Kelly Criterion Under Levy Processes

This is a very, very rough draft with many incomplete thoughts, typos, and errors

Austin Pollok

3rd May 2019

Contents

Part I

History and Possible Uses

Chapter 1

History of Kelly Criterion

The Kelly Criterion was first formalized by John Kelly in 1956 (Kelly, 1956) and later popularized by Ed Thorp when he used it to determine bet sizes in his newly discovered card counting system (Thorp, 1966). Kelly thought of this as an application of Claude Shannon's theory of communication while working together at Bell Labs. There was no thought of utility maximization, it was developed solely from an information theoretic interpretation of gambling with an edge. Thorp later used it when gambling in the securities market operating as the first quantitative market neutral hedge fund. Note it can be quite useful to think of investing in markets as gambling in casinos (quote from Thorp's autobiography)...

It was first called Kelly Criterion in paper by Thorp and Walden 1966, "A Favorable Side Bet in Nevada Baccarat."

The Kelly Criterion is nothing more than way to determine how to optimally bet when gambling in things with random returns so as to maximize our long term wealth. A so called "theory of re-investment," as opposed to a theory of investment (such as Markowitz theory), (Poundstone, 2006). This system only makes sense if we are placing a large number of bets and trying to compound our wealth. This opposes the Markowitz mean variance criterion which only applies to a single bet with no consideration of compounding. Intuitively, the Kelly Criterion tries to formalize the common gambler's saying of "try to play with the house's money."...Note briefly compare the two strategies from RedBloodedRisk...Also see NassimTalebs writings on KellyCriterion...Also see Bill Gross's notes on Kelly....other investors such as James Simons and Warren Buffett...Also note the scandal with Paul Samuelson......

....Need to note how it was thought about in context of gambling and how it was later thought about terms of managing a portfolio...not in the context of utility theory...effectively it says what is rational in the long run....

..."Kelly is not limited to two value payoffs, but applies generally to any gambling or investing situation **in which the probabilities are known or can be estimated** " (note this is a problem in estimation that possibly ML techniques can help with). "the investor or bettor generally avoids total loss; the bigger the edge, the larger the bet; the smaller the risk, the larger the bet." "The Kelly Criterion, not having been invented by the old-line academic economists, has generated considerable controversy." "Three caveats: the Kelly Criterion may lead to wide swings in the total wealth, so most users choose to bet some lesser fraction, typically one-half Kelly or less (this problem may be analyzed from fluctuation/ruin theory of Levy processes); for investors with short time horizons or who are averse to (volatility) risk, other approaches may be better; an exact application of Kelly requires exact probabilities of payoffs such as those in most casino games; to the extent these are uncertain, which is generally the case in the investment world, the Kelly bet should be based on a conservative estimate of the outcome." (this problem could maybe be resolved by thinking about model uncertainty framework, or non-stationary independent increment processes). "The Kelly Criterion highlights the perils of overbetting even when you have the edge." (Thorp, 2017) .....See Aaron Brown's description of Kelly in his book and on post on Quora...See all of Thorp's writings on it including his analysis of casino games book....., see papers in (Thorp, MacLean, & Ziemba, 2011).

Chapter 2

Possible Uses and Empirical Work

Here need to add a motivation as to how and why particular types of traders/gamblers/investors would use this system. (Include lots of pictures and possibly swap with first section). IN PARTICULAR THINK ABOUT HOW TO USE THE KELLY CRITERION IN A TAIL RISK STRATEGY/HEDGING PORTFOLIO.

Need to add empirical data analysis work here using cross-sectional, time-series, machine learning tools. Some simple top of the head strategies could be:

- (long vol, long gamma) Buying out of the money calls and puts.
- (short vol, short gamma) Selling out of the money calls and puts and collecting premiums.
- As Taleb mentioned, knowing how to determine when other people/firms/strategies are taking tail risk without knowing (could be from using thin tailed models to evaluate risk, or only evaluating risk based on past events) and somehow short them, or bet on them blowing up while remembering the quote from Keynes about the market remaining irrational longer than you can stay solvent.
- Is there a relation between leverage and chance of blowing up, and if so how to measure this relation and how to measure leverage in a given strategy.

These all have to do with how to value (whatever that means) tail risk, and knowing how to buy cheap insurance and lotto tickets and sell expensive insurance and lotto tickets that you know you can cover. One also has to be aware of non-linear time decay of our gambles, and the costs of our gambles. We should also investigate a survey of betting on unobserved events, or an empirical investigation of 'statistics of unobserved events' to help justify the fat tailed stable distribution assumptions.

Part II

Mathematical Formalization of Kelly Criterion

Chapter 3

Kelly Criterion in Discrete Time

We are essentially following (Thorp, 2006). Setup: We are repeatedly placing bets in a sequence of random gambles (such as stocks, bonds, casino games, etc) and we are trying to decide how much as percentage of current wealth, $f \in [0,1]$ (corresponding to no short selling and no leverage/use of margin), we should bet to maximize our long term wealth. Note that we are apriori **assuming** that f is a fixed constant as opposed to a predictable process $\{f_k\}_{k\geq 1}$. This is justified intuitively by the fact of iid gambles/returns. However there are rigorous justifications of this fact (using log utility yields optimal proportion/fraction is constant in particular cases), see papers by Browne and Whitt, Kalaba and Bellman, and paper by Kallsen (references to be added in later). Also, we are assuming the returns of the gambles are independent and identically distributed, $\{r_k\}_{k\geq 1}$, $r_k \geq -1, \forall k \geq 1$.

(Note that we are thinking of the returns $r_k = \frac{P_k - P_{k-1}}{P_{k-1}}$ for some price process P. One possible question is what happens when we assume non-iid gambles, such as an independent increment processes, or various distributions, such as fat tailed distributions?)

Assume we're sequentially placing bets/investments on a sequence of independent and identically distributed payoffs/returns. Let $\{W_1^{n,f}\}_{n\geq 1}$ be the wealth dynamics achieved from betting the fraction of current wealth f in the sequence of gambles/returns $\{r_n\}_{n\geq 1}$, with $r_n \geq -1$, and compounding my bankroll/wealth. In other words, the wealth dynamics are modeling the process of compounding returns with a constant exposure of (proportion) f to those returns.

