R} + \frac{D}{(1+R)^2} + \cdots + \frac{D}{(1+R)^\infty} = \frac{D}{R}$$

Where:

D = Is the dividend paid every year (assuming the same amount is paid).

R = The discount rate, this is the rate that is used to discount future value into today's value.

For Example, a company will issue 1$ dividend every year, and the discount rate is 5%, so the company's value is: PV = 1$ / 5% = 1$ / 0.05 = 20$

How Does Information Affect Price?

A study conducted by MacKinlay showed that news events affects equity prices, the reason for this is that news event influence on the future operation of a company, and its future profits, affecting present value.

In the same aspect there might be news that affects the entire local market (like droughts), or even the global economy (credit crises), in that case all equity prices will be affected.

**Capital Assets Pricing Model – CAPM**

Assumptions:

1) Return of a stock has two components:
   a. Systematic – The entire market.
   b. Residual – The Expected value of the residual is 0.

$$r_i = \beta_i * r_m + \alpha_i$$

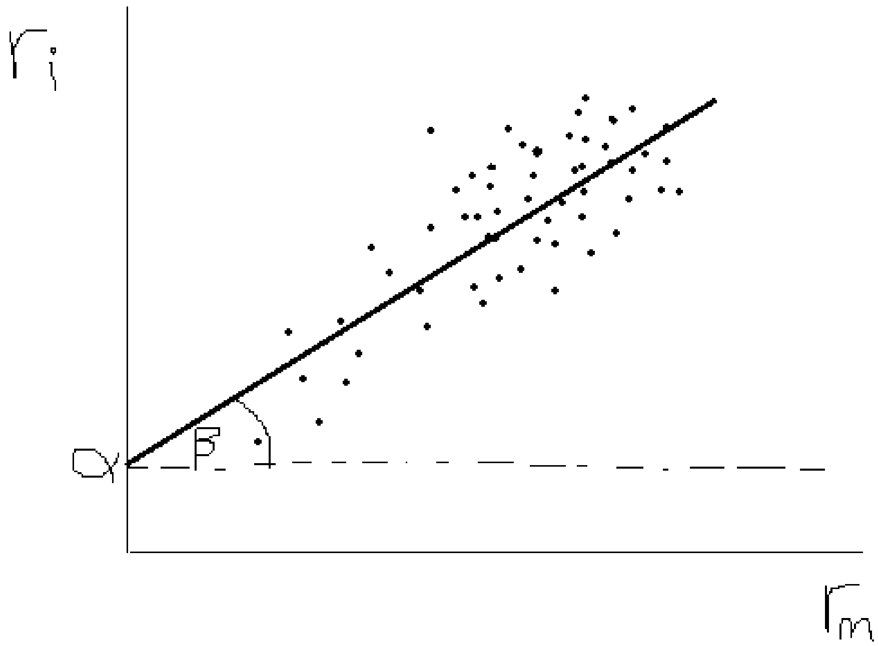
Where: Ri =The return of a certain stock.
Rm = The return of the entire market.
Beta = The sensitivity of a stock price to the market price.
Alpha = A residual (random value) who's expected return is 0, which means you can't leverage on that in order to generate excess return ("Generating Alpha").

2) The market we refer to in this model is the S&P 500.

Calculating the CAPM

The CAPM is calculated using a linear regression, where Beta is the slope of the linear regression, and Alpha is the intercept.

Basically the Y-Axis is the Ri, the return of a specific stock, and the X-Axis is the market return, now for each day we plot the combination of the market return and the stock return and we get a scatter chart, then we fit in a Linear Regression and this is the CAPM model for that stock.



Correlation

Correlation Coefficient determines how close the points to each other are, how well the two assets do goes together.  This is DIFFERENT from beta that is only the slope of the regression, there could be some stocks that have the same Beta but their correlation with the market is different.

CAPM vs Active Portfolio Management

The CAPM assumption is that Alpha is a random variable who's Expected value is zero and excess return is determined by Beta. If the stock Beta is less than 1 it means that the return will be smaller than the market return, and if the Beta is > 1 it means that you take higher risk than the market risk.

On the other hand, Active Portfolio Management claims that one can use different methods in order to take advantage of that Alpha and provide excess return over the market without necessarily increasing Beta. Alpha is not totally random in this theory and if you have a forecast for Alpha then you can exploit it to build a portfolio who's return is better than the market.

Portfolio

The return of a portfolio is:

$$r_p = \sum w_i * r_i$$

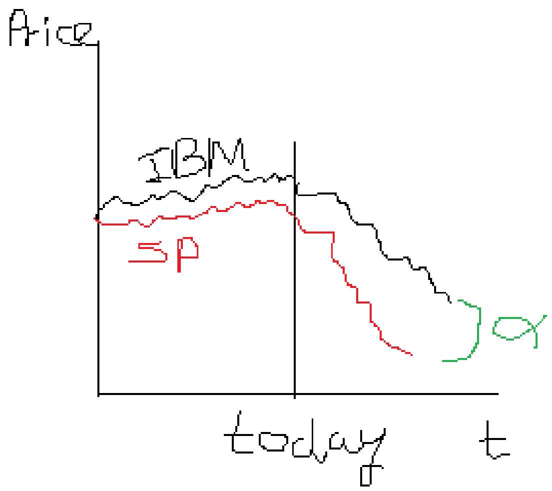
Where: Rp is the portfolio return.
Wi is the weight of the stock I in the portfolio.
Ri is the return of stock i.

On the same way the total Beta of a portfolio is:

$$\beta_p = \sum w_i * \beta_i$$

CAPM Market Risk – Example

Let's say we have a forecast that IBM is going to go up, and the chart ended up looking up like that:




So we can see that IBM and S&P 500 are going in the same direction where IBM is outperforming the S&P a little bit, and on the day we entered the market (with our information that IBM is going to go up) the market drops, but note that IBM drops LESS than the entire market (S&P).

Using CAPM we have a way to use our IBM information and still make money even thought the market is going down, we LONG IBM and SHORT the S&P.

