

# Understanding Kelly Criterion.

RK

November 27, 2013

## Abstract

The purpose of this document is to summarize my learnings from reading about Kelly Criterion for money management.

## Contents

<b>1</b>	<b>Background</b>	<b>2</b>
<b>2</b>	<b>Edward Thorpe's paper on Kelly Criterion</b>	<b>2</b>
2.1	Introduction . . . . .	2
2.2	Coin tossing . . . . .	2
2.3	Optimal growth: Kelly criterion formulas for practitioners . . . . .	4
2.4	The probability of reaching a fixed goal on or before $n$ trials . . . . .	4
2.5	The probability of ever being reduced to a fraction $x$ of this initial bankroll . . . . .	5
2.6	The probability of being at or above a specified value at the end of a specified number of trials . . . . .	6
2.7	Continuous approximation of expected time to reach a goal . . . . .	7
2.8	Comparing fixed fraction strategies: the probability that one strategy leads another after $n$ trials . . . . .	7
2.9	Blackjack . . . . .	7
2.10	Sports betting . . . . .	8
2.11	Wall street: the biggest game . . . . .	8
<b>3</b>	<b>Size Matters</b>	<b>8</b>
<b>4</b>	<b>Ernie Chan's blog post</b>	<b>11</b>
<b>5</b>	<b>Betting with the Kelly Criterion</b>	<b>11</b>
<b>6</b>	<b>Money and Risk Management</b>	<b>13</b>
<b>7</b>	<b>Risk Management</b>	<b>15</b>
7.1	Leverage . . . . .	15
7.2	Optimization of Expected Growth Rate using Simulated Returns . . . . .	15
7.3	Optimization of Historical growth rate . . . . .	17
7.4	Maximum Drawdown . . . . .	17
7.5	Constant Proportion Portfolio insurance . . . . .	17
7.6	Stop Loss . . . . .	17

# 1 Background

I remember having read “Fortune’s formula” 5 years back and I don’t remember a thing why Kelly’s criterion is used. I have spent the last two days going over Kelly Criterion. My immersion in math has given me skillset to understand Thorpe paper. I spent 6 to 8 hrs on Thorpe’s paper after which reading non-math content became easy and pleasant. This document will summarize my learnings from Thorpe’s paper, some general notes on the web about Kelly criterion and chapters from Ernie chan’s book.

## 2 Edward Thorpe’s paper on Kelly Criterion

Edward Thorpe, *The Kelly Criterion in Blackjack, Sportbetting and Stock market*, 1997.

### 2.1 Introduction

How much of capital must be invested on various profitable opportunities ? There were several approaches considered in the past. One approach is to choose a goal, such as to minimize the probability of total loss within a specified number of trials,  $N$ . Another example would be to maximize the probability of reaching a fixed goal on or before  $N$  trials. A different approach, much studied by economists and others, is to value money using a utility function. Thorpe was introduced to Kelly criterion by Shannon in 1960. Kelly’s criterion was a bet on each trial so as to maximize  $E(\log X)$  where  $X$  is the wealth of the investor.

### 2.2 Coin tossing

Imagine odds are in our favor in a coin tossing contest  $p > q$  and we come across an infinitely wealthy opponent. We want to decide how much we must bet each game ? If our initial capital is  $X_0$

$$X_k = X_{k-1} + T_k B_k$$

where  $T_k = \pm 1$ , depending on the outcome and  $B_k$  is the amount bet.

The expected value of the wealth after  $n$  steps is

$$E(X_n) = X_0 + \sum_{k=1}^n E(B_k T_k) = X_0 + (p - q) \sum_{k=1}^n E(B_k)$$

Hence the above strategy suggests *bet everything* but the probability of eventual ruin is 1. Hence this *bold strategy* is often undesirable. Using a timid strategy by taking minimum bet on each trial, also minimizes expected gain. Hence *timid strategy* is also no good. This suggests an intermediate strategy, somewhere between maximizing  $E(X_n)$  and minimizing  $E(X_n)$ . A midway approach is the Kelly’s Criterion based on logarithmic utility. One strategy is to bet a fraction  $f$  of your bankroll.

$$B_i = f X_i$$

Hence “ruin” in the technical sense cannot happen. Let  $G_n(f)$  represent the exponential rate of increase of capital per trial and  $g(f)$  represent the expected value of the same.

$$G_n(f) = \log \left[ \frac{X_n}{X_0} \right]^{1/n}$$

and

$$g(f) = E \left\{ \frac{S}{n} \log(1 + f) + \frac{F}{n} \log(1 - f) \right\}$$

Differentiating the above one gets the optimal fraction for bettings as  $p - q$ . For the graph we have  $g(0) = 1$ ,  $g(1^-) = -\infty$ ,  $g(f_c) = 0$ , for some  $f_c$  and  $g(f)$  takes the max value, for some  $f^*$ ,

- $0 < f < f^*$  guarantees a no ruin game
- $f = f^*$  gives the maximum growth rate
- $0 < f < f^*$  or  $f^* < f < f_c$  guarantees win but with slower rate of growth
- $f > f_c$  means a ruin

**Theorem 2.1** 1. If  $g(f) > 0$ , then  $\lim_{n \rightarrow \infty} X_n = \infty, a.s$

2. If  $g(f) < 0$ , then  $\lim_{n \rightarrow \infty} X_n = 0, a.s$

3. If  $g(f) = 0$ , then  $\limsup_{n \rightarrow \infty} X_n = \infty, a.s$ ,  $\liminf_{n \rightarrow \infty} X_n = 0, a.s$

4. Given a strategy  $\phi^*$ , which maximizes  $E \log X_n$  and any other strategy  $\phi$ , then

$$\lim_{n \rightarrow \infty} \frac{X_n(\phi^*)}{X_n(\phi)} = \infty, a.s$$

5. The expected time for the current capital  $X_n$  to reach any preassigned goal  $C$ , is asymptotically least with a strategy that maximized  $E(X_n)$ .