Remark. *Note that we interpret "discrete time" the same as discretely compounding over the unit interval $[0,1]$, which explains the notation $W_1^{n,f}$. This allows for an easy generalization to continuous time. The important part is that time is regarded as a point where we place a bet. So when we let $n \to \infty$, it can be interpreted as time (number of bets) increasing, or equivalently the frequency of bets per unit interval increasing.*

3.1 Discrete Time Wealth Dynamics

$W_0 = w =$ initial wealth/bankroll/portfolio size

$W_1^{1,f} = W_0 + r_1(fW_0) = W_0(1 + fr_1)$

$W_1^{2,f} = W_1^{1,f} + r_2(fW_1^{1,f}) = W_1^{1,f}(1+fr_2) = W_0(1+fr_1)(1+fr_2)$

$\vdots$

$$W_1^{n,f} = W_1^{n-1,f} + r_n(fW_1^{n-1,f}) = W_1^{n-1,f}(1+fr_n) = W_0 \prod_{k=1}^{n}(1+fr_k)$$

3.2 Kelly Criterion

$$\frac{W_1^{n,f}}{W_0} = \mathrm{e}^{n\ln(\frac{W_1^{n,f}}{W_0})^{\frac{1}{n}}} = \mathrm{e}^{n\frac{1}{n}\sum_{k=1}^{n}\ln(1+fr_k)} \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{\infty\mathbf{E}[\ln(1+fr_1)]}$$

a.s. by the Strong Law of Large Numbers.

We set $G_n(f) := \ln(\frac{W_1^{n,f}}{W_0})^{\frac{1}{n}}$ = growth rate at the n^{th} bet, and it measures the exponential rate of increase per bet. Also, note it is the log of the geometric average of my first n returns.

Also, set $g_{r_1}(f) := \mathbf{E}[\ln(1+fr_1)]$ = long term growth rate, which is the rate at which we expect, in the long run, to compound our wealth. So, intuitively, our goal is to maximize our long term wealth process which is equivalent to maximizing our long term growth rate.

This leads to the Kelly Criterion:

$$\max_f g_{r_1}(f) = \max_f \mathbf{E}[\ln(1+fr_1)]$$

Remark. *This setup lends itself to see how this problem is thought about and how to formalize it for more complicated return sequences. Also we want to make this a vector of returns for more general settings.*

Remark. *The Kelly criterion has many other names depending on which point of view you are coming from. Most in the gambling, trading, or information theory (see papers by Thomas Cover) community appear to call it the Kelly criterion, while most in the academic finance literature tend to call it optimal long term growth or log utility.*

(I need to investigate some of the other finance papers, such as those from Markowitz, Merton, Karatzas, Kardaras, etc)

3.3 Examples

Let $r_k = \begin{cases} \text{b} & \text{with probability p} \\ -1 & \text{with probability 1-p} \end{cases} = b^{Y_k}(-1)^{1-Y_k}$

Where $\{Y_k\}$ are iid Bernoulli(p) random variables.
We also assume we have an edge, i.e. $\mathbf{E}[r_k] > 0$.
$\implies \cdots \implies f^* = \frac{bp-q}{b} = \frac{\text{edge}}{\text{odds}}$

(Note we need to work out the details here and explain the implications, also include a simulation of wealth processes and include a graph of growth rate vs. fraction to see how bet sizes affects growth rates. Also include various theorems about the optimal properties of the Kelly as explored in Breiman and Cover, etc. Also include the paper about essentially different strategies and how a non-Kelly fraction can dominate in long run (Thorp, 2012))

3.4 Properties of Kelly Criterion in Discrete Time

I need to summarize the properties of the Kelly as investigated by Brieman and later Cover. Also, here I need to put the reason that having an edge implies positive long term growth, see (Thorp, 2006) or Ethier's Doctrine of Chances page 359. Basically, $g_r^{''}(f) < 0$ implying strict concavity and $g_r^{'}(0) = \mathbf{E}[r] > 0$ with $g_r(0) = 0$ and $g_r(1-) = -\infty$ which implies there is an $f \in [0, 1)$ such that $g_r(f) > 0$ by the intermediate value theorem. See Ethier's book page 361 for a CLT-type result for the geometric average growth rate, and try to get this using Discretization of Processes book for the continuous time results. See (Ethier, 2010).

Chapter 4

Continuous Time Wealth Dynamics Under Levy Processes

4.1 Why Go To Continuous Time?

explain uses of cts time and cts distribution and relation to investing and adding more realistic models..i currently believe it is to use the tools of stochastic calculus..are there gambling uses? Note how the realities of transaction costs and frequency can render continuous time methods and models useless, so these need to be considered, along with taxes!!

4.2 From Discrete to Continuous Time

Assume we are gambling/investing/betting on a sequence of payoffs/returns at an arbitrary (deterministic) frequency/rate n per period.

Assume the sequence of payoffs/returns that we're gambling/investing/betting in are independent and identically. In particular, $\forall n$, the sequence $\{r_{n,k}\}_{k=1}^{n}$, is independent and identically distributed, with $r_{n,k} \stackrel{d}{=} r_n$, and $r_{n,k} \geq -1$.

Remark. *NOTE: We are first assuming $\{r_{n,k}\}$ is i.i.d. (random walk), which should lead to Levy processes. Next, we should weaken i.i.d and assume $\{r_{n,k}\}$ independent, but not identically distributed, which should lead to additive or non-stationary independent increment processes. This would partially answer a question posed by Thorp. In this scenario, it would make sense for our f to be changing depending on the distribution we're currently betting on, so $\{f_{n,k}\}$. And finally, we would like to assume some dependence structure in the sequence $\{r_{n,k}\}$, for example Cov$(r_{n,k}, r_{n,j}) = \gamma_n(k,j)$, but the dependence doesn't need to be measured by the covariance(could be codifference or covariation for stable distributions). Finally, we would like to do these three scenarios but change the frequency n to be random N_n,i.e. a random time change. This is worked out in the i.i.d. case for the example of the binomial model in (Rachev & Rushendorf, 1994).*

Remark. *Here we only assume $\{r_{n,k}\} \in \mathbf{R}$, so we're making one gamble at a time as opposed to many simultaneous gambles at once, namely $\{\vec{r}_{n,k}\} = [r^1_{n,k}, \cdots, r^d_{n,k}]^\top \in [-1,\infty)^d$*

4.2.1 Single Period $[0, 1]$

Let $\{W_1^{n,f}\}_{n\geq 1}$ be the wealth dynamics achieved from betting the fraction of current wealth f in the sequence of gambles/returns $\{r_{n,k}\}_{k=1}^n$, and compounding my bankroll/wealth over a single period. In other words, the wealth dynamics are modeling the process of compounding returns with a constant exposure of (proportion) f to those returns.

