Introduction to Algorithmic Trading Strategies
Lecture 7

Portfolio Optimization

Haksun Li
haksun.li@numericalmethod.com
www.numericalmethod.com

Outline

- Sharpe Ratio
- Problems with Sharpe Ratio
- Omega
- Properties of Omega
- Portfolio Optimization

References

- Connor Keating, William Shadwick. A universal performance measure. Finance and Investment Conference 2002. 26 June 2002.
- Connor Keating, William Shadwick. An introduction to Omega. 2002.
- Kazemi, Scheeweis and Gupta. Omega as a performance measure. 2003.
- S. Avouyi-Dovi, A. Morin, and D. Neto. Optimal asset allocation with Omega function. Tech. report, Banque de France, 2004. Research Paper.

A Universal Performance Measure: Omega

Notations

- $r = (r_1, \dots, r_n)'$: a *random* vector of returns, either for a single asset over n periods, or a basket of n assets
- Q : the covariance matrix of the returns
- $x = (x_1, \dots, x_n)'$: the weightings given to each holding period, or to each asset in the basket

Portfolio Statistics

- Mean of portfolio
  - $\mu(x) = x'E(r)$
- Variance of portfolio
  - $\sigma^2(x) = x'Qx$

Sharpe Ratio

- $\mathrm{sr}(x) = \frac{\mu(x) - r_f}{\sigma^2(x)} = \frac{x'E(r) - r_f}{x'Qx}$
- r_f: a benchmark return, e.g., risk-free rate
- In general, we prefer
  - a bigger excess return
  - a smaller risk (uncertainty)

Sharpe Ratio Limitations

- Sharpe ratio does not differentiate between winning and losing trades, essentially ignoring their likelihoods (odds).
- Sharpe ratio does not consider, essentially ignoring, all higher moments of a return distribution except the first two, the mean and variance.

Sharpe's Choice

- Both A and B have the same mean.
- A has a smaller variance.
- Sharpe will always chooses a portfolio of the smallest variance among all those having the same mean.
  - Hence A is preferred to B by Sharpe.