6. Suppose the return on bet  $U_i$  is a binomial random variable with probability of success  $p_i$ ,  $1/2 < p_i < 1$ , then  $E(X_n)$  is maximized by choosing on each trial, the fraction  $p_i - q_i$  which maximized  $E(\log 1 + f_i U_i)$

**Example 1** Player A plays against an infinitely wealthy adversary. Player A wins even money on successive independent flips of a biased coin with a win probability of  $p = .53$  (no ties). Player A has an initial capital of  $X_0$  and capital is infinitely divisible

```
p <- 0.53
q <- 1 - p
(f.star <- p - q)

## [1] 0.06

fun <- function(fc) {
  return(p * log(1 + fc) + q * log(1 - fc))
}
(f.c <- uniroot(fun, interval = c(0.01, 0.99))$root)

## [1] 0.1197
```

- If the player bets a fraction larger than 6% up to  $f_c$ ,  $X_n$  will slowly edge towards  $\infty$ .
- If the player bets a fraction smaller than 6%,  $X_n$  will slowly edge towards  $\infty$ .
- If the player bets a fraction larger than 12% ,  $X_n$  will eventually hit 0
- After  $n$  successive bets, the player's bank roll reaches  $0.0018n$ , then the time to double the money is

```
(round(log(2)/fun(f.star)))
## [1] 385
```

The Kelly criterion can easily be extended to uneven payoff games. Suppose Player A wins  $b$  units for every unit wager. Further, suppose that on each trial the win probability is  $p > 0$  and  $pb - q > 0$  so the game is advantageous to Player A, the optimal fraction of capital is

$$f^* = \frac{bp - q}{b}$$

## 2.3 Optimal growth: Kelly criterion formulas for practitioners

This section compares Kelly criterion with other fixed fraction strategies

## 2.4 The probability of reaching a fixed goal on or before $n$ trials

The paper talks about the most useful single fact for dealing with diverse problems in gambling and in the theory of financial derivatives.

$$P(\sup[X(t) - (at + b)] \geq 0, 0 \leq t \leq T) = \mathcal{N}(-\alpha - \beta) + e^{-2ab}\mathcal{N}(\alpha - \beta)$$

where  $\alpha = a\sqrt{T}$  and  $\beta = b/\sqrt{T}$

Let  $f$  be the fraction bet. Assume iid trials  $Y_i, i = 1, 2, \dots, n$ , with  $P(Y_i = 1) = p > 1/2, P(Y_i = -1) = q < 1/2$ . Let  $V_k$  be the value of the gambler after  $k$  trials.

$$\begin{aligned} P(V_k \geq C, 1 \leq k \leq n) &= P(X_t \geq \log C - mt/s^2, 1 \leq t \leq s^2n) \\ &= \mathcal{N}(-\alpha - \beta) + e^{-2ab}\mathcal{N}(\alpha - \beta) \end{aligned}$$

with

$$\begin{aligned} T &= s^2n \\ b &= \log C \\ a &= -m/s^2 \\ \alpha &= -m\sqrt{n}/s \\ \beta &= \log C/(s\sqrt{n}) \end{aligned}$$

```
frac.strategy.1 <- function(f){
  C      <- 2
  n      <- 10^4
  p      <- 0.51
  q      <- 1-p
  mu     <- function(x){ p*log(1+x) + q*log(1-x) }
  sd2    <- function(x){ p*q*log( (1+x)/(1-x) )^2 }
  m      <- mu(f)
  s2     <- sd2(f)
  T      <- s2*n
  b      <- log(C)
  a      <- -m/s2
```

```
alpha <- a*sqrt(T)
beta  <- b/sqrt(T)
pnorm(-alpha-beta) + exp(-2*a*b)*pnorm(alpha-beta)
}
f      <- seq(0.001,0.99,length.out = 100)
probs <- sapply(f,function(z)frac.strategy.1(z))
(f[which.max(probs)])

## [1] 0.02098
```

```
plot(f, probs, xlab = "fraction",ylab = "probability",pch = 19, cex=0.4,col = "blue")
```