$W_0 = w =$ initial wealth/bankroll/portfolio size

$W_1^{1,f} = W_0 + r_{n,1}(fW_0) = W_0(1 + fr_{n,1})$

$$W_1^{2,f} = W_1^{1,f} + r_{n,2}(fW_1^{1,f}) = W_1^{1,f}(1 + fr_{n,2}) = W_0(1 + fr_{n,1})(1 + fr_{n,2})$$

$\vdots$

$$W_1^{n,f} = W_1^{n-1,f} + r_{n,n}(fW_1^{n-1,f}) = W_1^{n-1,f}(1 + fr_{n,n}) = W_0 \prod_{k=1}^{n}(1 + fr_{n,k})$$

Now, we wish to find the dynamics of the limiting wealth variable as the frequency of bets/investments per period gets larger and larger.

$$W_1^{n,f} \xrightarrow[n\to\infty]{d} W_1^{\infty,f} =: W_1^f$$

$$\iff \ln(\frac{W_1^{n,f}}{W_0}) = \sum_{k=1}^{n} \ln(1 + fr_{n,k}) \xrightarrow[n\to\infty]{d} \ln(\frac{W_1^f}{W_0})$$

We assume $f \in [0, 1)$ as $f = 1$ is overbetting and will lead to ruin almost surely. Thus $fr_{n,k} > -1 \implies 1 + fr_{n,k} > 0 \implies \ln(1 + fr_{n,k}) \in \mathbf{R}$.

Note that we need to prove convergence in distribution and justify that the limit, $\ln(\frac{W_1^f}{W_0})$, is infinitely divisible. Does this imply W_1^f is infinitely divisible? (Attempt to prove using both 507 characteristic function methods and using results from (Jacod & Shiryaev, 2003)) Next, get convergence rates, something like Berry-Esseen using Kolmogorov-Smirnoff distances, as there is a very practical interpretation to what games you can analyze with the continuous time results. Finally, we need many graphs and simulations showing tractability of these methods, including description and computation of the Levy characteristics of the limit from the approximating sequence.

4.2.2 Arbitrary Period $[0, t]$

Now we consider betting over a time horizon $[0, t]$ or, equivalently, betting over t-many periods, with n being the frequency/rate of bets being placed per period. Let $\{W_t^{n,f}\}_{n\geq 1}$ (note if t is free, we're dealing with a sequence of stochastic processes here as opposed to a sequence of random variables as above or if we fix t) be the wealth dynamics achieved from betting the fraction of current wealth f in the sequence of gambles/returns $\{r_{n,k}\}_{k=1}^n$, per period, and compounding my bankroll/wealth over t-many periods. Let $W_0 = w =$ initial wealth/bankroll/portfolio size. Then the wealth process given by compounding returns n-many times per period for $\lfloor t \rfloor$-many periods with a constant exposure of f to those returns is given by:

$$W_t^{n,f} = W_{t-1}^{n,f}\Big(\prod_{k=1}^{n}(1+fr_{n,k})\Big) \text{ , and } W_1^{n,f} = W_0\prod_{k=1}^{n}(1+fr_{n,k})$$

$$\implies W_t^{n,f} = W_0\Big(\prod_{k=1}^{n}(1+fr_{n,k})\Big)^{\lfloor t \rfloor} = W_0\prod_{k=1}^{\lfloor nt \rfloor}(1+fr_{n,k})$$

This time, instead of finding the limiting wealth variable, we wish to find the limiting wealth process as the frequency of my bets/trades per period gets larger and larger.

$$W_t^{n,f} \xrightarrow[n\to\infty]{\mathcal{L}} W_t^{\infty,f} =: W_t^f \quad \text{in } \mathbf{D}[0,\infty)$$

$$\iff \ln\Big(\frac{W_t^{n,f}}{W_0}\Big) = \sum_{k=1}^{\lfloor nt \rfloor}\ln(1+fr_{n,k}) =: \sum_{k=1}^{\lfloor nt \rfloor}X_{n,k}^f \xrightarrow[n\to\infty]{\mathcal{L}} \ln\Big(\frac{W_t^f}{W_0}\Big) \quad \text{in } \mathbf{D}[0,\infty)$$

We again note that we need to prove this convergence in the space of cadlag functions and also verify $\ln(\frac{W_t^{n,f}}{W_0})$ being a Levy process implies the actual wealth process is also Levy. One can hope to prove this convergence by the trick in section 2 of (Rachev & Rushendorf, 1994), or by using tightness, fdd convergence and characterization of limit by fdd. Or finally, what I am thinking is the most tractable as characteristics of Levy processes are supposedly tractable by FFT methods, prove tightness, convergence of triples, and characterization of limiting process by triples.

After establishing this, one would hope to get convergence rates that are uniform in t, and finally do many simulations.

Question: Should the sequence of returns depend on the time horizon? Maybe realistically, but for this model we assume each period, with number of gambles per period being fixed, we're betting on the same (statistically) sequence of gambles/investments.

Now that we get the continuous time wealth process, we should apply Ito's formula (for Levy processes or semi-martingales) to get the dynamics to see what SDE the wealth process satisfies and one could use tools of stochastic calculus to help answer some of the above and below questions, as well as for simulation.

Remark. *Here we list some examples that quickly come to light which fall under our setting of* $\forall n$, $\{r_{n,k}\}$ *i.i.d., and* $r_{n,k} \geq -1$. *First we can consider a biased coin,* $r_{n,k} \stackrel{d}{=} r = \begin{cases} 1, & w/p\ p \\ -1, & w/p\ q \end{cases}$ *with* $p > q$. *Or more generally, a binary payoff, which could model sports betting, loans with simple terms (contractually simple loans), etc.,* $r_{n,k} \stackrel{d}{=} r = \begin{cases} b, & w/p\ p \\ -1, & w/p\ q \end{cases}$. *These were worked out in (Thorp, 2006), and also above.*

We could also consider betting on a financial product with some price or value process. Specifically, we would be placing bets on the random returns of this process. In this scenario, we have

$$r_{n,k} = \frac{P_{k/n} - P_{(k-1)/n}}{P_{(k-1)/n}} = \mathrm{e}^{R_{k/n} - R_{(k-1)/n}} - 1 \stackrel{d}{=} \mathrm{e}^{R_{1/n}} - 1 =: r_n \quad \geq -1$$

We could also consider horse racing, roulette, or maybe approximately blackjack, see (Thorp, 2006). One should also consider trading/managing a sequence of derivatives, and consider making loans or investing in MBS.

Finally, we also have the binomial model as a primary example, and can be used in numerical approximation.

4.3 Example: Using Binomial Tree (Discrete Time, Discrete Space) Approximations to Continuous Time

Here we also first follow (Thorp, 2006).

It suffices to use continuous time dynamics when we can think of our wealth at any point in time as the result of compounding many bets. Note, compounding bets could simply be thought of as readjusting portfolio holdings based off of a proportional strategy, i.e. a strategy that gives a percentage of current bankroll.