For this example we are going to use the SPY ETF, we assume the IBM Beta is 1 (SPY Beta is obviously 1 as this ETF follows the market exactly), and we are going to build an equally weighted portfolio.

So:

**R (IBM) = 0.5*Beta IBM + 0.5*Alpha IBM** (which we believe we have info on and is positive).
Since we short SPY its weight is negative 0.5
**R (SPY) = -0.5*Beta SPY + Alpha SPY** (which we don't have info on and we assume on average is zero).

So giving the assumption that both Betas are 1 and Alpha of SPY is 0, we can sum up the return of both components: **R (IBM) + R (SPY) = 0.5*1 + 0.5 Alpha IBM + (-0.5*1) + 0 = <u>0.5*Alpha IBM</u>**.

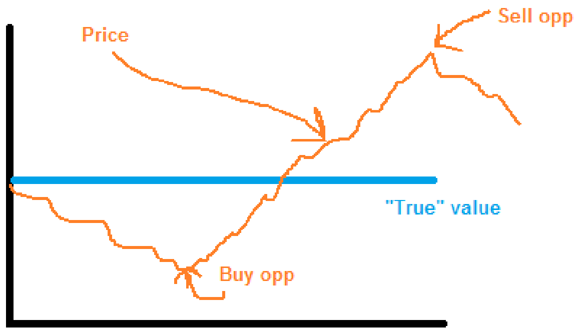
And since we believe IBM Alpha is greater than zero we can make money no matter what the market and IBM stock are doing (we basically invest only in IBM's alpha).
This is how we remove the market risk.

## **Week 3** – Was only about NumPy.

## **Week 4** – 03/17/2013

**<u>Information and Efficient Market Hypothesis</u>**

Most models are based on some arbitrage modeling, we have our “true” value of the equity (or at least what we think is true value), we assume that the market should be trading around that true value, and when price diverge from that true value we get a buy/sell opportunity, expecting the market to come back to the “true” value.




<u>Assessing the “true” value</u>

There are different sources of information that help assess stock’s value, the most common are:

1) Technical Analysis – Price & Volume
2) Fundamental Analysis – The company reports (P/E, Cash, Dividends etc).
3) News – Those are exogenous events that affect the stock price.

<u>Efficient Market Hypothesis</u>

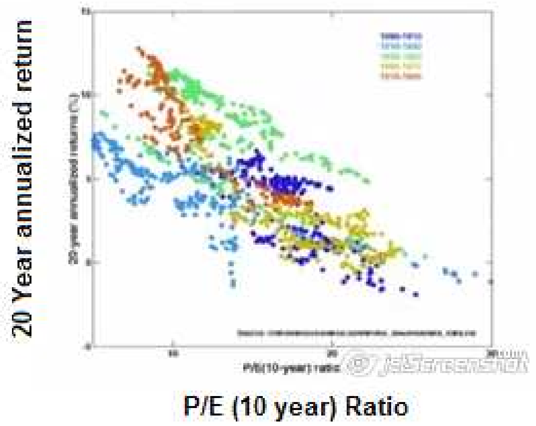
The main idea of that hypothesis is that there is no difference between the stock’s “true” value and its price, hence there is no real arbitrage opportunity. The “true” value of a stock is not a flat line (like the draw above), but rather is moving all the time and price match that value continuously.

There are 3 versions of that hypothesis:

1) Weak – Prices reflect all **past** publicly available information
   a) This prohibits any technical analysis value as all past information (like price & value) are already priced in.
   b) Fundamental Analysis does have some value in this approach as NEW information has not been priced in yet and can be exploited.
2) Semi Strong – like weak but adds that price **instantly** change to reflect **new** public information.
   a) Prohibit Technical Analysis.
   b) Prohibits Fundamental Analysis value as well since NEW information is instantly gets priced in.
3) Strong – like the Semi-Strong but adds that prices instantly change to reflect **FUTURE** information like “inside” information etc.
   a) This prohibits ANY information from being valuable, include inside information.

Is the EMH True?

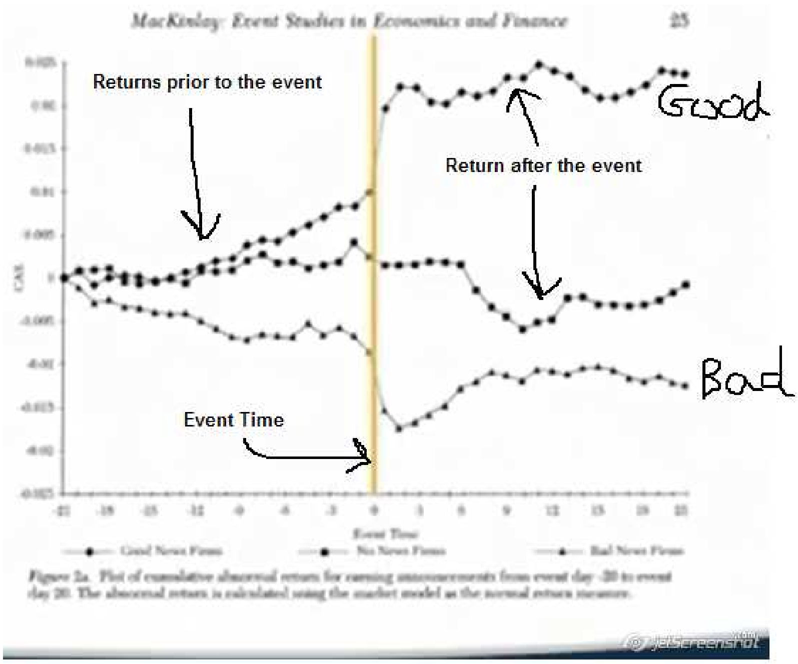
There are some evidences it is NOT. Robert Shiller in Irrational Exuberance showed there is inverse relation between the P/E ration and future returns, meaning a low current P/E ration resulted higher rate of return in the future, this gives some value to Fundamental Analysis.




A new argument against EMH is Behavioral Economics which says there are cognitive biases that affects the stock price and those can be exploited (like over-reaction to news etc).