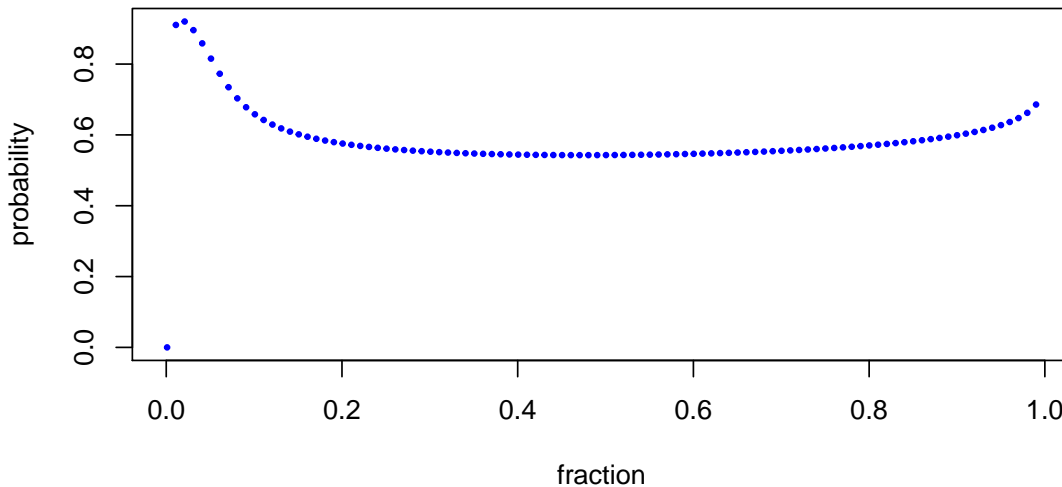


Figure 2.1: Strategy - 1

## 2.5 The probability of ever being reduced to a fraction $x$ of this initial bankroll

To address this problem , the probability turns out to

$$Prob(\cdot) = e^{2ab}, \quad a = -m/s^2, b = -\log x$$

```
frac.strategy.2 <- function(f,x){
  p      <- 0.51
  q      <- 1-p
  mu     <- function(x){ p*log(1+x) + q*log(1-x) }
  sd2    <- function(x){ p*q*log( (1+x)/(1-x) )^2 }
  m      <- mu(f)
```

```
s2      <- sd2(f)
b       <- -log(x)
a       <- -m/s2
exp(2*a*b)
}
frac.strategy.2(0.02,0.3)

## [1] 0.2999
```

For the limiting continuous approximation and the Kelly optimal fraction

$$P(V_k(f^*) \leq x, k \geq 1) = x$$

Most cautious gamblers or investors who use Kelly find the frequency of substantial bankroll reduction to be uncomfortably large. We can see why now. To reduce this, they tend to prefer somewhat less than the full betting fraction

## 2.6 The probability of being at or above a specified value at the end of a specified number of trials

$$P(X(T) \geq aT + b) = \mathcal{N}(-\alpha - \beta)$$

```
frac.strategy.3 <- function(f){
  C      <- 2
  n      <- 10^4
  p      <- 0.51
  q      <- 1-p
  mu     <- function(x){ p*log(1+x) + q*log(1-x) }
  sd2    <- function(x){ p*q*log( (1+x)/(1-x) )^2 }
  m      <- mu(f)
  s2     <- sd2(f)
  T      <- s2*n
  b      <- log(C)
  a      <- -m/s2
  alpha  <- a*sqrt(T)
  beta   <- b/sqrt(T)
  pnorm(-alpha-beta)
}
(frac.strategy.3(.0117))

## [1] 0.7947

(frac.strategy.3(.02))

## [1] 0.7433
```

## 2.7 Continuous approximation of expected time to reach a goal

$$n(C, f) = \frac{\log C}{g(f)}$$

## 2.8 Comparing fixed fraction strategies: the probability that one strategy leads another after $n$ trials

One can compute the expected value and the variance for rate of increase of capital per trial

$$\begin{aligned} G(f) &= (W_n/n) \log(1+f) + (1-W_n)/n \log(1-f) \\ E(G(f)) &= p \log(1+f) + q \log(1-f) \\ \text{Var}(G(f)) &= \frac{pq}{n} \left\{ \log \left( \frac{1+f}{1-f} \right)^2 \right\} \end{aligned}$$

Hence using the above moments, one can compare the strategies in two cases, one where the players are using two different fractions on the same trial, and second case is where the betting is on iid trials.

One can compare the Kelly strategy with other fixed fractions to determine the probability that Kelly leads after  $n$  trials. Note that this probability is always greater than 0.5 because  $g(f^*) - g(f) > 0$  where  $f^* = p - q, f \neq f^*$ . But some small values of  $n$  this might not be the case. So, when does Kelly start to dominate? It is important to understand that “the long run”, the time it takes for  $f^*$  to dominate a specified neighbor by a specified probability, can vary without limit. Each application requires a separate analysis. The paper gives an example where dominance is slow.

## 2.9 Blackjack

The analysis is more complicated than that of coin tossing because the payoffs are not simply one to one. In particular the variance is generally more than 1 and the Kelly fraction tends to be less than for coin tossing with the same expectation. Moreover, the distribution of various payoffs depends on the player advantage.

The betting in blackjack has a lot of nuances. The author takes a simple example where the player must make small “waiting” bets on the unfavorable situations in order to be able to exploit the favorable situations. On these he will place “large” bets. There are two cases considered.

CASE 1: Bet  $f_0$  on unfavorable situations and find the optimal  $f^*$  for favorable situations. In this case the solution is the old Kelly criterion  $p - q$ . Since the side bets are fixed quantity, there is no influence of it on the optimal bet fraction via Kelly.

CASE 2: Bet  $f$  in favorable situations and  $af, 0 < a < 1$  in unfavorable situations. As  $a$  increases from 0 to 1, the Kelly fraction decreases from  $p - q$  to 0. Thus the Kelly fraction for favorable situations is less in this case when bets on unfavorable situations reduce the overall advantage of the game.