It is of interest to determine how quickly the discrete approximation converges to the continuous time wealth process to give us a precise gauge of about how many bets/readjustments are needed to use the continuous time dynamics(or at least the frequency of bets per unit time, whatever time may be: seconds, minutes, hours, days, weeks, months, years) (possibly see papers by Remi and Fanhui to get convergence rates of discretized SDEs).

Fix an arbitrary $t \in [0, \infty)$, and let n be the frequency/rate of bets placed in the window $[0, \lfloor nt \rfloor]$. In particular, n is the number of bets per unit time (i.e. in $[0, 1]$), and nt is the number of bets per time t.

Let $\mu \in \mathbf{R}$, $\sigma > 0$, $\{Y_k\}$ iid Bernoulli($\frac{1}{2}$),

$r_{n,k} = \begin{cases} \frac{\mu}{n} + \frac{\sigma}{\sqrt{n}} & \text{with probability} \frac{1}{2} \\ \frac{\mu}{n} - \frac{\sigma}{\sqrt{n}} & \text{with probability} \frac{1}{2} \end{cases} = (\frac{\mu}{n} + \frac{\sigma}{\sqrt{n}})^{Y_k}(\frac{\mu}{n} - \frac{\sigma}{\sqrt{n}})^{1-Y_k}$

Note $r_{n,k} = \frac{\mu}{n} + \zeta_{n,k}$ where $\zeta_{n,k} = (\frac{\sigma}{\sqrt{n}})^{Y_k}(-\frac{\sigma}{\sqrt{n}})^{1-Y_k}$

Assume we have an edge (note that this is the hard part to find in practice), $\mathbf{E}[r_{n,k}] = \frac{\mu}{n} > 0$

Consider the approximation $\forall t$: (note we need to prove this convergence in law on the space $\mathbf{C}$ or $\mathbf{D}$ following methods, use characteristics, in (Jacod & Shiryaev, 2003))

$W_t^{n,f} = W_0 \prod_{k=1}^{\lfloor nt \rfloor}(1 + f r_{n,k}) \xrightarrow[n\to\infty]{d} W_t^f$

Equivalently, consider, $\forall t$, the log returns of wealth: (note need to fix equation spacing)

$$\ln(\frac{W_t^{n,f}}{W_0}) = \sum_{k=1}^{\lfloor nt \rfloor} \ln(1 + f r_{n,k}) = \sum_{k=1}^{\lfloor nt \rfloor}(f r_{n,k} - \frac{f^2}{2} r_{n,k}^2 + H.O.T)$$

$$= (f\mu t) + f\sum_{k=1}^{\lfloor nt \rfloor}\zeta_{n,k} - \frac{f^2}{2}(\frac{\mu^2 t}{n} + 2\frac{\mu}{n}\sum_{k=1}^{\lfloor nt \rfloor}\zeta_{n,k} + \sigma^2 t) + H.O.T.$$

$$\xrightarrow[n\to\infty]{d} f\mu t + f\sigma B_t - \frac{f^2}{2}\sigma^2 t = f(\mu t + \sigma B_t) - \frac{f^2}{2}\sigma^2 =: \ln(\frac{W_t^f}{W_0})$$

$$\iff W_t^{n,f} \xrightarrow[n\to\infty]{d} W_t^f = W_0 e^{f(\mu t + \sigma B_t) - \frac{f^2}{2}\sigma^2} = W_0 \mathcal{E}(f(\mu t + \sigma B_t))$$

We see this is true since

$$\varphi_{\sum_{k=1}^{\lfloor nt\rfloor}\zeta_{n,k}}(u) = (\varphi_{\zeta_{n,k}}(u))^{\lfloor nt\rfloor} = [\cos(\frac{u\sigma}{\sqrt{n}})^n]^t \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-t(\frac{u^2\sigma^2}{2})} = \varphi_{\sigma B_t}(u)$$

Note, $\{B_t\}$ is Brownian motion, and $\mathcal{E}$ is the stochastic exponential operator, see (Jacod & Shiryaev, 2003).

Also in terms of the interpretation, note, $W_t^{n,f}$ is the wealth/bankroll/portfolio value corresponding to compounding at frequency/rate n per period, for t many periods, at (random) rate of return $\ln(\frac{W_t^{n,f}}{W_0})$. Then, as we send $n \to \infty$, i.e. the frequency of betting per period is getting larger and larger, we get W_t^f is the wealth/bankroll/portfolio value corresponding to continuously compounding over t many periods at random rate of return $\ln(\frac{W_t^f}{W_0})$.

Applying Ito's formula to the limiting wealth process gives us the linear (controlled) SDE:

$$\begin{cases} dW_t^f = fW_t^f(\mu dt + \sigma dB_t) \\ W(0) = W_0 \end{cases}$$

To extend this idea to more general distributions, not just lognormal, we follow (Rachev & Rushendorf, 1994) which characterizes all the infinitely divisible limits of the binomial model when you randomize some of the parameters. In particular, the frequency (number) of bets per period, and the up and down sizes of the returns.

(I think it could be worth determining limits of other classes, such as non-iid summands leading to independent increment processes, or correlated summands possibly leading to models like fractional Brownian motion)

Let $r_{n,k} = \begin{cases} u_n & \text{w.p.} p_n \\ d_n & \text{w.p.} q_n \end{cases} = (u_n)^{Y_k}(d_n)^{1-Y_k}$

Where $\{Y_k\}$ are iid Bernoulli(p_n) random variables.

Again, we consider $W_t^{n,f} = W_t^{n-1,f}(1 + fr_{n,\lfloor nt\rfloor}) = W_0 \prod_{k=1}^{\lfloor nt\rfloor}(1 + fr_{n,k})$

$$\implies \ln(\frac{W_t^{n,f}}{W_0}) = \sum_{k=1}^{\lfloor nt\rfloor} \ln(1 + fr_{n,k}) =: \sum_{k=1}^{\lfloor nt\rfloor} X_{n,k}^f$$

Where $X_{n,k}^f := \ln(1 + fr_{n,k}) = (Y_k)\ln(1 + fu_n) + (1 - Y_k)\ln(1 + fd_n)$, again, with $\{Y_k\}$ being iid Bernoulli(p_n), for $1 \le k \le \lfloor nt\rfloor$. Of course we also need to make some assumptions on our up and down probabilities, namely we need our summands to be uniformly asymptotically negligible (UAN), so we are actually looking in the class of infinitely divisible limits of the binomial model. Namely, $\forall\epsilon > 0$, $\max_{1\le k\le\lfloor nt\rfloor} \mathbf{P}(|X_{n,k}| > \epsilon) \underset{n\to\infty}{\longrightarrow} 0$. The condition of UAN puts conditions on the limit of the sequence of ups and downs, $\{u_n\}$ and $\{d_n\}$, by definition of $X_{n,k}$.