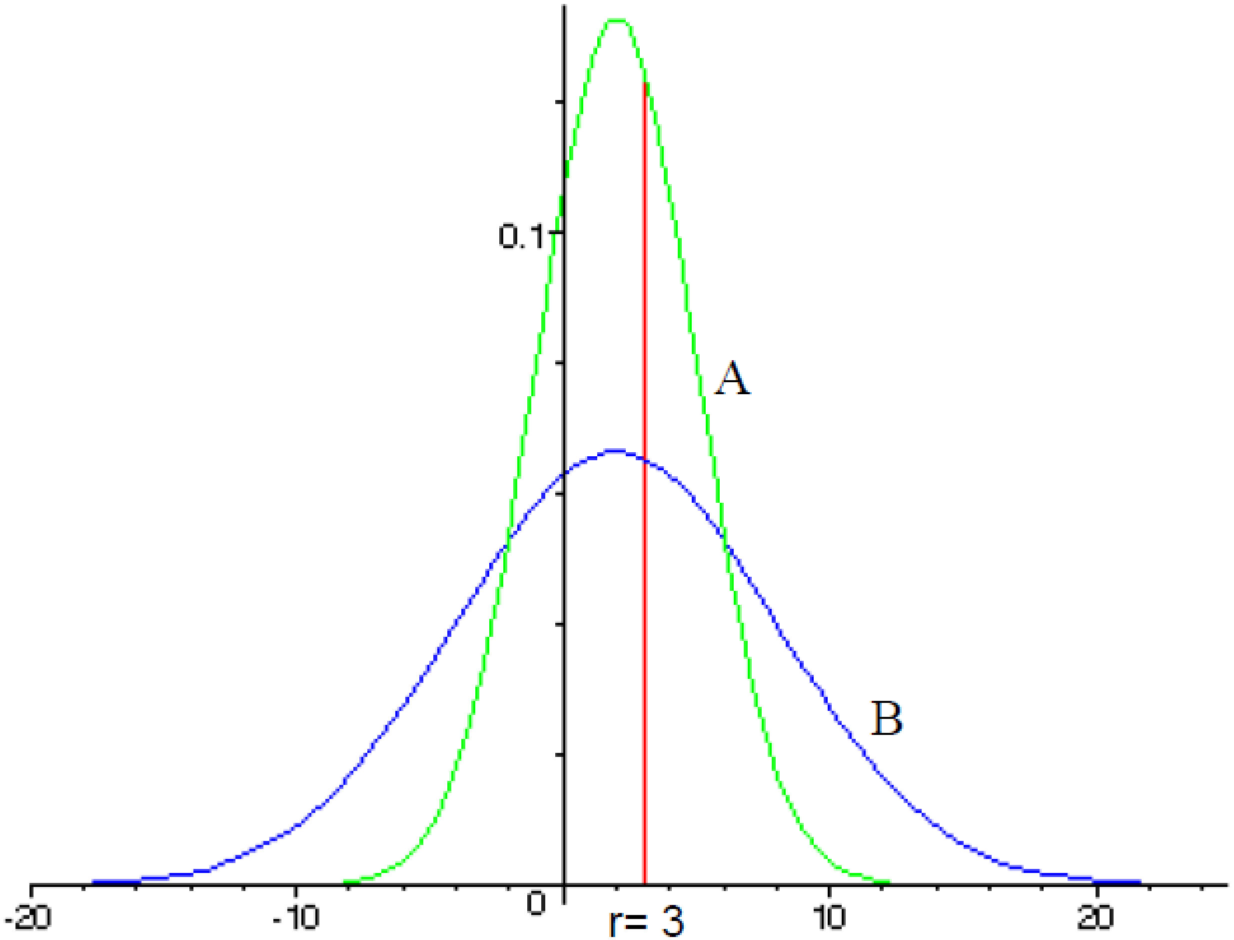
Avoid Downsides and Upsides

- Sharpe chooses the smallest variance portfolio to reduce the chance of having extreme losses.
- Yet, for a Normally distributed return, the extreme gains are as likely as the extreme losses.
- Ignoring the downsides will inevitably ignore the potential for upsides as well.

Potential for Gains

- Suppose we rank A and B by their potential for gains, we would choose B over A.
- Shall we choose the portfolio with the biggest variance then?
  - It is very counter intuitive.

Example 1: A or B?



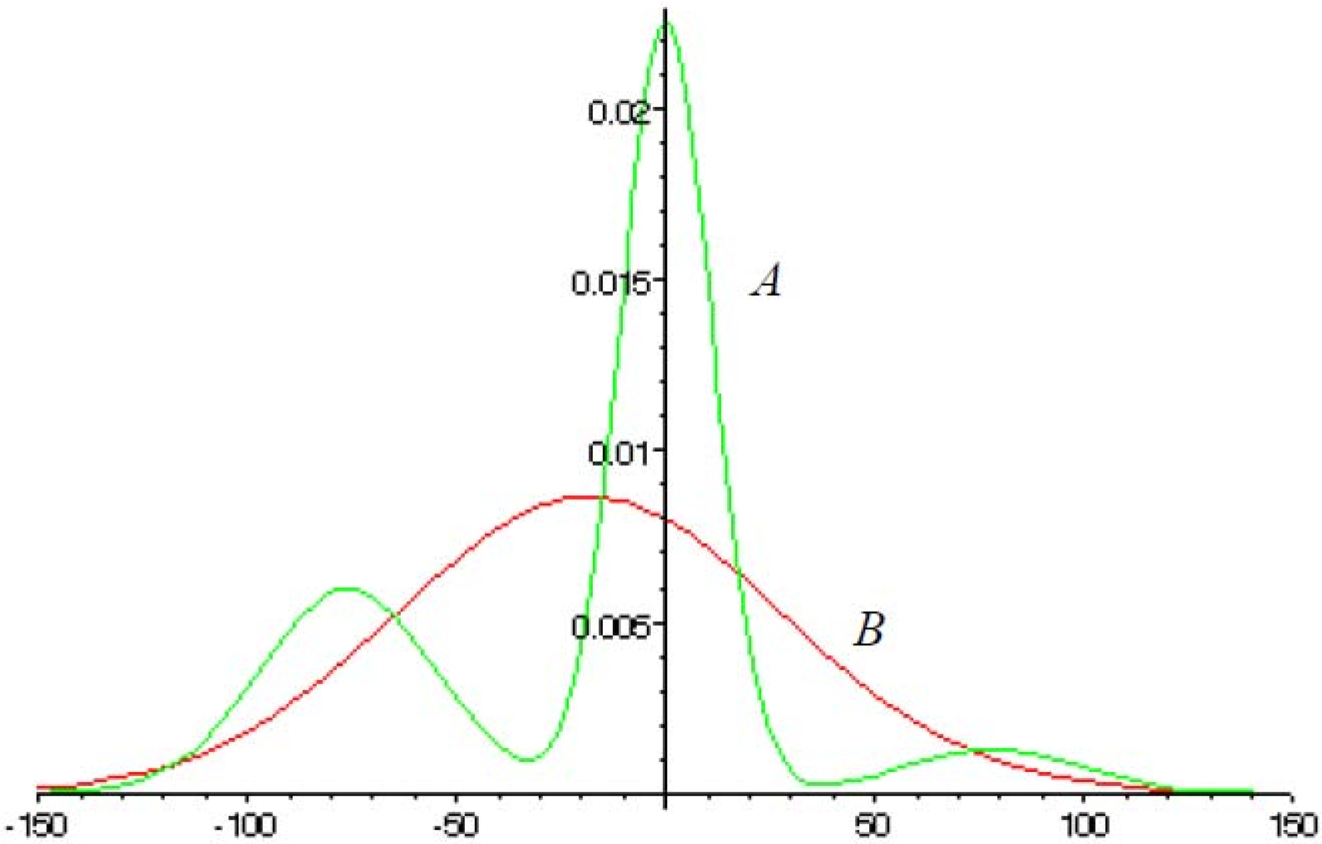
Example 1: L = 3

- Suppose the loss threshold is 3.
- Pictorially, we see that B has more mass to the right of 3 than that of A.
  - B: 43% of mass; A: 37%.
- We compare the likelihood of winning to losing.
  - B: 0.77; A: 0.59.
- We therefore prefer B to A.

