Assessing the Risk and Return of Financial Trading Systems - a Large Deviation Approach

Nicolas NAVET – INRIA, France
nnavet@loria.fr

René SCHOTT – IECN, France
Schott@loria.fr

CIEF2007 - 07/22/2007

1

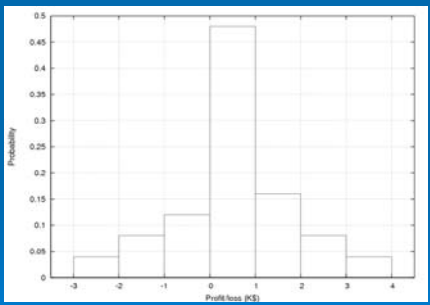
Trading System

- Defined by :
  - Entry condition(s)
  - Exit condition(s)
  - Position sizing
- Implemented in an Automated Trading System (ATS) or executed by a trader

2

Performance of a trading system

- Performance metric : return (P&L), Sharpe ratio, ..
- Reference period – e.g: day, week, …



P&L interval	Probability	k
$(-3,-2]$	$1/25$	-2.5
$(-2,-1]$	$2/25$	-1.5
$(-1,-0]$	$3/25$	-0.5
$(0,1]$	$12/25$	0.5
$(1,2]$	$4/25$	1.5
$(2,3]$	$2/25$	2.5
$(3,4]$	$1/25$	3.5

Distribution of the performance metric

3

Obtaining the distribution of the performance metric

1. Prior use of the TS
2. Back-testing on historical data, but :
   - Does not account for slippage and delays
   - Data-mining bias if a large number of systems are tested
   - ➢ Performance must be adjusted accordingly!

4

Notations and Assumptions

X_i : performance at period *i*

$X_1, X_2, \ldots, X_n$: 1) are mutually independent and identically distributed

2) obey a distribution law that does not change over time

How to assess the risks ?

- We want to estimate $p = P[\sum_i^n X_i < x\$]$

1. Monte-Carlo simulation
2. Analytic approaches:
   1. Markov's, Tchebychev's, Chernoff's upper bounds
   2. Large deviation

Monte-Carlo simulation

- Generate *n* random trading sequences and compute an estimate of the probability
- CLT tells us that the estimate will convergence to *p* but slowly and

$$\text{percentage error} = \frac{\sqrt{p(1-p)}}{\sqrt{np}} \cdot 100$$

- Error bound of 1% with $p=10^{-5}$ requires $n=10^9$
- Problem : random number generators are not perfect ..

7

Analytic approaches

- Weak law of large numbers :

$$\lim_{n\to\infty} P\left(\left|\frac{\sum_i^n X_i}{n} - E[X]\right| < \epsilon\right) \quad \forall\epsilon > 0$$

But the rate of convergence is unknown ..

- Elements of solution:

1. Markov's inequality : $\forall\alpha \quad P\left(X \geq \alpha E[X]\right) \leq \frac{1}{\alpha}$
2. Tchebychev's inequality : $P\left(|X - E[X]| \geq k\sigma[X]\right) \leq \frac{1}{k^2}$

➢ Not tight enough for real-world applications

8

Large deviation: main result

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$: mean performance over *n* periods

- Cramer's theorem : if X_n i.i.d. r.v.

$$P(M_n \in G) \asymp e^{-n \inf_{x \in G} I(x)}$$

e.g. $G = (-\infty, -k\$]$

with *I(x)* the rate function

$$I(x) = \sup_{\tau > 0}[\tau x - \log E(e^{\tau X})] = \sup_{\tau > 0}[\tau x - \log \sum_{k=-\infty}^{+\infty} p_k e^{k\tau}]$$

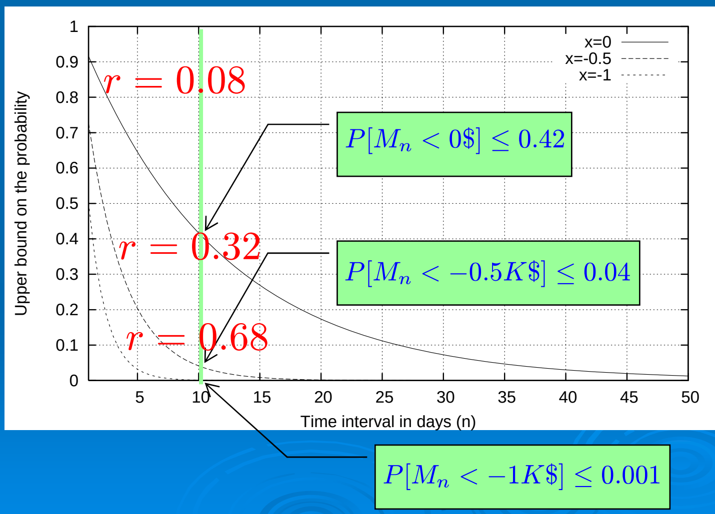
9

Technical contribution

- Can deal with distributions given as frequency histogram (no closed-form)
  - I(x) is the sup. of affine functions and thus convex
  - Computing the point where first derivative equal zero is thus enough
  - Can be done with standard numerical methods

10

Risk over a given time interval



11

Quantifying the uncertainty

- The uncertainty of trading system *Sp* to achieve a performance *x* over *n* time periods is

$$\mathcal{U}(x,n) = P[M_n \leq x/n] \leq e^{-n \inf_{y \leq x/n} I(y)}$$

- *Sp* is with performance objective x over n time periods is less uncertain than *Sp'* with return objective x' over n' time periods if

$$\mathcal{U}(x,n) \leq \mathcal{U}'(x',n')$$

12

Detecting changing market conditions

- Idea: if a TS performs way below what was foreseeable, it suggests that market conditions have changed

E.g., if the current performance level had a probability less than 10^{-6}

13

Portfolio of Trading Systems

- **Assumption:** TS are independent
- Comes to evaluate : $P[\frac{1}{n}\sum_{i=1}^{n}(X_i^1 + X_i^2 + ... + X_i^m) < x\$]$
- Sum of 2 id. r.v. = convolution, computed using Fast Fourier Transform :

$$f \star g = FFT^{-1}(FFT(x) \cdot FFT(y))$$

14

Conclusion

- LD is better suited than simulation for rare events ($<10^{-4}$)
- LD can serve to validate simulation results
- LD helps to detect changing market conditions
- Our approach is practical :
  - No need for closed-form distributions
  - Easily implementable
  - Work for portfolio of TS
  - Can be embedded in a broader analysis

15

Extensions

- There are ways to address the cases:
  - There are serial dependencies in the trade outcomes
  - The market conditions are changing over time ➔ p.d.f. non-homogeneous in time

16