**Events Studies**

AC MacKinlay had published a paper on how information affects price over time, what he did was analyzing ~1700 stocks and documented the date of Positive news event, Negative events and Neutral events, and monitored how news affected price. Whenever there was an event he looked what was the future return for that stock, and then plotted the mean return for all stocks and all events.



Notice the returns BEFORE the event, note how it goes up for positive events and how it goes down for negative events, one of the possible reasons is that there is some sort of leakage before the news is published and some participants take advantage of that unpublished info.

Also note the spike on the day of the event and then it balances (for negative event it's actually recovers first, but that's just a very general behavior).

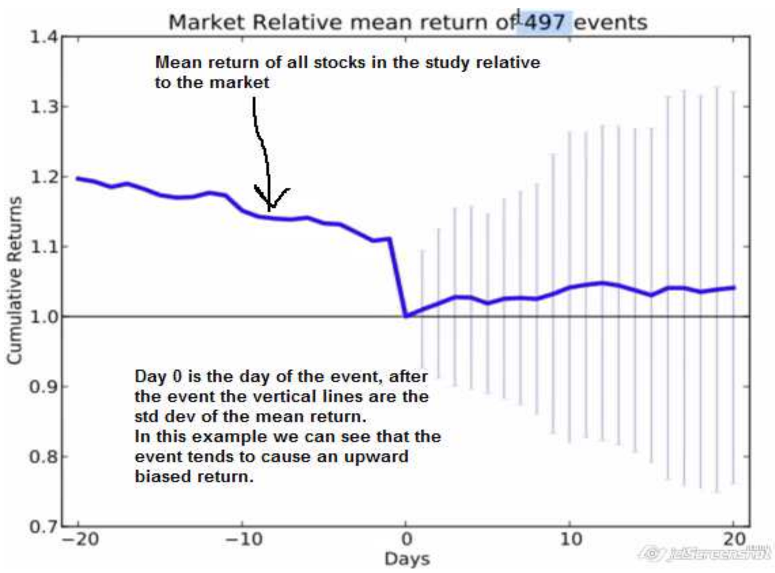
Event Profiler in QSTK

The event profiler is described in Tutorial 9.
Basically the event profiler gets a matrix of Dates/Symbols that contains for each Date/Symbol combination whether or not there was an event, 1 is True and NaN is False.
What is event? Event could be anything you like, fundamental news or a technical event like moving average cross etc.

Arguments passed to the event profiler:

1) The event matrix (Pandas DataFrame)
2) The data DataFrame (the data frame that contains the raw prices for each symbol)
3) Lookback – the number of days BEFORE the event to plot the mean return.
4) LookForward – The number of days AFTER the event to plot the mean return.
5) FileName – The output chart file.
6) MarketNeutral – If true it plots the returns relative to the market, meaning subtracting the market return from the stock return (assuming Beta = 1), NOTE that when the look forward returns are positive it is relative to the market, meaning absolute return can still be negative, but less negative than the market.
7) ErrorBars – whether or not plot the std dev of returns.
8) Market Symbol – The symbol that represents the market (for Market Neutral calculations).

Important to remember is that the event profiler will IGNORE events that happened "too" early or "too" late because there is not enough look back and look forward data.
Here is a sample output for the event profiler.



**Portfolio Optimization and the Efficient Frontier**

Portfolio optimization is finding the "best" (i.e minimum risk) allocation to a set of stocks (or combination of stocks) that will achieve the target return.
What do we mean by Risk? See week 1 (Std dev of daily return , max draw-down etc).
So now we can classify equities into 2 dimensions, return & risk:



Usually the greater the risk (std dev of return) the greater the return, so it is quite "difficult" to be on the upper left corner of the chart, now a portfolio can be put on that 2 dimensions (draw on the right).

Using different allocations we can control where the portfolio would be on the Return/Risk space.
Nobel Prize winner, **Harry Markowitz**, is the developer of the Mean Variance Optimization method which allows us to achieve a portfolio whose risk is lower than any of parts.

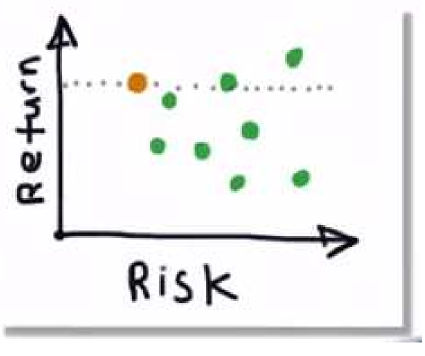
The Input and Output of a Portfolio Optimizer

Inputs:

1) Expected return for each equity – You need to provide some estimation of the future return for each equity; this can be done using different methods you believe in (technical, fundamental etc).
2) Expected Volatility (Risk) for each equity – this can be done by analyzing past volatility.
3) Target return – what's the return we want for this portfolio.
4) Covariance matrix – the correlation of each stock to the others.

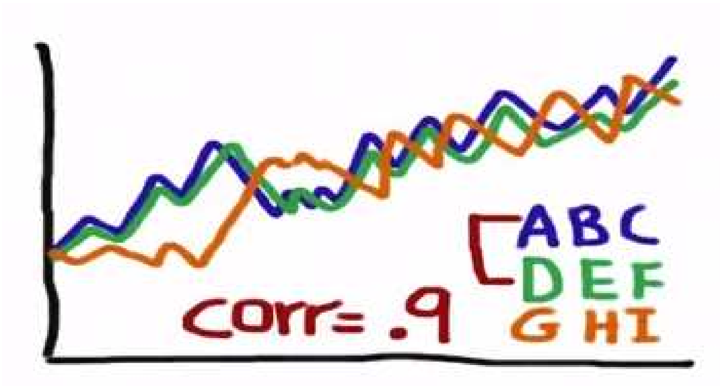
Output:

A set of weights we should apply for each stock in the portfolio. The result would be a portfolio with lower risk than any of its parts.