## 2.10 Sports betting

Thorp introduces multiple betting situation via sports betting examples. The section starts off with an example of betting on two independent favorable coins with betting fractions  $f_1$  and  $f_2$ .

$$g(f_1, f_2) = p_1 p_2 \log(1 + f_1 + f_2) + p_1 q_2 \log(1 + f_1 - f_2) + p_2 q_1 \log(1 - f_1 + f_2) + q_1 q_2 \log(1 - f_1 - f_2)$$

Optimizing the above expectation, one gets

$$f_1^* = \frac{m_1(1 - m_2^2)}{1 - m_1^2 m_2^2}$$

$$f_2^* = \frac{m_2(1 - m_1^2)}{1 - m_1^2 m_2^2}$$

where  $m_1$  and  $m_2$  ideal kelly fractions, i.e.  $m_i = p_i - q_i$ . This shows that the fractions in this case are reduced in the case iid simulatenous bettings.

The first glimpse of using kelly comes in the form of a question,

Our simultaneous sports bets were generally on different games and typically not numerous so they were approximately independent and the appropriate fractions were only moderately less than the corresponding single bet fractions. Question: Is this always true for independent simultaneous bets? Simultaneous bets on blackjack hands at different tables are independent but at the same table they have a pairwise correlation that has been estimated at 0.5. This should substantially reduce the Kelly fraction per hand. The blackjack literature discusses approximations to these problems. On the other hand, correlations between the returns on securities can range from nearly -1 to nearly 1. An extreme correlation often can be exploited to great advantage through the techniques of “hedging”. The risk averse investor may be able to acquire combinations of securities where the expectations add and the risks tend to cancel. The optimal betting fraction may be very large.

This means one might have to take covariance structure between trials while betting. Using a simple coin toss example, Thorp shows that Kelly criterion going increases as the correlation of outcomes in two trials become negative.

## 2.11 Wall street: the biggest game

I had to plod through the first 20 pages of the paper, understand the author’s arguments to give me enough preparation to go over this section. This gives the Kelly criterion applied to a security of stocks. It connects the covariance of stocks, the excess mean returns of the stocks and come up with a formula for kelly fraction for a set of stocks.

This paper is very well organized with adequate examples, solid math , intuition, practical results. Such papers are rare I guess and hence it might be worth the time and effort that goes on in reading and thinking about various aspects of the paper.

## 3 Size Matters

I stumbled on to an article by titled [Size Matters](#) and found that the paper hits on all a variety of important points relating to using this criterion. The paper has the following ideas



- Not many portfolio managers have position sizing as their forte
- Based on information theory, the Kelly Criterion says an investor should choose the investment(s) with the **highest geometric mean return**. This strategy is distinct from those based on mean/variance efficiency.
- Kelly's fraction for discrete world is

$$f = \frac{\text{Edge}}{\text{Odds}}$$

A simple example illustrates the point. Assume you can participate in a coin toss game where heads pays \$2 and tails costs \$1. You start with a \$100 bankroll and can play for 40 rounds. What betting strategy will allow you to achieve the greatest probability of the most money at the end of the fortieth round?

```
b      <- 2
p      <- 1/2
q      <- 1-p
(f.star <- (p*b-q)/b)
## [1] 0.25
```

```
set.seed(1234)
ftest  <- seq(0.01,0.99,length.out=100)
wins   <- 20
lost    <- 40 - wins
result  <- ((1+ftest*2)^20)*((1-ftest)^20)
ftest[which.max(result)]
## [1] 0.2476
```

```
plot(ftest,result, type="l", col="blue", ylab="Multiple of Original Bankroll",xlab="f")
```