Example 1: L = 1

- Suppose the loss threshold is 1.
- A has more mass to the right of L than that of B.
- We compare the likelihood of winning to losing.
  - A: 1.71; B: 1.31.
- We therefore prefer A to B.

Example 2



Example 2: Winning Ratio

- It is evident from the example(s) that, when choosing a portfolio, the likelihoods/odds/chances/potentials for upside and downside are important.
- Winning ratio $\frac{W_A}{W_B}$:
  - 2σ gain: 1.8
  - 3σ gain: 0.85
  - 4σ gain: 35

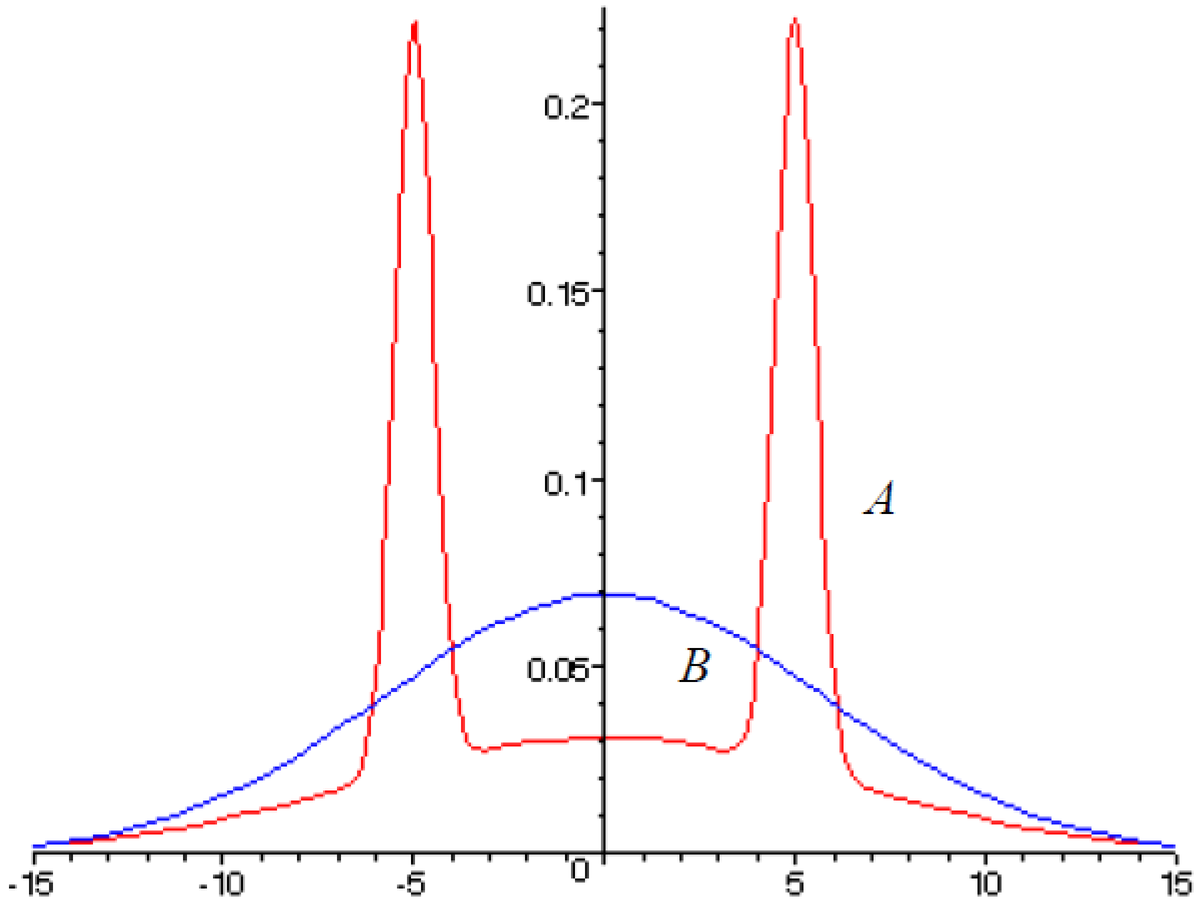
Example 2: Losing Ratio

- Losing ratio $\frac{L_A}{L_B}$:
  - 1σ loss: 1.4
  - 2σ loss: 0.7
  - 3σ loss : 80
  - 4σ loss : 100,000!!!