The importance of Correlation and Covariance

The key factor than enables the portfolio optimizer reduce the risk is by using the covariance matrix, i.e exploiting the correlation (or un-correlation) of the various daily returns of the equities.
In the figure below you can see that ABC and DEF are quite correlated with correlation coefficient of 0.9, while ABC and GHI are quite un-correlated (when one goes up the other goes down).



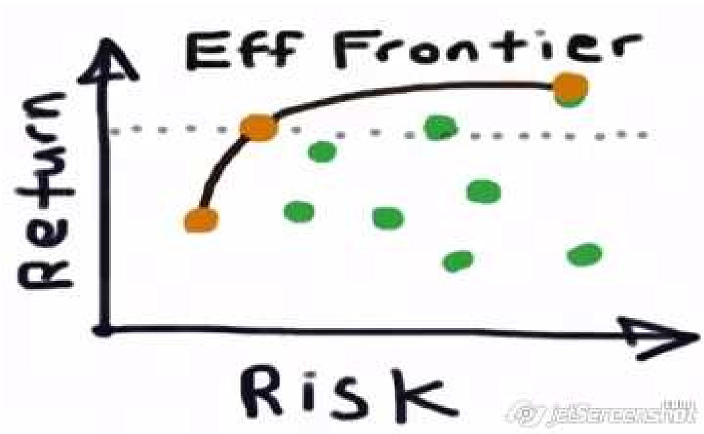

Now consider we build a portfolio of 50% ABC and 50% DEF => the result would be a portfolio that moves similarly to the equities themselves.

But now assume we allocate 50% to GHI and 25% to ABC and DEF each, the resulting portfolio would be a smoother line (less volatility) as the anti-correlation smoothes the moves in the opposite directions (when ABC and DEF goes down and drag to portfolio down, GHI would move up and drag the portfolio up).

The Efficient Frontier

Remember that the portfolio optimizer guarantees to find the lowest risk for a given expected return; that is, finding the allocation that would minimize the risk while keeping the return the same.

Now instead of leaving the target Return as constant we could change it, and for each target return we would get a different combination that would give the lowest risk for that return, the combination of all those target returns and lowest risk is the Efficient Frontier.
NOTE: It is NOT guaranteed that the lowest risk portfolio has positive return!!



How an optimizer works?

QSTK has a built-in optimizer in it, more details are in tutorial 8.

General Idea of how optimizer works:

1) Define variables – those are "tweakable" variable that the optimizer plays with in order to get to the best result. In our case these are the Equity Weights (and target return).
2) Define Constraints – Those are constraints on the tweakable variable, for example the sum of all weights should be 1. Or minimum/maximum weight per equity etc.
3) Define optimization criteria – A function that gives a value for each set of variables, in our case it is the "risk"

Optimizer Algo:

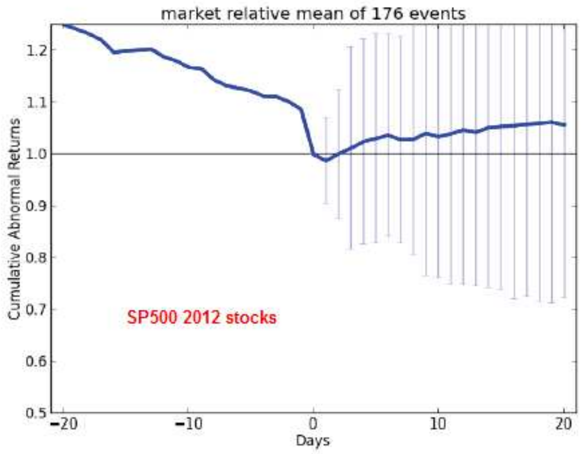
1) Tweak weights.
2) Check constraints.
3) Get a "Score" for this combination.
4) Repeat until it finds the best score. Now this can be done by looping thru all available combinations (like we did in HW1, or use some other sort of algorithm).

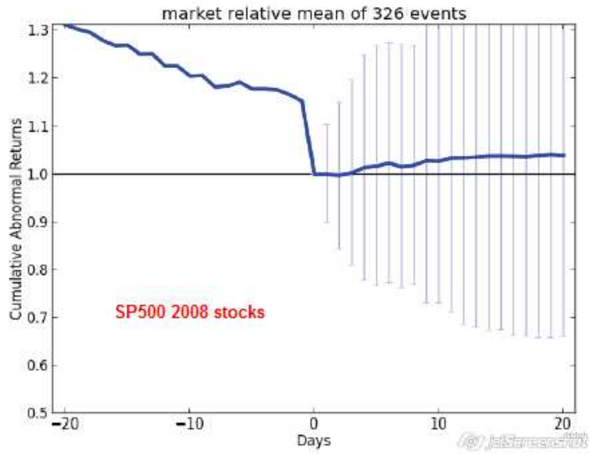
QSTK uses the CVXOPT optimizer (Convex optimization), open source developed by UCLA. Basically this algorithm assumes that a local maximum is also the global maximum, so once it started to go in a certain direction and finds a maximum it stops as this is also THE highest maximum possible.

## **Week 5** – 24/03/2013

**Survivor Bias**

In HW2 we used 2 lists, one is the SP500 components as of 2008, and the other the SP500 components as of 2012, and our study period was 2008-2009. We got different results just by using different list of stocks, why is that?




We can see that the list of stocks as of 2012 produced better results ON THE SAME TEST PERIOD (2008-2009), the reason for that is that by using the 2012 list we guarantee to include only stocks that are alive today, we don't include all the stocks that went down to 0, which is the case when we use the 2008 list, some stocks there just disappeared (62 equities, some acquired and some died).

What can we do?

Use data that is Survivor Bias free, meaning use BOTH stocks that are alive and that are dead, so your data is not only the "survivals" stocks which tend to skew up the results.
One of the issues with using companies that have died is database indexing. Usually a database is indexed by symbol which we tend to think of as unique, this is true only for a specific date, when we use historical data symbols tend to get recycled by different companies (for example JAVA was used by Mr Coffee company, and then by SUN, which used the symbol SUNW before etc). So when we index the database it is important to use other thing than just symbol (like company id or something like that).