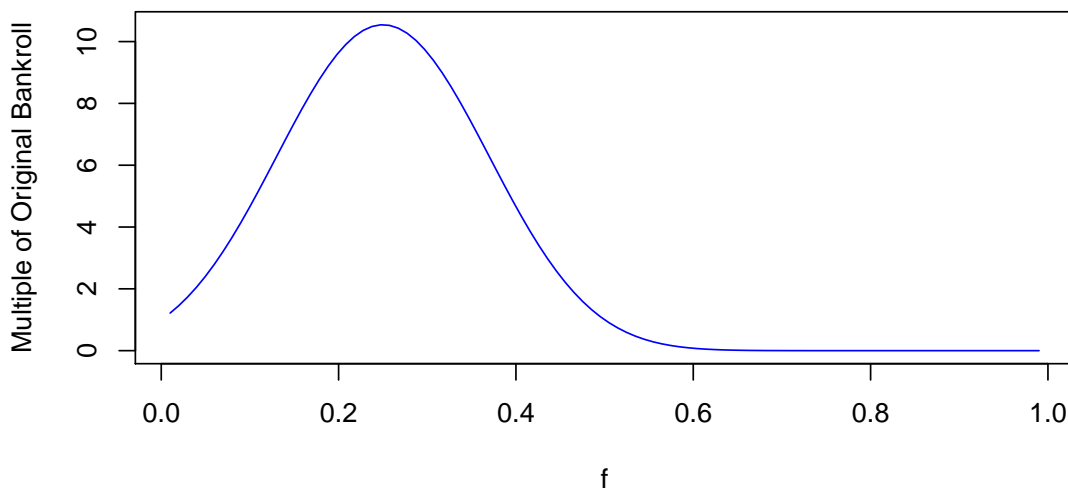


Figure 3.1: The Kelly Formula Solves for the Optimal Betting Strategy

- Connection with Shannon's work - For Shannon, the incompressible part of a message relates to its unpredictability. The less probable a message, the more bandwidth it requires. Shannon's theory also considers equivocation - the chance the message is wrong and shows - you must subtract equivocation from the channel capacity to determine the information rate. John Kelly, recognized another application for information theory's ideas: gambling. Information in a betting setting is something the market does not already know. Consistent with the idea of equivocation, true information is also probabilistic. Kelly imagined a system where you have an edge; a set of expectations that differs from those of the market. He then developed a formula, based on Shannon's work, showing the exact amount of your bankroll you should bet in order to maximize your capital over the long term. Consistent with the theory, the maximum rate of return comes when you know something the market doesn't.
- via "The Mathematics of Gambling" : *The chance of ruin is small* - Because the Kelly system is based on proportional bets, losing all of your capital is theoretically impossible (assuming money is infinitely divisible). Even so, a small chance of a significant drawdown remains. *The Kelly system is highly likely to grow a bankroll faster than other systems*. Provided comparably attractive opportunities continue to appear, there is a high probability the system will generate a bankroll that exceeds other systems by a determinable multiple. *You tend to reach a specified level of winnings in the least average time*. If you have a financial end goal in mind and continuous opportunities, the Kelly system will likely allow you to achieve the objective in a shorter time than other systems.
- the Kelly Criterion works well when you parlay your bets, face repeated opportunities, and know what the underlying distribution looks like.
- the central message for investors is that standard mean/variance analysis does not deal with the compounding of investments. If you seek to compound your wealth, then maximizing geometric returns should be front and center in your thinking.
- Problems with Practical application of Kelly
  - the geometric mean maximization strategy does not assure that you will end up with more wealth than other strategies. Since the approach is based on probability, there remains a very small chance an investor will do poorly
  - success of geometric mean maximization depends on investors staying in the market for the long run. If an investor needs access to the funds in the near term, the benefits of compounding do not apply.
  - the system assumes the investment payoffs remain steady and the investment opportunities set is large enough to accommodate a rising asset base. Shifting investment payoffs undermine the system.
- Geometric mean maximization simply does not make sense for a portfolio manager in this short term mindset.
- Good for closed ended funds but firms are increasingly offering open ended funds
- Loss aversion can lead investors to suboptimal decisions and sticking with Kelly is difficult for psychological reasons.
- Mean variance is not the best way to think about maximizing long-term wealth if you are reinvesting your investment proceeds

## 4 Ernie Chan's blog post

Ernie Chan has a dedicated chapter on Kelly criterion in one of his books on quant trading. On his blog he has summarized his view about it via the blog post, [How much leverage should you use? Maximizing growth without risking bankruptcy](#). The main points from his posts and the list of comments for his post are :

- Continuous version of Kelly fraction

$$f = \frac{m - r}{s^2}$$

- . This quantity looks like the familiar Sharpe ratio, but it is not.
- The common recommendation is that you should halve your expected returns estimated from backtests when calculating  $f$ . This is often called the half-Kelly criterion.
- Kelly criterion requires you to sell into a loss (assuming you have a long-only portfolio here), and buys into a profit — something that requires steely discipline to achieve. It also runs counter to the usual mean-reversion expectation.
- if your goal is to maximize your wealth (which equals your initial equity times the maximum growth rate possible using your strategy), Kelly criterion is the way.
- If you adopt the Kelly criterion, there may be long periods of drawdown, highly volatile returns, low Sharpe ratio, and so forth. The only thing that Kelly guarantees (to an exponentially high degree of certainty), is that you will maximize the growth potential of your strategy in the long run, and you will not be bankrupt in the interim because of the inevitable short-term market fluctuations.
- comments to the blog post
  - The basic idea of the Kelly criterion, optimize geometric return of your portfolio, makes a lot of sense, only the optimal leverage will be substantially lower than a naive application of the Kelly criterion suggests. You can convince yourself of this if you do a Monte-Carlo simulation of differently leveraged portfolios with actual return distribution instead of Gaussian returns, periodic instead of continuous rehedgeing, and trading costs
  - The half factor in fractional kelly comes from collective gut feeling of quant traders
  - The reason full kelly betting isn't used is not because it's "wrong" (i.e. lognormal vs. fat tailed, though that causes problems), but because we usually don't really know the expected return nor variance. We might be able to make good guesses, but they are still guesses. Kelly shows that overbetting leads to ruin, whereas underbetting is "merely" suboptimal.