Higher Moments Are Important

- Both large gains and losses in example 2 are produced by moments of order 5 and higher.
  - They even shadow the effects of skew and kurtosis.
  - Example 2 has the same mean and variance for both distributions.
- Because Sharpe Ratio ignores all moments from order 3 and bigger, it treats all these very different distributions the same.

How Many Moments Are Needed?



Distribution A

- Combining 3 Normal distributions
  - N(-5, 0.5)
  - N(0, 6.5)
  - N(5, 0.5)
- Weights:
  - 25%
  - 50%
  - 25%

Moments of A

- Same mean and variance as distribution B.
- Symmetric distribution implies all odd moments (3rd, 5th, etc.) are 0.
- Kurtosis = 2.65 (smaller than the 3 of Normal)
  - Does smaller Kurtosis imply smaller risk?
- 6th moment: 0.2% different from Normal
- *8th moment: 24% different from Normal*
- *10th moment: 55% bigger than Normal*

Performance Measure Requirements

- Take into account the odds of winning and losing.
- Take into account the sizes of winning and losing.
- Take into account of (all) the moments of a return distribution.

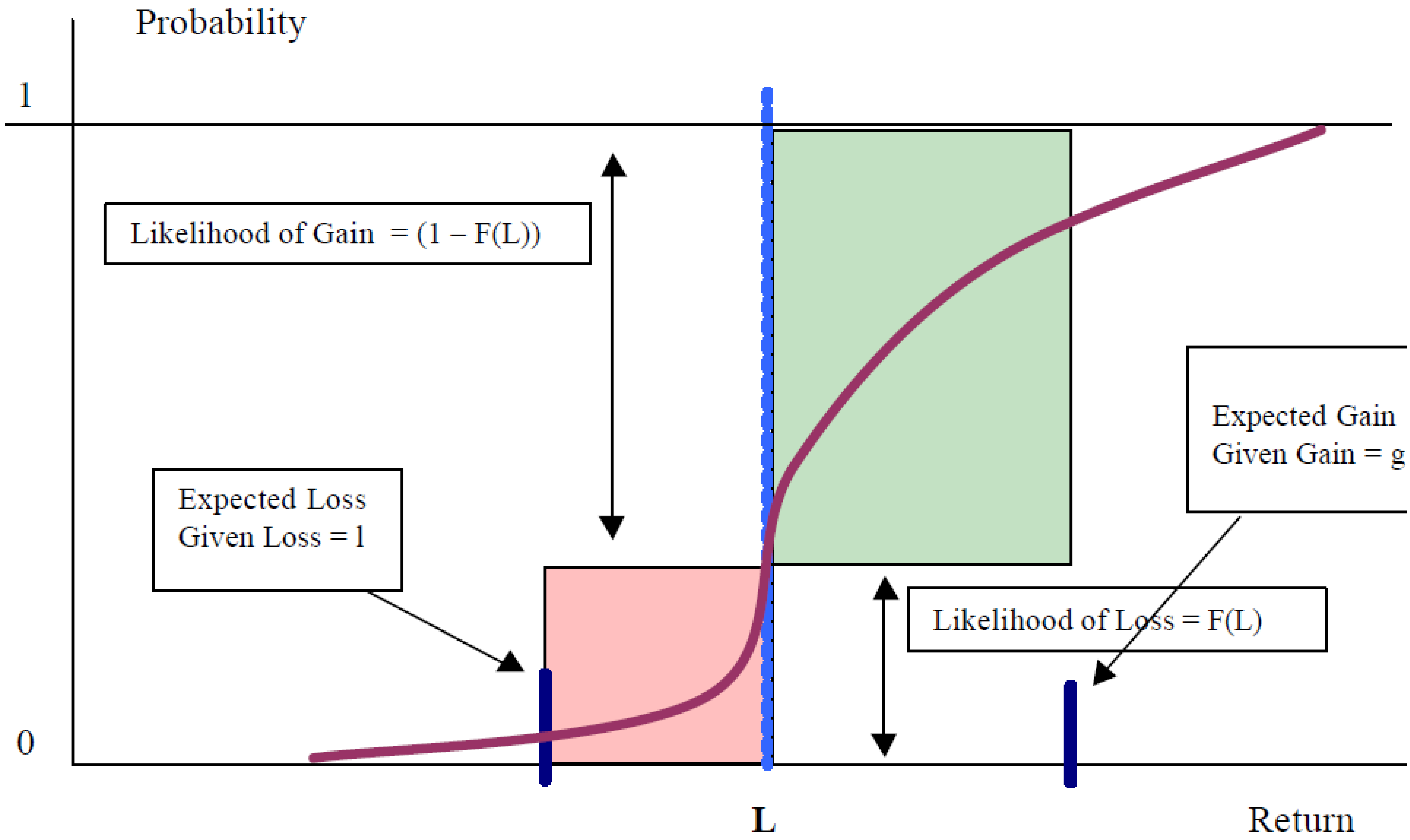
Loss Threshold

- Clearly, the definition, hence likelihoods, of winning and losing depends on how we define loss.
- Suppose L = Loss Threshold,
  - for return < L, we consider it a loss
  - for return > L, we consider it a gain