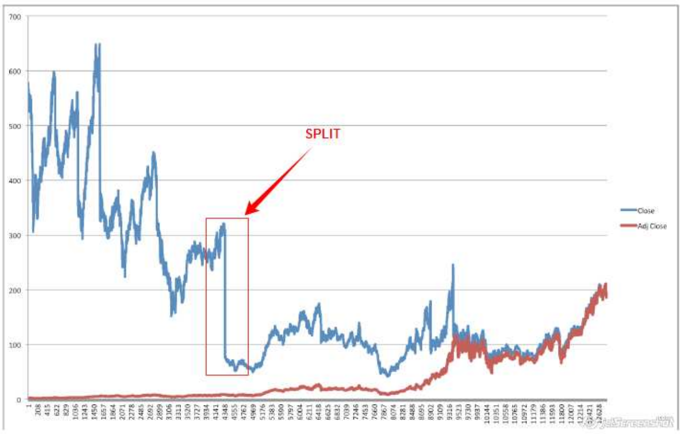
Another possibility is to create several random portfolio of survival stocks, which we know tends to perform better over time, and average those portfolios returns, then look for a strategy that can beat this average, then it means the strategy might be good (like instead of using the performance relative to the market (SPY) test it relative to that averaged random portfolios).

But it is very important that your strategies work with companies that failed, there might be patterns there you'll discover and want to avoid.

Actual vs. Adjusted Close

Actual – This is the raw price that was on that day (“what was on the paper that day”).
Adjusted – A revised price that count in dividends payments (which gives you more money had you held the stock), and account in for split or reverse split.

Split – This is when a company decides to increase the amount of stocks that are available, so for example a 1:2 split means that if you held 1 share of XXX now you hold 2 shares, and since there was no change to the market cap of the company (market cap=number of shares * last price), we have to cut price in half in order to maintain the same market cap.
So this split would result a sharp drop in the Actual Price, but for the Adjust Price the transition would be smooth, this is because the adjustment now goes and cut all historical prices by half. Over time the very historical data gets adjusted many times which usually results the adjusted price to be very small, see in the chart below of IBM the Actual vs Adjusted in the beginning (~50 years ago):




Dividends – Over time companies tends to pay dividend, what happens after the close on the dividend day the exchange subtract the dividend amount from the closing price so the next day the stock would open lower just because of that subtraction (this is done in order to reflect the change in the company’s market cap due to the dividend payment).
Adjusting for dividends would smooth out that artificial price drop which in hand would reflect the fact that investors got “richer” by the dividend amount.

Gaps in Data (NaN)

It is possible that there is no data for some stocks in the dates your are querying (like the stock didn’t exist or died or just couldn’t trade due to regulations etc).
What should we do? In Pandas there is a function to “Fill Forward” and “Fill Backwards”.
Fill Forward – Take the last price that was available before the NaN and push it forward instead of the NaN.
Fill Backward – Take the very first price after the NaN and push it backwards instead of the NaN.
We want to use fill FORWARD, the reason is that we don’t want to “peak” into the future, if we do a backward fill we actually take a FUTURE price and push it back before it was actually existed.
After we do a Forward fill we might still have NaN if the very first dates are NaN (say the stock didn’t exist), so only there we do a Backfill, this is somewhat of “peaking” as well but since it is on the start and the price is flat it’s not that bad.
**So Forward Fill and then Backward Fill.**

## Data Sanity and Data Scrubbing

Historical data might contain some errors that are a result of miscalculations, like fail to adjust for splits or dividends, orders of magnitude drops followed by offsetting orders of magnitude climbs, missing data of dates/symbols, etc.


Why bad data is bad? Automated strategies may exploit bad data and then fail on the real data. All simulation calculations would be wrong.

### Sanity Checks

1) When you download historical data scan it for big spikes for the upside/downside.
2) Look for missing data (NaN).
3) Prices that are under a penny.

### Scrubbing

It is easiest to remove the bad data, but it is better if you can get multiple sources of data and complement the bad data with good one.

**<u>Week 6</u>** – 31/03/2013

**<u>The Fundamental Law of Active Portfolio Management</u>**

Richard Grinold, in his “Active Portfolio Management” book, offers this fundamental law where it helps up make quantitative decisions like “should we expand efforts make our forecast better, or should we expand effort making more trading opportunities”, both of these are important to make overall better performance.

<u>Coin Flip Experiment</u>

Betting on a coin flip is like making a trade, when we make a trade decision we assume we have a tiny knowledge, or expectation, about the trade outcome; this is like having a biased coin to one side.
So in our experiment we assume that the coin has 51% chance of getting “head”, so the <u>Uncertainty</u> is what we don’t know, this is like the whole market, and that is Beta, while the coin bias, what we do know, is Alpha.
Uncertainty = Beta
Coin Bias = Alpha (51% heads).

<u>Single vs Multiple bets</u>

Assume our account size is 1000$.
Case 1:
We make 1 bet for 1000$, if we win we end up with 2000$, if we lose we end up with 0$.

Case 2:
Make 1000 bets for 1$ each, when we win we get 1$, when we lose we give 1$.

Let’s compute our <u>Reward</u> for each case:
Case 1 (Single bet) – Expected Return = Reward = 51% * 1000$ + 49% * (-1000$) = 20$.
Case 2 (1000 bets) – Expected Return = Reward = (51% * 1000$ + 49% * (-1000$)) * 1000 bets = 20$.

Let’s compute our <u>Risk</u> for each case:
We define Risk as the possibility we would lose the entire account.
Case 1 (Single bet) - We have 49% losing everything.
Case 2 (1000 bets) – We need 1000 losers in a row to lose everything = 49%*49%*49%.... = $49\%^{1000}$

Another measure of <u>Risk</u> is the standard deviation of return, for the multiple bets case (case 2), the returns are either +1 or -1 so:
STDEV(1, -1, 1, -1, 1, 1, ….., -1) ≈ 1$

On a single bet case (case 2) it is not really possible to compute the STDEV as it is a single event, so we assume the returns are either (1000$, 0, 0, 0,….., 0), or (-1000$, 0, 0, 0, …., 0), so:
STDEV(1000, 0, 0, ….,0) ≈ 31.62$

Finally let’s compute the Reward/Risk measure (Sharpe Ratio):
Case 1 (single bet): 20$ / 31.62$ = **0.63**
Case 2 (multi-bet): 20$ / 1$ = **<u>20</u>**

More generally (there is no proof given for that formula):
**Sharpe Ratio = C * alpha * SQRT(bets)**

C – Is some coefficient
Alpha – Is some knowledge we have about the bet, in our case is the 2% bias towards "heads"

So you can improve your Sharpe if you increase your Alpha or increase the number of bets.