## 5 Betting with the Kelly Criterion

This is a brief note with a simulation exercise for using Kelly criterion. First thing that I realized from reading this write up is that, Kelly took the base 2 logarithm of capital as his utility function in his paper. Somehow my mind tricked me in assuming natural logarithm. I just thought a simulation similar to the one done in the note.

```
set.seed(1234)
p      <- 0.55
q      <- 1-p
b      <- 10/11
f.star <- (b*p - q)/b
n      <- 250
```

```
X.0      <- 100
f.const  <- 0.1
outcome  <- sample(c(0,1),n,prob=c(q,p), replace=T)
growth11 <- as.ts(X.0*cumprod(ifelse(outcome==1,1+b*f.star,1-f.star)))
growth21 <- as.ts(X.0*cumprod(ifelse(outcome==1,1+b*f.const,1-f.const)))
```

```
set.seed(1234)
p      <- 0.55
q      <- 1-p
b      <- 10/11
f.star <- (b*p - q)/b
n      <- 500
X.0    <- 100
f.const <- 0.1
outcome <- sample(c(0,1),n,prob=c(q,p), replace=T)
growth12 <- as.ts(X.0*cumprod(ifelse(outcome==1,1+b*f.star,1-f.star)))
growth22 <- as.ts(X.0*cumprod(ifelse(outcome==1,1+b*f.const,1-f.const)))
```

```
plot(as.zoo(cbind(growth11,growth21)/X.0),plot.type="s",col=c("red","black"),xlab="",
      ylab="multiple of initial capital",cex.lab = 0.7,cex.axis=0.7)

plot(as.zoo(cbind(growth12,growth22)/X.0),plot.type="s",col=c("red","black"),xlab="",
      ylab="multiple of initial capital",cex.lab = 0.7,cex.axis=0.7)
```

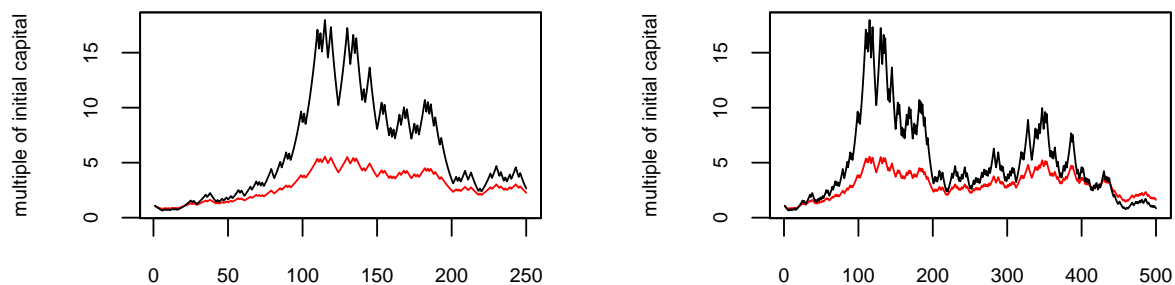


Figure 5.1: 250 vs. 500 bets

For the random realization, the kelly criterion and an arbitrary fraction does not make any difference. In fact following Kelly criterion, the portfolio becomes quite volatile.

## 6 Money and Risk Management

Chapter 6 of the book, "Quantitative Trading, How to build your own algorithmic trading business?", deals with Kelly criterion and some backtesting results on SPY. Here are some points from the chapter

The question that this chapter answers is, "how should capital be allocated amongst a set of portfolio of strategies?". Denoting the optimal fractions to be allocated as  $F^* = (f_1^*, f_2^*, \dots, f_n^*)^T$ , Thorpe shows that the optimal allocation is

$$F^* = C^{-1}M$$

Here,  $C$  is the covariance matrix such that matrix element  $C_{ij}$  is the covariance of the returns of the  $i$ th and  $j$ th strategies, and  $M = (m_1, m_2, \dots, m_n)^T$  is the column vector of mean returns of the strategies.

If we assume that the strategies are all statistically independent, the covariance matrix becomes the diagonal matrix, with the diagonal elements equal to the variance of the individual strategies. This leads to

$$f_i = \frac{m_i}{s_i^2}$$

If there is a restriction on your maximum overall leverage  $l$ , then you might have to reduce each  $f_i$  by the same factor

$$\frac{l}{|f_1| + |f_2| + \dots + |f_n|}$$

By adopting the above fraction for each strategy, the long term compounded growth rate of strategy becomes

$$g = r + S^2/2, \quad \text{where } S \text{ is the Sharpe ratio of the strategy}$$

The takeaway is that the higher the Sharpe ratio of your portfolio (or strategy), the higher the maximum growth rate of your equity (or wealth), provided you use the optimal leverage recommended by the Kelly formula. An example from the book that involves optimal allocation using Kelly formula