An Attempt

- To account for
  - the odds of wining and losing
  - the sizes of wining and losing
- We consider
  - $\Omega = \frac{E(r|r>L)\times P(r>L)}{E(r|r\leq L)\times P(r\leq L)}$
- $\Omega = \frac{E(r|r>L)\left(1-F(L)\right)}{E(r|r\leq L)F(L)}$

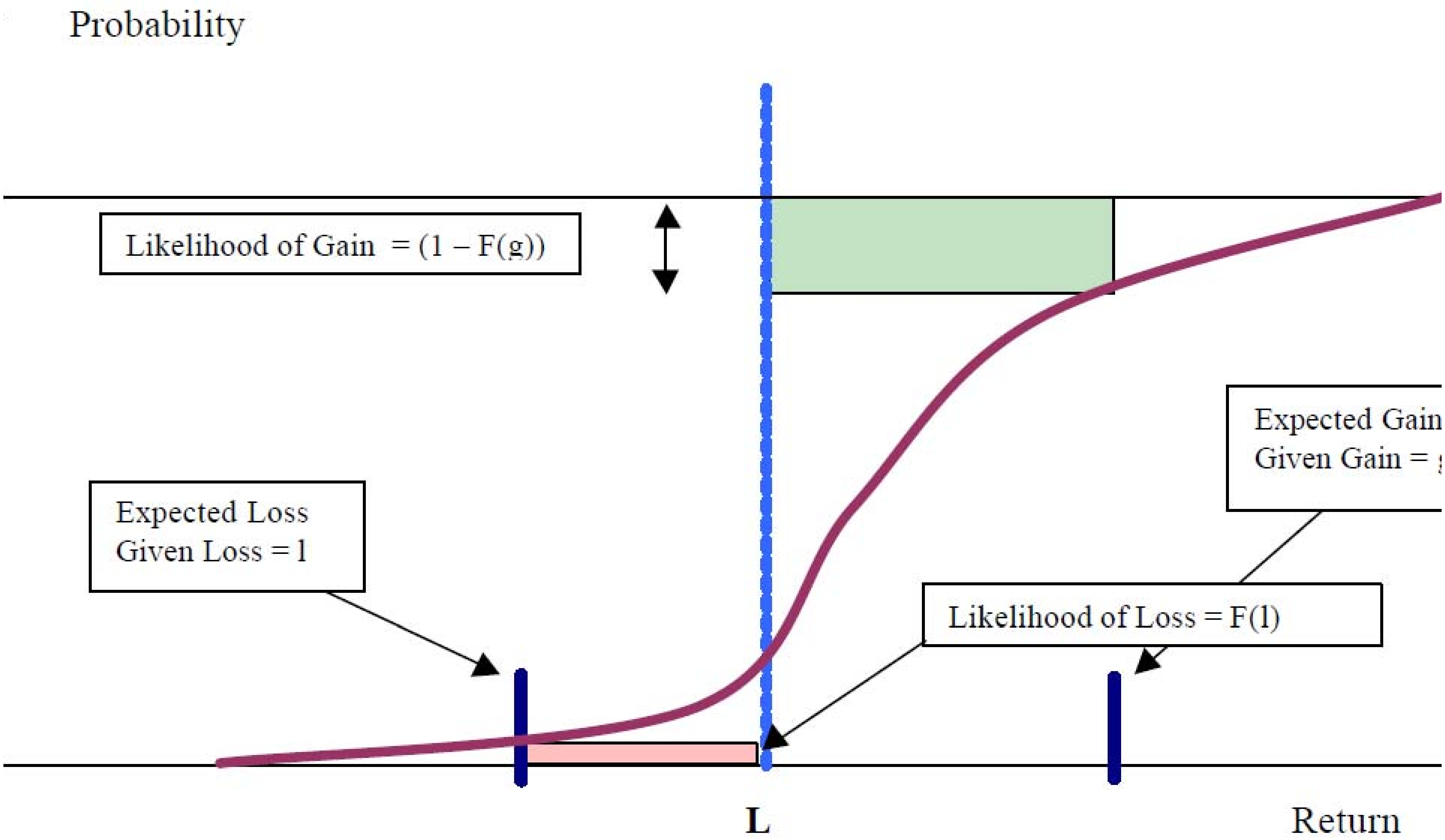
First Attempt



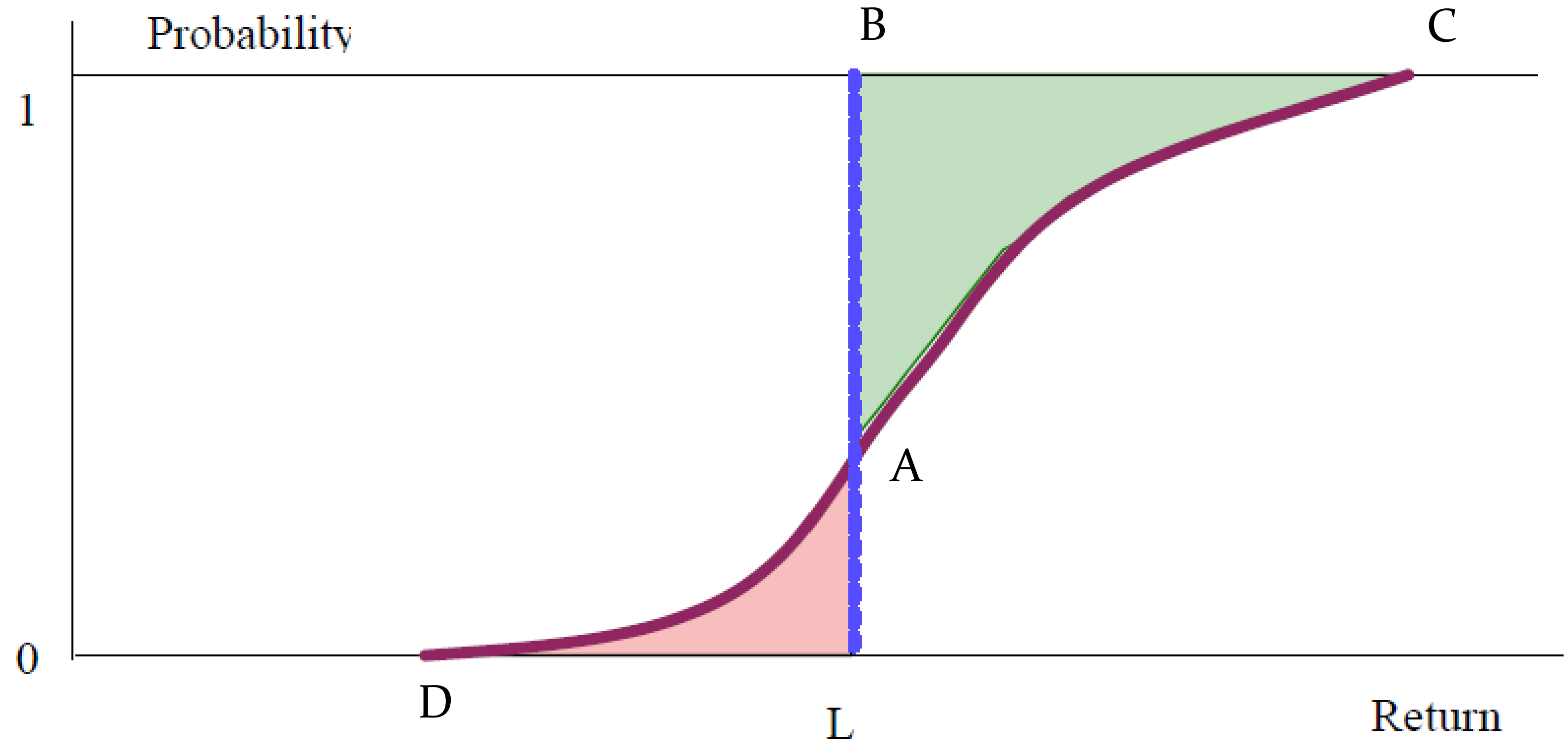
First Attempt Inadequacy

- Why F(L)?
- Not using the information from the entire distribution.
  - hence ignoring higher moments

Another Attempt



Yet Another Attempt



Omega Definition

- Ω takes the concept to the limit.
- Ω uses the whole distribution.
- Ω definition:
  - $\Omega = \frac{ABC}{ALD}$
- $\Omega = \frac{\int_{L}^{b=\max\{r\}}[1-F(r)]dr}{\int_{a=\min\{r\}}^{L} F(r)dr}$