Overall when we compare the Single case vs Multi bets case we see that:
* We have the same expected return, but
* On Multi bet case:
    * Much lower risk to lose everything.
    * Much lower standard deviation of return.
    * Much higher Reward/Risk ratio.

<u>The Fundamental Law components</u>

Let's compare at two well-known funds mangers, Warren Buffet and Jim Simons. Buffet is very selective and have very small number of companies in his portfolio, in 2010 54% of the fund was in only 3 stocks. On the other day Jim Simons fund makes 100Ks of trades a day.

The overall performance of both funds is more or less the same, where Jim Simons have better performance.

Recall the CAPM model:
$r_i = \beta_i * r_m + \alpha_i$

The Alpha is the added value of each fund manager to the overall performance, this is due to skill.
When we look at the RISK of the CAPM model we can break it into two parts:
Part 1 is the STDEV of the Beta, which is the <u>Market Risk</u>.
Part 2 is the STDEV of the Alpha which is the <u>Residual Risk</u>.

- Information Risk

  $$IR = \frac{Mean(\alpha)}{Stdev(\alpha)} = \frac{E(\alpha)}{\sigma(\alpha)}$$

- Information Coefficient
  IC = Correlation of forecast to actual return, how much time does the fund manager is "right".
- Breadth
  BR = Number of opportunities to execute.

<u>The Fundamental Law</u>

$$IR = IC \times \sqrt{BR}$$

Where we can say that, IR is the overall <u>performance</u>, IC is the manager <u>skill</u>, and BR is the number of <u>opportunities</u> your methodology produces. So there is a tradeoff between Skill and number of trades.

So let's go back to the comparison of Warren Buffet and Jim Simons funds. Warren Buffet has very few investments but on each one he conducts a deep analysis and as a result has high alpha. Reason that Buffet does not have more investments is because his approach does not scale, he needs to perform a deep and long analysis and he can't do it for a large number of companies.

Jim Simons on the other hand, has a computational approach that scales easily, once the criteria of the methodology is defined it can be applied to all stocks that are traded on the exchange, this approach has lower alpha but generated much more trading opportunities.

So using the fundamental law we know that performance is the product of IC (Skill) and BR (opportunities), and that gives us the two ways to succeed: <u>Skill</u> and <u>Breadth</u>.

## CAPM for Portfolios

The expected return on a single asset is depended on the overall market return plus some residual (skill).

$$r_i = \beta_i * r_m + \alpha_i$$

When we come to extend the CAPM for a portfolio we let:

$$h_i = \%\ holdings\ in\ i$$

$$r_p = \sum_{i=1}^{n} h_i \times r_i$$

Where R(p) is the portfolio return and R(i) is the individual asset return.

Example:

h(1) = 0.25, b(1) = 3

h(2) = 0.75, b(2) = 1

r(p) = h(1) * [b(1)*r(m) + a(1)] + h(2) * [b(2)*r(m) + a(2)]

= h(1)*b(1)*r(m) + h(2)*b(2)*r(m) + h(1)*a(1) + h(2)*a(2)

= **1.5*r(m) + 0.25*a(1) + 0.75*a(2)**

This is our portfolio total expected return.

Our goal in trading is finding stocks with higher alpha and minimizing our exposure to the market as this is something we don't have control on.

Reducing Market Risk

Assume we invest in 2 stocks with the following characteristics:
Stock 1 – Beta = 2, and we think it will underperform the market by 2% (alpha = -2%).
Stock 2 – Beta = 1, and we think it will outperform the market by 2% (alpha = +2%).
We SHORT stock 1 and LONG stock 2 with an allocation of 50%/50%.

Now assume the market went UP 10%, since stock 1 has a Beta of 2 its return was +20% but let's assume our prediction of underperformance of 2% was correct, so the overall return of stock 1 is +18%, and we shorted it.
Stock 2 has beta of 1 so it went up with the market 10%, but assume our prediction of alpha +2% was correct, so the overall return of stock 2 is +12%, we longed it.
So the overall performance of our portfolio is: 0.5 * (-18%) + 0.5 * (+12%) = -3%

So although our prediction was correct we lost money! Why? Because of the high Beta of stock 1, the correlation with the market (Beta) overwhelmed our Skill (Alpha), what can we do? Using the portfolio CAPM:

$$r_p = \sum_{i=1}^{n} h_i \times r_i$$

We can choose a combination of weightings (h) that will zero out Beta:
b(portfolio) = h(1)*b(1) + h(2)*b(2)
So let's assume now we SHORT stock 1 (negative holding) with holding of 33% and LONG stock 2 with holding of 66%, so the total portfolio bets = (-0.33)*2 + 0.66*1 = **0**.
So now our portfolio expected return = (-0.33) * a(1) + 0.66 * a(2)

Going back to our example, the market went up 10% so stock 1 went up 18%, and stock 2 went up 12%, our total return = (-0.33)*18% + 0.66*12% = 1.98%.