*Pick three sector-specific ETFs and see how we should allocate capital among them to achieve the maximum growth rate for the portfolio. The three ETFs are: OIH (oil service), RKH (regional bank), and RTH (retail). The daily prices are downloaded from Yahoo! Finance.*

```
OIH <- read.csv("./data/OIH.csv", stringsAsFactors= F)
OIH$Date <- as.Date(OIH$Date, format = "%m/%d/%Y")
OIH <- OIH[,c(1,7)]
OIH.xts <- xts(OIH[,2], OIH[,1])
colnames(OIH.xts) <- c("OIH")

RTH <- read.csv("./data/RTH.csv", stringsAsFactors= F)
RTH$Date <- as.Date(RTH$Date, format = "%m/%d/%Y")
RTH <- RTH[,c(1,7)]
RTH.xts <- xts(RTH[,2], RTH[,1])
colnames(RTH.xts) <- c("RTH")

RKH <- read.csv("./data/RKH.csv", stringsAsFactors= F)
RKH$Date <- as.Date(RKH$Date, format = "%m/%d/%Y")
```

```

RKH          <- RKH[,c(1,7)]
RKH.xts      <- xts(RKH[,2],RKH[,1])
colnames(RKH.xts) <- c("RKH")

data         <- merge(OIH.xts,RTH.xts,RKH.xts)
data         <- data[complete.cases(data),]
head(data)

##           OIH    RTH    RKH
## 2001-05-17 84.60 87.09 93.63
## 2001-05-18 85.30 86.87 93.31
## 2001-05-21 86.06 88.83 93.43
## 2001-05-22 84.76 88.96 95.15
## 2001-05-23 81.58 87.27 95.04
## 2001-05-24 79.84 88.13 95.43

returns      <- (data -lag(data,1))/lag(data,1)
returns      <- returns[complete.cases(returns),]
excess.ret   <- returns - repmat(0.04/252,dim(returns)[1],dim(returns)[2])
(M           <- 252* apply(excess.ret,2,mean))

##           OIH          RTH          RKH
## 0.139568 -0.007346 0.029400

(C           <- 252*cov(excess.ret) )

##           OIH          RTH          RKH
## OIH 0.11090 0.01825 0.02001
## RTH 0.01825 0.04197 0.02689
## RKH 0.02001 0.02689 0.03716

(F           <- solve(C)%*%M )

##           [,1]
## OIH 1.292
## RTH -1.488
## RKH 1.172

(g           <- -0.04 + t(F)%*%C%*%F/2 )

##           [,1]
## [1,] 0.1529

(S           <- sqrt(t(F)%*%C%*%F) )

##           [,1]
## [1,] 0.4751

(oih.g <- (sqrt(252)*mean(excess.ret[,1])/sd(excess.ret[,1]))^2/2 + 0.04)

## [1] 0.1278

```

The compounded growth for the three stocks is 15.29 %. that is greater than the best individual growth rate (OIH) of 12.78 %.

The chapter talks about a limitation of using half-kelly :Sometimes, even taking the conservative half-Kelly formula may be too aggressive, and traders may want to limit their portfolio size further by additional constraints. This is because the application of the Kelly formula to continuous finance is premised on the assumption that return distribution is Gaussian.

The chapter is certainly gungho about Kelly,as can be inferred from this quote,

My newfound discipline and faith in the Kelly formula has so far prevented similar disasters from happening again.

At the beginning of the chapter, the author borrows something from Behavioral finance that I found very interesting

research has suggested that the average human being needs to have the potential for making \$2 to compensate for the risk of losing \$1, which may explain why a Sharpe ratio of 2 is so emotionally appealing

## 7 Risk Management

Chapter 8 of the book,"Algorithmic Trading Winning Strategies and Their Rationale " , deals with risk management and the author touches upon Kelly criterion.

The chapter discusses the key concept in risk management, "prudent use of leverage". However there are times when reality forces to limit the max drawdown of an account. In such cases, stoploss or portfolio insurance can be followed.

### 7.1 Leverage

How does one go about analyzing the leverage to be applied to a set of stocks or set of strategies ? The answer is Kellys criterion. In the author's previous book on algorithmic trading, there is a full chapter dedicated to Kelly criterion. Though it is not mathematically rigorous, the author touches upon all the important points that one need to understand about Kelly fraction. The key idea is the idea of geometrical returns and the optimal leverage to maximixe the networth is given by kelly fraction. In practice, traders follow half-Kelly to incorporate the estimation error of mean and covariances of stocks/strategies that go in to the computation of kelly fraction. Obviously there is also an error in assuming gaussian returns but for that original kelly formula by Thorpe does not given any guidance. In particular if there is a limit on leverage, which is usually the case, the author makes an important point

When we have two or more strategies with very different independent growth rates, and when we have a maximum leverage constraint that is much lower than the Kelly leverage, it is often optimal to just apply all of our buying power on the strategy that has the highest growth rate.

### 7.2 Optimization of Expected Growth Rate using Simulated Returns

This section deals with the situation where student  $t$  distribution is used to compute the kelly fraction. There is no neat closed form after relaxing gaussian assumption. One must do it numerically. The author uses the returns from the mean reversion strategy described in one of the previous chapters in the book.

```

pathname      <- file.path(".", "/data/AUDCAD_unequal_ret.mat")
data          <- readMat(pathname)
mu            <- mean(data$ret)
stdev         <- sd(data$ret)
(f.star       <- (252*mu)/(sqrt(252)*stdev)^2)

## [1] 23.85

```

Thus optimal kelly under gaussian is 24. To relax the gaussian distribution, one can fit a  $t$  distribution and then simulate from that distribution

```

fit           <- fitdistr(data$ret, "t")
returns.sim   <- fit$estimate[1] + fit$estimate[2] * rt(10000, fit$estimate[3])
g             <- function(f, R){ sum(log(1+f*R))/length(R)}
myf           <- 0:40;
myg           <- rep(NA, 41)
for(f in seq_along(myf)){
  myg[f]      <- g(f, returns.sim)
}
(myf[which.max(myg)])

## [1] 15

```

```

plot(myf, myg, type="l", col = "blue", xlab = "f", ylab = "g", cex.lab=0.7, cex.axis=0.7)