We need to prove that

$$\ln(\frac{W_t^{n,f}}{W_0}) = \sum_{k=1}^{\lfloor nt\rfloor} \ln(1 + fr_{n,k}) =: \sum_{k=1}^{\lfloor nt\rfloor} X_{n,k} \underset{n\to\infty}{\xrightarrow{\mathcal{L}}} \ln(\frac{W_t^f}{W_0}) \text{ on the space } \mathbf{D}([0,\infty)).$$

NOTE: First, to prove the above convergence in distribution, for fixed $t \in [0, \infty)$, we need the tools of characteristic functions and infinitely divisible distributions. Second, to prove this convergence in law as random elements of function space, we can use two methods as described in the introduction of (Jacod & Shiryaev, 2003). The traditional Skorokhod way is done in section 2 in (Rachev & Rushendorf, 1994). We need to establish this convergence using the convergence of the characteristics as described in (Jacod & Shiryaev, 2003). Finally consider a random time change and check if it falls into the framework of the above sections.

(Need to Add computer simulations of these results)

4.4 Example: Using Continuous Time, Continuous Space (Stochastic Calculus) Approximations to Continuous Time

We can use stochastic exponential Levy models and relies more on the theory of stochastic calculus and methods of (Jacod & Shiryaev, 2003). Note the stochastic exponential can be non-positive, so one needs to ensure that this doesn't happen.

Or we can use ordinary exponential Levy models which makes the stochastic calculus more difficult, though it makes more sense from a modeling and estimating from discrete returns point of view. Moreover, it turns out these are equivalent. See chapter 4 of (Barndorff-Nielsen & Shiryaev, 2010).Claims stochastic exponential is simple interest (seems like discrete compounding but not sure) and ordinary exponential is compound interest (seems like continuous compounding). Also see Essentials of Stochastic Finance by Shiryaev, page 83 I think.

Assume a price process, which we're betting/gambling on the incremental returns or investing in, follows an ordinary exponential, $P_t = P_0 \mathrm{e}^{R_t}$, for R a Levy process with characteristics (b, c, F), and thus a Levy-Ito decomposition of

$$R_t = \underbrace{bt + \sqrt{c}W_t}_{\text{Brownian motion with drift}} + \underbrace{x1_{\{|x|\leq 1\}} * (\mu^R - \mathrm{mF})_t}_{\text{square integrable (compensated) martingale with countably many jumps}} + \underbrace{x1_{\{|x|>1\}} * \mu^R_t}_{\text{compound Poisson process}}$$

Equivalently, $dP_t = P_{t-}(dR_t + \frac{c}{2}dt + \mathrm{e}^{\Delta R_t} - 1 - \Delta R_t)$ with an initial price P_0. Note that $\mu^R(\omega; ds, dx) = \sum_{s>0} \delta_{(s, \Delta R_s(\omega))}(ds, dx)$ is the jump measure of the process R (Eberlein, 2009).

Then consider compounding at frequency n over a period $[0, t]$, and set $r_{n,k} := \dfrac{P_{k/n} - P_{(k-1)/n}}{P_{(k-1)/n}} = e^{R_{k/n} - R_{(k-1)/n}} - 1$,

$$W_t^{n,f} = W_0 \prod_{k=1}^{\lfloor nt \rfloor} (1 + f r_{n,k}) = W_0 \prod_{k=1}^{\lfloor nt \rfloor} \left(1 + f\left(\frac{P_{k/n} - P_{(k-1)/n}}{P_{(k-1)/n}}\right)\right)$$

$$\iff \ln\big(\frac{W_t^{n,f}}{W_0}\big) = \sum_{k=1}^{\lfloor nt \rfloor} \ln(1+f\big(\frac{P_{k/n}-P_{(k-1)/n}}{P_{(k-1)/n}}\big)) = \sum_{k=1}^{\lfloor nt \rfloor} \ln(1+f\big(e^{R_{k/n}-R_{(k-1)/n}}-1\big))$$

By Ito's formula, for $\frac{k-1}{n} \leq s \leq \frac{k}{n}$

$$\ln(1+f\big(e^{R_s-R_{(k-1)/n}}-1\big)) = \int_{\frac{k-1}{n}}^{s} \frac{fe^{R_{u-}-R_{(k-1)/n}}}{1+f(e^{R_{u-}-R_{(k-1)/n}}-1)} dR_u +$$

$$\frac{c}{2} f(1-f) \int_{\frac{k-1}{n}}^{s} \frac{fe^{R_{u-}-R_{(k-1)/n}}}{(1+f(e^{R_{u-}-R_{(k-1)/n}}-1))^2} du +$$

$$\int_{\frac{k-1}{n}}^{s} \int_{\mathbf{R}} \ln(1+f(e^{R_{u-}+x-R_{(k-1)/n}}-1)) - \ln(1+f(e^{R_{u-}-R_{(k-1)/n}}-1)) -$$

$$fx \frac{e^{R_{u-}-R_{(k-1)/n}}}{1+f(e^{R_{u-}-R_{(k-1)/n}}-1)} \mu^R(du,dx)$$

Now, let $s = \frac{k}{n}$, then

$$\ln\big(\frac{W_t^{n,f}}{W_0}\big) = \sum_{k=1}^{\lfloor nt \rfloor} \ln(1+f\big(e^{R_{k/n}-R_{(k-1)/n}}-1\big)) = \sum_{k=1}^{\lfloor nt \rfloor} \Big(\int_{\frac{k-1}{n}}^{\frac{k}{n}} h_1(\frac{k-1}{n}, R_{s-}) dR_s +$$

$$\int_{\frac{k-1}{n}}^{\frac{k}{n}} h_2(\frac{k-1}{n}, R_{s-}) ds + \int_{\frac{k-1}{n}}^{\frac{k}{n}} \int_{\mathbf{R}} h_3(\frac{k-1}{n}, x, R_{s-}) \mu^R(ds,dx) \Big)$$

$$\implies \ln\big(\frac{W_t^{n,f}}{W_0}\big) = \int_0^t H_1^n(s, R_{s-}) dR_s + \int_0^t H_2^n(s, R_{s-}) ds + \int_0^t \int_{\mathbf{R}} H_3^n(s,x,R_{s-}) \mu^R(ds,dx)$$

Where $H_i^n(s,x) := \sum_{k=1}^{\lfloor nt \rfloor} h_i(\frac{k-1}{n}, x) 1_{[(k-1)/n, k/n)}(s)$, for $i = 1, 2$, and $H_3^n(s,x,y) = \sum_{k=1}^{\lfloor nt \rfloor} h_3(\frac{k-1}{n}, x, y) 1_{[(k-1)/n, k/n)}(s)$

Now we need to prove convergence of three stochastic integrals by showing $H_i^n \underset{n\to\infty}{\longrightarrow} H_i$ implies the corresponding stochastic integrals converge in some sense, $\forall 1 \leq i \leq 3$. Remi suggests trying Lenglart's inequality to get convergence in probability or convergence in expectation with appropriate moment assumptions. I also believe one could try to prove this by keeping track of the characteristics (b, c, F) through the discretizations, limiting operations, and transformations, and prove that they converge to the limiting characteristics. This could also make simulation and approximation more tractable. Also, see chapter 4 of Applebaums book.