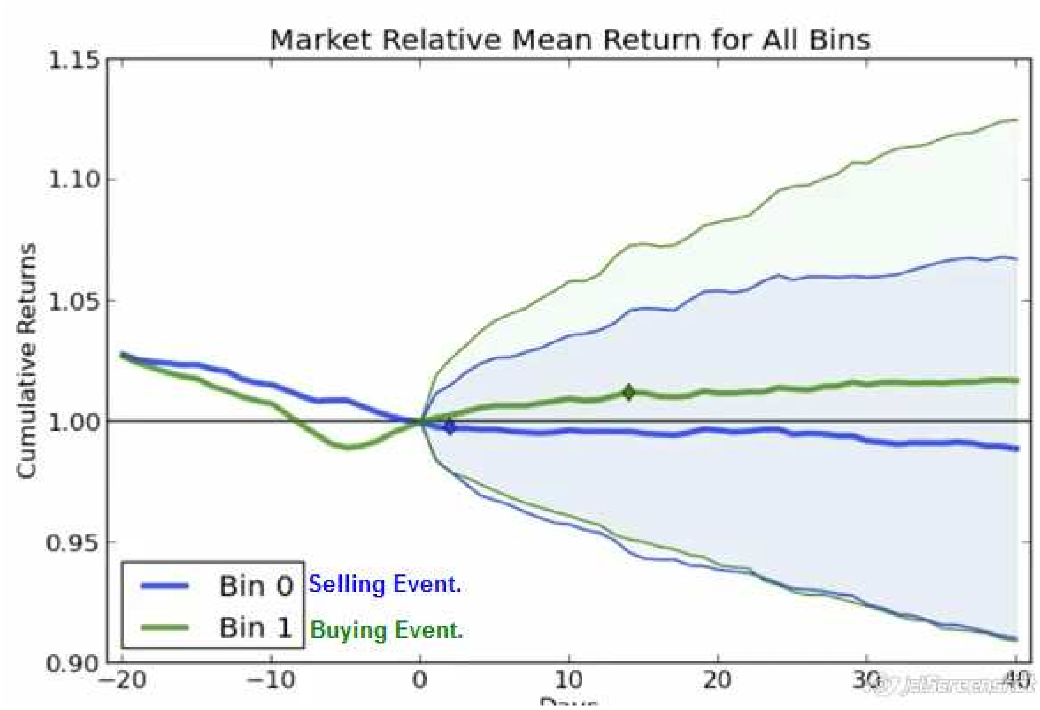
## Week 7 – 07/04/2013

**Information and Technical Analysis**

### Information Feeds

There are different feeds providers, these are example of 3 information feeds the professor has worked with:

1) Thomson Reuters Machine Readable News – An ultra-low latency feed of xml-based news, the news are tagged as either positive/negative/neutral, so that can be used for events studies.
2) StarMine – This feed delivers Assessment of Analysts, they have a predictive signal of future changes in analyst sentiment.
3) InsiderInsights – A daily feed of all insider buying/selling activity, they also add their own assessment of the importance of that buying/selling.
   Here is an event study done on InsiderInsights data:




What to look for when looking for information feed?

1) Historical data - so you could run several strategies back test.
2) Survivor bias free – Should include stocks that are dead.
3) Ease of integration – What is the structure of the data, does it fit your system.
4) Low latency – How quickly does the information get to you?

What is Technical Analysis (TA)?

- TA uses historical price and volume data ONLY to compute "indicators".
- Indicators are calculated from recent prices and volume in order to predict future prices. Indicators can be calculated by using either daily data (I.e End of Day data), or intra-day data (i.e every second/minute/hours etc).
- Indicators are "Heuristics", meaning that although they are estimates built using quantitative information it is built from a person insight of how he thinks the market works.
- Are indicators information source? Does it have any added value in predicting future price moves? Maybe…

TA is very controversial; many believe it is "voodoo" and "superstition". TA depends on information in historical price and the fact that this information can be exploited (the markets are not fully efficient).
The main reason technical analysts believe TA works is that TA is only a proxy to the Psychology of Investors which drives the markets. Another reason that make TAs believe TA works is what they call "Markets Physics", there are some basic "rules" controlling price moves (like there is a limit of how quick a price can move, price that moved "too far" etc).

The professor view

"I just look at the evidence", from several studies the Prof and his students conducted he thinks that individual indicators are not predictive by themselves, they were predictive in the 80s & 90s but not anymore as the market is more efficient today.

However, from studies done that combined several indicators together he found some predictive power.

Why it might work? There are lots of people who pay attention to different indicators, and if they act on it this can become a "self fulfilling prophecy".
**The best case** of using TA is when a stock's indicators are contrary to the market, like when the market indicators are pointing down and a stock's indicators are pointing up there is some potential value there.

For further study of technical analysis see "Technical Analysis Explained" by Martin Pring.

Pring's definition of TA: "Identify a trend reversal at a relatively early stage and ride on that trend until the weight of the evidence shows or proves that the trend has reversed".
The assumption is that "people will continue to make the same mistakes they have made

in the past". It's the people's attitudes and emotions that cause the market inefficiencies which TA tries to exploit.

Pring divide the indicators into three "branches":

1) Sentiment indicators – "Emotions of investors"
2) Flow-of-Funds indicators – How much cash has been flowing "in" or "out".
3) Market Structure Indicators – These are based on price and volume.

<u>Time Scales and Trend Analysis</u>

(The professor opinion):

- Short term price movements is driven by market mechanics and can be predicted by tick and order book information, but in order to act on it you have to be quick.
- The longer term price movement is driven by exogenous events and is not always observable by computer; this is where you need human analysis.

<u>Indicators Examples</u>

The main "theme" for technical indicators is that there is some real value for an asset and emotions and market structure cause price to deviate from the "true" value.
If we can estimate the "true" value we can find arbitrage opportunities.
Those indicators are also called as "Mean Reversion" indicators.

Now there is a review of the following indicators (pretty basic info nothing special):

1) Simple Moving Average
2) MACD
3) Bollinger Bands

Each indicator has its own values and interpretation, in order to be able to easily use them in events studies or Machine Learning it is recommended to Normalize their value into a value between -1 and +1, that way it is much easier to perform studies on a collection of indicators (for example in Bollinger Bands "-1" is when price touches the lower band and "+1" is when price touches the upper band).