```

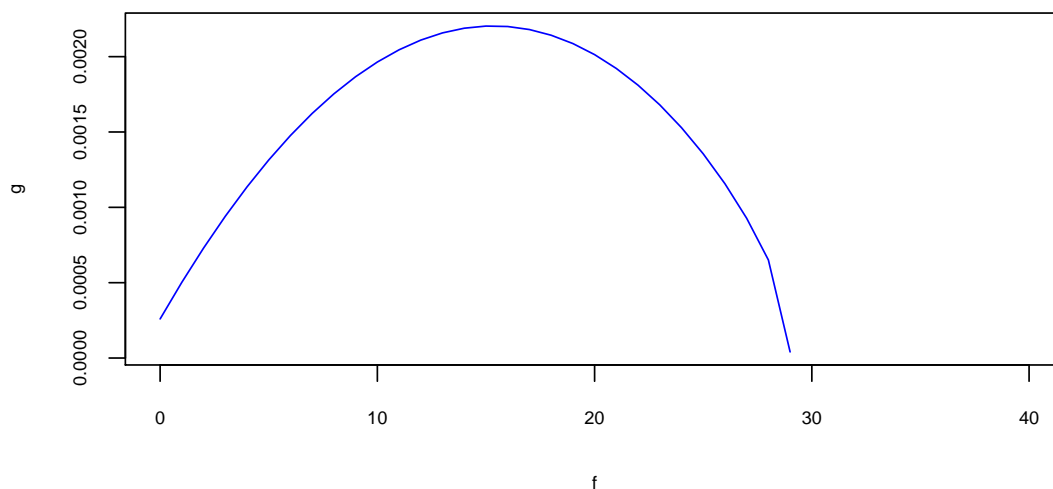


Figure 7.1: Expected Growth Rate  $g$  as Function of  $f$ .

By simulating via a  $t$  distribution, the optimal fraction is around 15.



### 7.3 Optimization of Historical growth rate

Instead of optimizing the expected value of the growth rate using our analytical probability distribution of returns, one can of course just optimize the historical growth rate in the backtest with respect to the leverage.

```
g          <- function(f){ sum(log(1+f*data$ret))/length(data$ret)}  
(f.hist    <- optimize(g,interval=c(1,40),maximum=TRUE)$maximum)  
  
## [1] 23.72
```

This method suffers the usual drawback of parameter optimization in backtest: data-snooping bias. In general, the optimal leverage for this particular historical realization of the strategy returns won't be optimal for a different realization that will occur in the future. Unlike Monte Carlo optimization, the historical returns offer insufficient data to determine an optimal leverage that works well for many realizations. Despite these caveats, brute force optimization over the backtest returns sometimes does give a very similar answer to Kelly leverage.

### 7.4 Maximum Drawdown

In the case when a constraint of max drawdown is imposed, it is not straight forward to compute the kelly fraction. Hence one might have to do trial and error to get to the right leverage. As an input to this exercise, one can use the historical returns or the simulated returns. Each of the input has its own advantages and disadvantages. A good compromise may be a leverage somewhere in between those generated by two methods.

### 7.5 Constant Proportion Portfolio insurance

This strategy entails applying the kelly criterion to the max drawdown  $D$ . The rest of the capital sits in cash. In this way, the max drawdown is adhered to by design. If the trading strategy is profitable and the total account equity reaches a new high water mark, then we can reset our subaccount equity so that it is again  $D$  of the total equity, moving some cash back to the "cash" account. However, if the strategy suffers losses, we will not transfer any cash between the cash and the trading subaccount. Of course, if the losses continue and we lose all the equity in the trading subaccount, we have to abandon the strategy because it has reached our maximum allowed drawdown of  $-D$ . Therefore, in addition to limiting our drawdown, this scheme serves as a graceful, principled way to wind down a losing strategy. One problem using CPPI is that you can't prevent a big drawdown happening during the overnight gap.

### 7.6 Stop Loss

If you believe in mean reverting signal, then the concept of the spread reaching a new low or new high might seem a temporary aberration. But the problem is that there could have been a regime shift in which case one has to close out the position. There is always this risk when one trades a mean reverting strategy. The author suggests a simple method of having a stop loss that is bigger than the backtest intraday drawdown. As an aside, momentum strategies benefit from stoploss. If a momentum strategy is losing, it means that momentum has reversed, so logically we should be exiting the position and maybe even reversing the position. Thus, a continuously updated momentum trading signal serves as a de facto stop loss. This is the reason momentum models do not present the same kind of tail risk that mean-reverting models do.

Key summary points from the chapter

- Maximization of long-term growth rate:
  - Is your goal the maximization of your net worth over the long term? If so, consider using the half-Kelly optimal leverage.
  - Are your strategy returns fat-tailed? You may want to use Monte Carlo simulations to optimize the growth rate instead of relying on Kelly's formula.
  - Keeping data-snooping bias in mind, sometimes you can just directly optimize the leverage based on your backtest returns' compounded growth rate.
  - Do you want to ensure that your drawdown will not exceed a preset maximum, yet enjoy the highest possible growth rate? Use constant proportion portfolio insurance.
- Stop Loss:
  - Stop loss will usually lower the backtest performance of mean-reverting strategies because of survivorship bias, but it can prevent black swan events.
  - Stop loss for mean-reverting strategies should be set so that they are never triggered in backtests.
  - Stop loss for momentum strategies forms a natural and logical part of such strategies.
- Risk indicators:
  - Do you want to avoid risky periods? You can consider one of these possible leading indicators of risk: VIX, TED spread, HYG, ONN/OFF, MXN.
  - Be careful of data-snooping bias when testing the efficacy of leading risk indicators.
  - Increasingly negative order flow of a risky asset can be a short-term leading risk indicator