I claim that (need to prove)

$$\ln\big(\frac{W_t^{n,f}}{W_0}\big) \underset{n\to\infty}{\longrightarrow} fR_t + \frac{1}{2}f(1-f)[R^c,R^c]_t + \int_0^t\int_{\mathbf{R}}\big[\ln(1+f(e^x-1)) - fx\big]\mu^R(ds,dx)$$

Thus we get the continuous time wealth dynamics, when we're continuously betting/trading (readjusting portfolio) on the increments of the price, for a price process following a log-Levy (or ordinary exponential of a Levy) process.

$$\implies W_t^f := W_0 e^{fR_t+\frac{1}{2}f(1-f)[R^c,R^c]_t+\int_0^t\int_{\mathbf{R}}\big[\ln(1+f(e^x-1))-fx\big]\mu^R(ds,dx)}$$

One question needing to be answered is what are the characteristics of the continuous time wealth process? I think this will be answered if we try to prove using convergence of characteristics. If not, we should be able to track the characteristic starting from the price process and see how a C^2 mapping of them alters the characteristics to get the wealth process. We also need to do this for a predictable process $f = \{f_t\}$, not just a constant f.

Note, for $Q = Q(\omega, s, x)$, we have equivalent notation for the stochastic integral of Q with respect to some random measure (in particular the jump measure of a process $\mu^X(\omega; ds, dx) := \sum_{s>0}\delta_{(s,\Delta X_s(\omega))}(ds,dx)$).

$\int_0^t\int_{\mathbf{R}} Q(\omega,s,x)\mu^X(\omega;ds,dx) = (Q*\mu^X)_t(\omega) = \sum_{s\le t} Q(\omega,s,\Delta X_s(\omega))$. See (Jacod & Shiryaev, 2003)

Now that we have conjectured the continuous time wealth dynamics, corresponding to continuous compounding, we apply Ito's formula to see what type of SDE it solves. Recall, for a Levy process X, we have $d(e^{X_t}) = e^{X_{t-}}(dX_t + \frac{1}{2}d[X^c,X^c]_t + e^{\Delta X_t} - 1 - \Delta X_t)$.

$$\implies dW_t^f = fW_{t-}^f\big(dR_t + \frac{1}{2}d[R^c,R^c]_t + e^{\Delta R_t} - 1 - \Delta R_t\big)$$

Assume a price process, which we're betting/gambling on the returns or investing in, follows a stochastic exponential, $P_t = P_0\mathcal{E}(\tilde{R})_t$, for $\tilde{R}$ a Levy process $\iff dP_t = P_{t-}d\tilde{R}_t$, with initial price P_0. Note the relation between ordinary exponential and stochastic exponential, $e^{R_t} = \mathcal{E}(\tilde{R})_t$, where $\tilde{R}_t = R_t + \frac{1}{2}[R^c,R^c]_t + \sum_{s\le t} e^{\Delta R_s} - 1 - \Delta R_s$, for more details, see either (Jacod & Shiryaev, 2003), or (Cont & Tankov, 2004). By these results, the two methods should be equivalent with the theorem relating ordinary versus stochastic exponential serving as the bridge. Then consider the compounding at frequency n, and set $r_{n,k} := \frac{P_{k/n} - P_{(k-1)/n}}{P_{(k-1)/n}}$,

$$W_t^{n,f} = W_0\prod_{k=1}^{\lfloor nt\rfloor}(1 + fr_{n,k}) = W_0\prod_{k=1}^{\lfloor nt\rfloor}\Big(1 + f\big(\frac{P_{k/n} - P_{(k-1)/n}}{P_{(k-1)/n}}\big)\Big)$$

$$\iff \ln\big(\frac{W_t^{n,f}}{W_0}\big) = \sum_{k=1}^{\lfloor nt\rfloor}\ln(1 + f\big(\frac{P_{k/n} - P_{(k-1)/n}}{P_{(k-1)/n}}\big))$$

Then, $W_t^{n,f} \underset{n\to\infty}{\longrightarrow} W_0\mathcal{E}(f\tilde{R})_t =: W_t^f$ in the sense of uniform convergence in probability. This implies $dW_t^f = fW_{t-}^f d\tilde{R}_t$

See (Kardaras & Platen, 2013).

I also claim that $\mathcal{E}(f\tilde{R})_t = e^{fR_t + \frac{1}{2}f(1-f)[R^c,R^c]_t + \int_0^t \int_{\mathbf{R}} \left[\ln(1+f(e^x-1))-fx\right]\mu^R(ds,dx)}$.
Thus these two modeling approaches are eqivalent, yet when working with data, one would want to use the ordinary exponential to get the distribution of the log returns.

Goal: Verify this setup and the equivalency of these two methods with a few simple processes with increasing complexity/behavior. First take a (simple) Poisson process as my return process, $R_t = N_t = \sum_{s\le t} \Delta N_s \implies \tilde{R}_t = \sum_{s\le t} e^{\Delta N_s} - 1$. Note this doesn't make sense to use as a return process as this says returns are increasing by a unit at random times. Next, let Z_t be a process with finite variation and consider a finite variation process with drift $R_t = bt + Z_t$. In particular if $\displaystyle\int_0^t \int_{\mathbf{R}} |x|\mu^R(ds,dx) < \infty$, then as an example, take $Z_t = \displaystyle\int_0^t \int_{\mathbf{R}} x\mu^R(ds,dx) = \sum_{s\le t} \Delta R_s$, so $R_t = bt + \displaystyle\sum_{s\le t} \Delta R_s \implies \tilde{R}_t = bt + \displaystyle\sum_{s\le t}(e^{\Delta R_s} - 1)$. Finally, we consider a continuous, infinite variation process with drift. Let $R_t = (b - \frac{c}{2})t + \sqrt{c}W_t \implies \tilde{R}_t = bt + \sqrt{c}W_t$

Remark. *Note we need to somehow be able to estimate the distributions of the $r_{n,k}$ bets we're gambling in. Whether it be a nonparametric method, or we build a model and fit parameters from data and check goodness of fit. Finally, we need to be able to estimate the expected return or edge in each of the gambles.*

Remark. *We also need to redo this for $f = \{f_t\}$ predictable fraction processes and consider when this would be necessary and also if the optimal f is still a constant or actually varies with time.*

Chapter 5

Kelly Criterion in Continuous Time Under Levy Processes

Summarize Kelly's conclusion and chapter of book about using it in investing.