**Week 8** – 14/04/2013

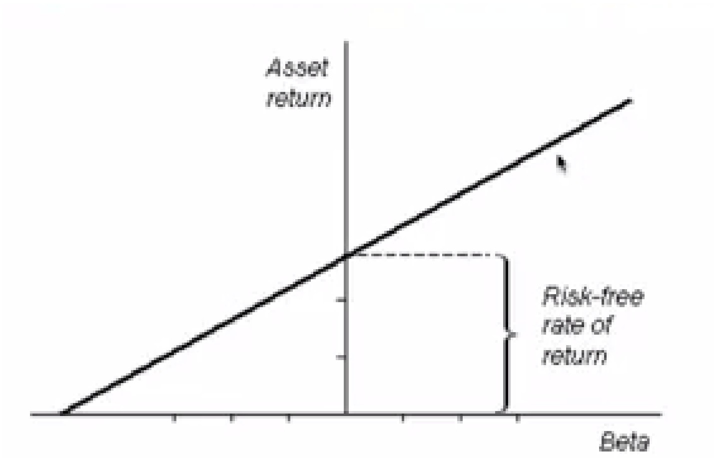
**Jensen's Alpha**

This measure was created by Michael Jensen in order to measure fund managers performance. Some of the funds were invested in high beta stocks, which were volatile and when there were higher returns it was due to the fact of the high beta and not necessary due to the manager skill. Jensen goal was to come with a measure of the manager skill (i.e. neutralizing the returns due to volatility).

CAPM suggests that:

$$E(R_i) = R_f + \beta_i\big(E(R_m) - R_f\big)$$

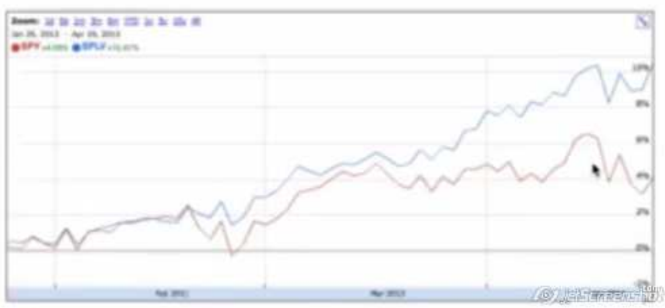
As we can see, if we have a positive expected market return, the return of our investment will be larger as beta is larger. The following chart shows the expected investment return (Ri) as function of beta assuming the expected market return is positive.




Now if we build a model out of the CAPM "expected" model we would get that:

$$R_i = R_f + \beta_i\big(R_m - R_f\big)$$

Now if we face this model with real results it does not add up exactly, there are (big) deviations between the real returns and the return the model suggests.
Let's have a look at some real results, here we have the SPLV and the SPY, SPLV is a low volatility ETF that tracks the SPY, as such its Beta is 0.75. According the model SPLV should have lower return (beta < 1) than the market, in reality SPLV (blue line) has higher cumulative return than SPY (red line).



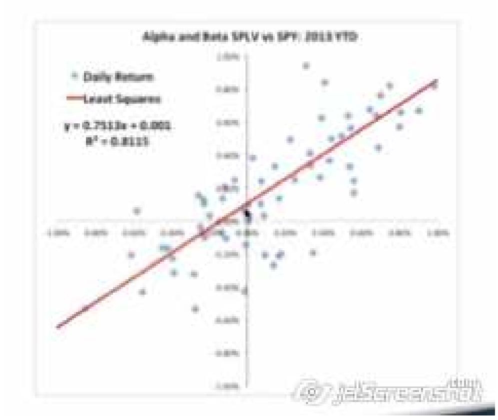
This different is explained by Alpha.

$$R_i = R_f + \beta_i(R_m - R_f) + \alpha$$

CAPM asserts that E(alpha) = 0, meaning alpha is random and expected to be 0.
Others assert that alpha is a measure of management skill.

In order to estimate Alpha (and Beta) we are using linear regression of SPLV vs SPY, the slope of the linear regression is Beta and the Y-Axis intersect is Alpha.
(Each dot represents the daily return of SPLV and SPY).

Note that Alpha could be less than zero ("bad" manager) but the fund could still have a positive return due to beta and market return - This is Jensen's Alpha which measure the fund manager added value.




**Back-Testing**

Back testing is the process of testing a trading strategy on prior time periods.
In back testing you are taking your strategy back in time and test its performance, it is a quick way to assess the strategy "effectiveness". The alternative method is "forward testing" which is starting to simulate trades from the current point in time forward, the problem with this method is that it requires a lot of time before you can get enough results and drive into meaningful conclusions.

<u>Risks in Back Testing</u>

1) The market is changing and the conditions that were during the back testing period is not necessarily the current market conditions, and it might affect its future results.
2) "Data mining fallacy" – The strategy performance was due to "luck" and not as a result of some fundamental reason, this could be due to some specific parameter that worked it the past and won't work now.
3) Over fitting – this is similar to the Data Mining Fallacy, once you find a strategy that works ok you are starting to fine tune the parameters until it fits the data perfectly and it will show great past performance, but it won't work in the future.

How to avoid it?
The best way to asses a strategy in back testing is to divide the data into 2 (or more) segments, for example you build the strategy using data from stocks symbols A-M, then you run the strategy on symbols N-Z and see if the results are the same.
Another approach is building the strategy on a subset of symbols AND subset of time (like A-M and 1990-2001), and then validate the strategy on the rest of the symbols (still only a subset of time 1990-2001), finally you validate the strategy over all the data.
If it still works well after all those validations then it is likely there is some fundamental advantage in this strategy and it's not just "luck" – This is called Cross Validation.

Components of back tester

1) Historical data – it is important to have good a clean data free of survivor-bias.
2) Strategy definition – Those are the rules that generates the buy/sell orders, you want a back tester that is easy in defining strategies.
3) Market Simulator – This part is responsible for simulating the actual buying/selling transactions, it is important the simulator will be accurate and include all transaction costs, commissions and market impact (A paper by Almgren et al 2005 suggests a model to estimate the price impact of your order based on the transaction volume and the stock daily volume).
4) Analysis Engine – This part is responsible for analyzing the overall performance of the strategy. It is important it will include all the important metrics including comparison with a benchmark.

# ----- THE END -----