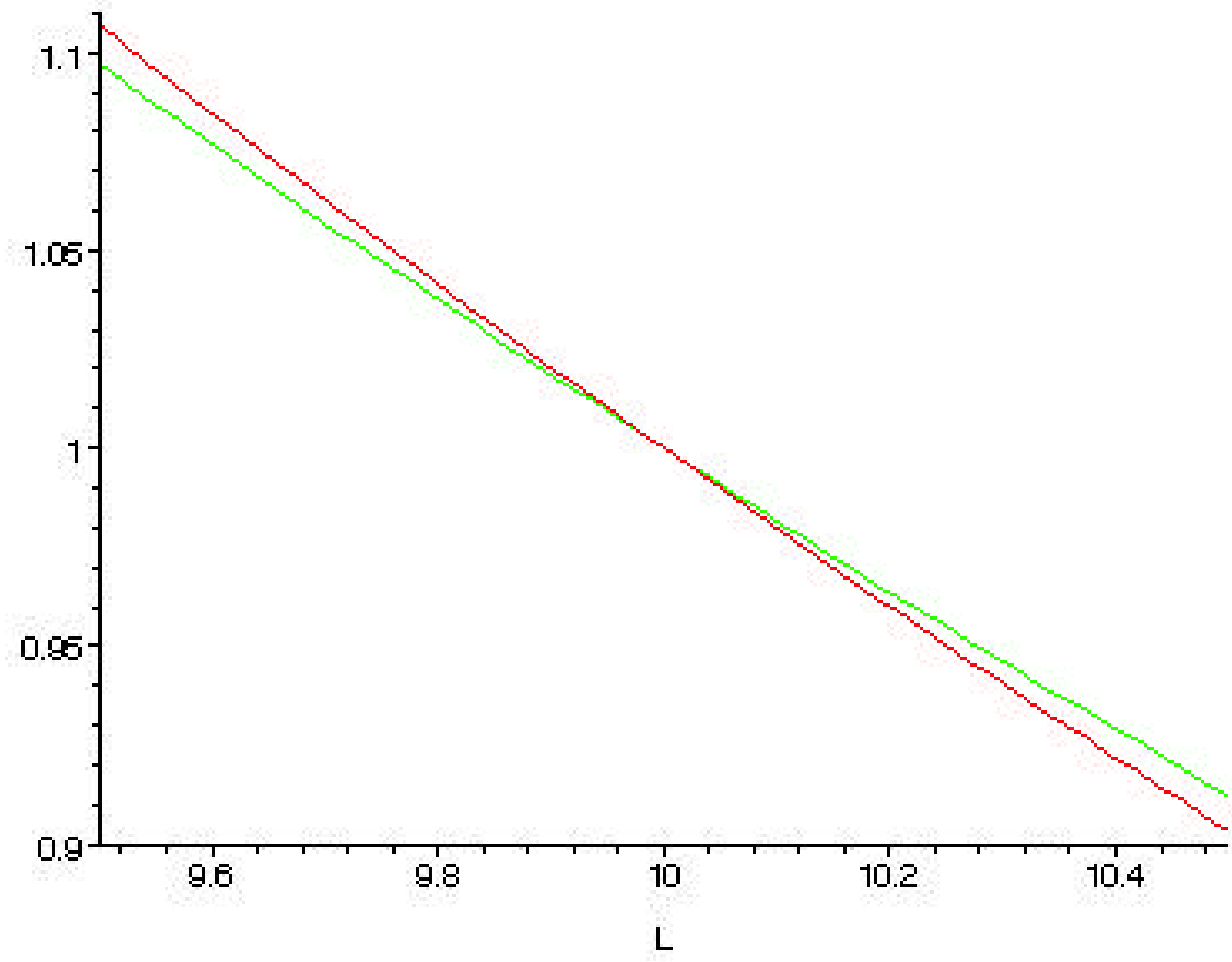
Intuitions

- Omega is a ratio of winning size weighted by probabilities to losing size weighted by probabilities.
- Omega considers size and odds of winning and losing trades.
- Omega considers all moments because the definition incorporates the whole distribution.

Omega Advantages

- There is no parameter (estimation).
- There is no need to estimate (higher) moments.
- Work with all kinds of distributions.
- Use a function (of Loss Threshold) to measure performance rather than a single number (as in Sharpe Ratio).
- It is as smooth as the return distribution.
- It is monotonic decreasing.

Omega Example



Affine Invariant

- $\varphi: r \rightarrow Ar + B$, iff $\widehat{\Omega}\big(\varphi(L)\big) = \Omega(L)$
- $L \rightarrow AL + B$
- We may transform the returns distribution using any invertible transformation before calculating the Gamma measure.
- The transformation can be thought of as some sort of utility function, modifying the mean, variance, higher moments, and the distribution in general.

Numerator Integral (1)

- $\int_L^b d\left[x\left(1 - F(x)\right)\right]$
- $= \left[x\left(1 - F(x)\right)\right]^b{}_L$
- $= b\left(1 - F(b)\right) - L\left(1 - F(L)\right)$
- $= -L\left(1 - F(L)\right)$

Numerator Integral (2)

- $\int_L^b d\left[x\left(1 - F(x)\right)\right]$
- $= \int_L^b \left(1 - F(x)\right)dx + \int_L^b xd\left(1 - F(x)\right)$
- $= \int_L^b \left(1 - F(x)\right)dx - \int_L^b xdF(x)$

Numerator Integral (3)

- $-L(1 - F(L)) = \int_L^b (1 - F(x))dx - \int_L^b xdF(x)$
- $\int_L^b (1 - F(x))dx = -L(1 - F(L)) + \int_L^b xdF(x)$
- $= \int_L^b (x - L)f(x)dx$
- $= \int_a^b \max(x - L, 0)f(x)dx$
- $= E[\max(x - L, 0)]$

undiscounted call option price

Denominator Integral (1)

- $\int_a^L d[xF(x)]$
- $= [xF(x)]^L{}_a$
- $= LF(L) - a\big(F(a)\big)$
- $= LF(L)$

Denominator Integral (2)

- $\int_a^L d[xF(x)]$
- $= \int_a^L F(x)dx + \int_a^L xdF(x)$

Denominator Integral (3)