The Kelly Criterion, as named by Thorp in CITE, is a decision rule that is meant to be used when I want to know how to size bets so as to best compound my (random) returns.

5.1 Kelly Criterion Setup

Now that we have our continuous time wealth dynamics, we would like to see how the wealth dynamics evolve in the long run while continuously having exposure f to the returns/gambles.

$$\frac{W_t^f}{W_0} = \mathrm{e}^{t\ln(\frac{W_t^f}{W_0})^{\frac{1}{t}}} = \mathrm{e}^{t\frac{1}{t}\ln(\frac{W_t^f}{W_0})} \underset{t\to\infty}{\longrightarrow} \exp\left(\infty \lim_{t\to\infty}\frac{1}{t}\ln(\frac{W_t^f}{W_0})\right)$$

We see if we wish to choose a constant exposure f^* which maximizes my long term wealth, we need choose f^* to maximize my long term growth rate, or equivalently maximize the long term log-geometric average. More formally,

$$f^* = \underset{f\in[0,1)}{\arg\sup}\lim_{t\to\infty}\frac{W_t^f}{W_0} = \underset{f\in[0,1)}{\arg\sup}\lim_{t\to\infty}\frac{1}{t}\ln(\frac{W_t^f}{W_0})$$

Set

$$g(f) := \lim_{t\to\infty}\frac{1}{t}\ln(\frac{W_t^f}{W_0})$$

Where $g(f)$ is the long term growth rate if we can continuously compound (at least approximately, i.e. $n\to\infty$ is justified) over an infinite (or at least long enough) horizon. Note in a finance context, $n\to\infty$ means continuous trading or continuously updating/re-balancing a portfolio's position.

The continuous time Kelly Criterion is, therefore,

$$\sup_{f\in[0,1)} g(f) = \sup_{f\in[0,1)}\lim_{t\to\infty}\frac{1}{t}\ln(\frac{W_t^f}{W_0}) = \sup_{f\in[0,1)}\mathbf{E}\left[\ln(\frac{W_1^f}{W_0})\right]$$

Where the third equality is true for Levy processes with a first moment (essentially SLLN's).

Remark. *Note that we need to establish convergence rates as $t \to \infty$ to tell us how long we have to keep placing our bets for this approximation to be relevant. Another question to be answered is whether or not this strategy is myopic, in other words, is maximizing my growth rate at any horizon/time period the same as maximizing my long term growth rate? For example, is maximizing my wealth over one year the same as maximizing my wealth in the long run? If not, how are they related? More precisely, how is $sup_f \mathbf{E}\big[\frac{1}{t}\ln(\frac{W_t^f}{W_0})\big]$ related to $\sup_f g(f)$?*

Note that working in continuous time means we are continuously placing bets, or continuously trading and adjusting our portfolio's exposure. Obviously this is a mathematical idealization. Thus to justify our use of this idealization, we need to verify:

$$\lim_{t\to\infty}\lim_{n\to\infty}\frac{1}{t}\ln(\frac{W_t^{n,f}}{W_0}) = \lim_{n\to\infty}\lim_{t\to\infty}\frac{1}{t}\ln(\frac{W_t^{n,f}}{W_0})$$

The left hand side corresponds to continuous trading and then letting the horizon go to infinity, a mathematical idealization. While the right hand side corresponds to discretely trading, which is what actually happens, and letting the horizon get larger and larger, then letting the frequency of compounding go to infinity. Note that we should verify that $f_n^* \to f^*$, and that $f_t^* \to f^*$.

Set,

$$\begin{aligned} G_t^n(f) &:= \frac{1}{t}\ln\left(\frac{W_t^{n,f}}{W_0}\right) = \ln\left(\frac{W_t^{n,f}}{W_0}\right)^{\frac{1}{t}} \\ G_t(f) &:= \lim_{n\to\infty} G_t^n \\ G^n(f) &:= \lim_{t\to\infty} G_t^n \end{aligned}$$

We note that $G_t^n(f)$ is the log-geometric average over horizon $[0,t]$ corresponding to discretely compounding n-many times per period $[0,1]$ with constant exposure of f to the bets/returns. In other words, it is the growth rate at time t corresponding to discretely many compounds. Then $G^n(f)$ is the long term growth corresponding to compounding at rate n per period with exposure of f to the gambles. And $G_t(f)$ is the growth rate at time t corresponding to continuous compounding with exposure of f to the continuous returns (i.e. continuous trading).

Remark. *Note that the conditions $n \to \infty$ and $t \to \infty$ means we are placing lots and lots of bets for a long long time. (Is this reasonable, and if so, what types of gambles? Stock market, debt/bond/mortgage markets, casino games, currencies, making loans?)*

5.2 Example: Using Binomial Tree Approximations

Follow idea in(Thorp, 2006).

5.3 Example: Using Characteristics of Exponential of Levy Processes (Wrong Title)

Recall for a price process $P_t = P_0 e^{R_t}$, where R is a Levy process with characteristics (b, c, F) and thus decomposition

$$R_t = bt + \sqrt{c}W_t + \int_0^t \int_{\mathbf{R}} x 1_{|x|\leq 1}(\mu^R - mF)(ds, dx) + \int_0^t \int_{\mathbf{R}} x 1_{|x|>1}\mu^R(ds, dx)$$

NOTE: For this section, as of now, we are assuming all the moments and smoothness we need from the processes involved for the operations we use to make sense. For example, differentiating under integral, finiteness of integrals, or LLN.