- $LF(L) = \int_a^L F(x)dx + \int_a^L xdF(x)$
- $\int_a^L F(x)dx = LF(L) - \int_a^L xdF(x)$
- $= \int_a^L (L - x)f(x)dx$
- $= \int_a^b \max(L - x, 0)f(x)dx$
- $= E[\max(L - x, 0)]$

undiscounted put option price

Another Look at Omega

- $\Omega = \frac{\int_{L}^{b=\max\{r\}}[1-F(r)]dr}{\int_{a=\min\{r\}}^{L} F(r)dr}$
- $= \frac{E[\max(x-L,0)]}{E[\max(L-x,0)]}$
- $= \frac{e^{-r_f} E[\max(x-L,0)]}{e^{-r_f} E[\max(L-x,0)]}$
- $= \frac{C(L)}{P(L)}$

Options Intuition

- Numerator: the cost of acquiring the return above L
- Denominator: the cost of protecting the return below L
- Risk measure: the put option price as the cost of protection is a much more general measure than variance

Can We Do Better?

- Excess return in Sharpe Ratio is more intuitive than $C(L)$ in Omega.
- Put options price as a risk measure in Omega is better than variance in Sharpe Ratio.

Sharpe-Omega

- $\Omega_S = \frac{\bar{r} - L}{P(L)}$
- In this definition, we combine the advantages in both Sharpe Ratio and Omega.
  - meaning of excess return is clear
  - risk is bettered measured
- Sharpe-Omega is more intuitive.
- *Ω_S ranks the portfolios in exactly the same way as Ω.*

Sharpe-Omega and Moments

- It is important to note that the numerator relates only to the first moment (the mean) of the returns distribution.
- It is the denominator that take into account the variance and all the higher moments, hence the whole distribution.

Sharpe-Omega and Variance

- Suppose $\bar{r} > L$. $\Omega_S > 0$.
  - The bigger the volatility, the higher the put price, the bigger the risk, the smaller the Ω_S, the less attractive the investment.
  - We want smaller volatility to be more certain about the gains.
- Suppose $\bar{r} < L$. $\Omega_S < 0$.
  - The bigger the volatility, the higher the put price, the bigger the Ω_S, the more attractive the investment.
  - Bigger volatility increases the odd of earning a return above L.

Portfolio Optimization

- In general, a Sharpe optimized portfolio is different from an Omega optimized portfolio.

Optimizing for Omega

- $$\begin{cases} \max\limits_{x} \Omega_S(x) \\ \sum_i^n x_i E(r_i) \geq \rho \\ \sum_i^n x_i = 1 \\ {x_i}^l \leq x_i \leq 1 \end{cases}$$
- Minimum holding: $x^l = \left({x_1}^l, \dots, {x_n}^l\right)'$

Optimization Methods

- Nonlinear Programming
  - Penalty Method
- Global Optimization
  - Tabu search (Glover 2005)
  - Threshold Accepting algorithm (Avouyi-Dovi et al.)
  - MCS algorithm (Huyer and Neumaier 1999)
  - Simulated Annealing
  - Genetic Algorithm
- Integer Programming (Mausser et al.)

3 Assets Example

- $x_1 + x_2 + x_3 = 1$
- $R_i = x_1 r_{1i} + x_2 r_{2i} + x_3 r_{3i}$
- $= x_1 r_{1i} + x_2 r_{2i} + (1 - x_1 - x_2) r_{3i}$

Penalty Method

- $F(x_1, x_2) = -\Omega(R_i) + \rho\{[\min(0, x_1)]^2 + [\min(0, x_2)]^2 + [\min(0,1 - x_1 - x_2)]^2\}$
- Can apply Nelder-Mead, a Simplex algorithm that takes initial guesses.
- F needs not be differentiable.
- Can do random-restart to search for global optimum.

Threshold Accepting Algorithm

- It is a local search algorithm.
  - It explores the potential candidates around the current best solution.
- It “escapes” the local minimum by allowing choosing a lower than current best solution.
  - This is in very sharp contrast to a hilling climbing algorithm.

Objective

- Objective function
  - $h: X \rightarrow R, X \in R^n$
- Optimum
  - $h_{\text{opt}} = \max_{x \in X} h(x)$

Initialization

- Initialize n (number of iterations) and $step$.
- Initialize sequence of thresholds th_k, $k = 1, \dots, step$
- Starting point: $x_0 \in X$

Thresholds

- Simulate a set of portfolios.
- Compute the distances between the portfolios.
- Order the distances from smallest to biggest.
- Choose the first $step$ number of them as thresholds.

Search

- $x_{i+1} \in N_{x_i}$ (neighbour of x_i)
- Threshold: $\Delta h = h(x_{i+1}) - h(x_i)$
- Accepting: If $\Delta h > -th_k$ set $x_{i+1} = x_i$
- Continue until we finish the last (smallest) threshold.
  - $h(x_i) \approx h_{opt}$
- Evaluating h by Monte Carlo simulation.