Now, consider the average geometric rate of growth (growth rate) over the period $[0, t]$

$$G_t(f) := \frac{1}{t}\ln\big(\frac{W_t^f}{W_0}\big) = \frac{1}{t}\Big(fR_t + \frac{f(1-f)}{2}[R^c, R^c]_t + \int_0^t \int_{\mathbf{R}} (\ln(1 + f(e^x - 1)) - fx)\mu^R(ds, dx)\Big)$$

Then by a Law of Large Numbers result (need finite mean) for Levy processes, we get convergence in some sense (a.s. or in $\mathbf{P}$)

$$\begin{aligned} G_t(f) \underset{t\to\infty}{\longrightarrow} f\mathbf{E}[R_1] + \frac{1}{2}f(1-f)\mathbf{E}\big[[R^c, R^c]_1\big] + \mathbf{E}\big[\int_0^1 \int_{\mathbf{R}} (\ln(1 + f(e^x - 1)) - fx)\mu^R(ds, dx)\big] \\ =: g_R(f) \end{aligned}$$

We can express the long term growth rate, defined above, in terms of the characteristics of R, so that $g_R(f) = g_R(f; b, c, F)$

$$g_R(f) = f\big(b + \int_{\mathbf{R}} x1_{|x|>1}F(dx)\big) + \frac{1}{2}f(1-f)c + \int_{\mathbf{R}} (\ln(1 + f(e^x - 1)) - fx)F(dx)$$

The Kelly criterion seeks to maximize this long term growth rate function. Here we assume we can differentiate under the integral and that the integral is finite, which really is a a condition of the small jumps according to F as we've already assumed things about the large jumps when we assumed finite mean.

$$\begin{aligned} \max_f g_R(f) \implies \frac{d}{df}g_R(f) &= \big(b + \int_{\mathbf{R}} x1_{|x|>1}F(dx)\big) + \frac{c}{2} - fc + \int_{\mathbf{R}} \big[\frac{e^x - 1}{1 + f(e^x - 1)} - x\big]F(dx) = 0 \\ \iff \frac{d}{df}g_R(f) &= \big(b + \frac{c}{2}\big) - fc + \int_{\mathbf{R}} \big[\frac{e^x - 1}{1 + f(e^x - 1)} - x1_{|x|\leq 1}\big]F(dx) = 0 \end{aligned}$$

We need to somehow solve this above equation for $f^* = f^*(b, c, F)$ and see how it depends on the the characteristics (b, c, F). We also need to plot $g_R(f; b, c, F)$.

Chapter 6

Properties of Kelly Criterion in Continuous Time

Need to investigate more of the math finance literature (as they frequently work in continuous time) such as papers by Markowitz, Merton, Karatzas, Kallsen, Goll, and Kardaras).
If you choose to use this decision criterion, what are some properties of your wealth you can expect to happen?

See Ethier's book page 361, (Ethier, 2010), for a CLT-type result for the geometric average growth rate, and try to get this using Discretization of Processes book for the continuous time results.

Chapter 7

Kelly Criterion in Continuous Time with Drawdown Constraint

Here, we need to discuss how major drawdowns can affect your wealth aand long term growth/compounding rate. We need to formally flush out the idea that a single ten percent drop is way worse than ten, one percent drops spread out over some time because of the non-linearity of the wealth process. Cite the example of tax avoidance in Fortune's Formula book. This idea leads us to wanting to add a drawdown constraint to the Kelly criterion.

7.1 Motivation and History

There are lots of writings about fluctuation theory and gambling, see (Feller, 1968), (Feller, 1971), (Wilson, 1970), maybe Ziemba's Beat the Racetrack. Basically says your wealth process can experience surprisingly large amounts of fluctuations away from the long term expected value. So one has to be able to weather these possible bad streaks, and take advantage of the good streaks.

7.2 Formalism/Formulation

Note the difference between a relative (to either largest point or to starting point) and an absolute draw-down. My guess is I would consider relative draw-down if I was working with the wealth process, but I would consider absolute draw-down if I was working with the log-wealth process, namely the log-return process. We would like the probability of a "draw-down" being smaller than a prescribed threshold over a particular window of time to be close to one. One way of formalizing this is: Fix $S, T, \beta \in (0,1), \epsilon \in (0,1)$,

$$\begin{aligned}
&\mathbf{P}(\text{drawdown}_S(t) \leq \beta, \forall t \in [S,T]) > 1-\epsilon \\
\iff &\mathbf{P}\left(\frac{\sup\limits_{S\leq s\leq t} W_s^f - W_t^f}{\sup\limits_{S\leq s\leq t} W_s^f} \leq \beta, \forall t \in [S,T]\right) > 1-\epsilon \\
\iff &\mathbf{P}\left((1-\beta)\sup_{S\leq s\leq t} W_s^f - W_t^f \leq 0, \forall t \in [S,T]\right) > 1-\epsilon
\end{aligned}$$

$$\iff \mathbf{P}\bigg(\sup_{S\le t\le T}\big\{(1-\beta)\sup_{S\le s\le t} W_s^f - W_t^f\big\} \le 0\bigg) > 1-\epsilon$$

$$\iff \mathbf{P}\bigg(\sup_{S\le t\le T}\big\{(1-\beta)\sup_{S\le s\le t} W_s^f - W_t^f\big\} > 0\bigg) \le \epsilon_{S,T}$$

$$\iff \mathbf{P}\bigg((1-\beta)\sup_{S\le t\le T}\sup_{S\le s\le t} W_s^f > \inf_{S\le t\le T} W_t^f\bigg) \le \epsilon_{S,T}$$

Now we need to solve the left hand side for which will give relations between S, T, β, f.

Note, it appears (need to prove) that there is no uniform ϵ bound, meaning we conjecture (which should follow from iid increments) that the following

$$\mathbf{P}\Bigg(\sup_{0\le t\le T}\frac{\sup\limits_{0\le s\le t} W_s^f - W_t^f}{\sup\limits_{0\le s\le t} W_s^f} \le \beta\Bigg) > 1-\epsilon, \quad \forall T\in(0,\infty), \forall\epsilon\in(0,1)$$

is impossible to satisfy since it is (?) equivalent to

$$\lim_{T\to\infty}\mathbf{P}\Bigg(\sup_{0\le t\le T}\frac{\sup\limits_{0\le s\le t} W_s^f - W_t^f}{\sup\limits_{0\le s\le t} W_s^f} \le \beta\Bigg) = 0 > 1-\epsilon$$

Consider the space

$$(\beta, f, \mathbf{P}\bigg(\sup_{S\le t\le T}\big\{(1-\beta)\sup_{S\le s\le t} W_s^f - W_t^f\big\} > 0\bigg))$$

and for a given ϵ, β we will get a set of admissible f (i.e. a constrained region) to which we maximize $g_R(f)$ over the admissible set.

Another weaker (but maybe easier to solve) formulation (as suggested by Stephan Strum) is:

$$\iff \mathbf{P}\bigg((1-\beta)W_{t-1}^f > \inf_{t-1\le s\le t} W_s^f\bigg) \le \epsilon, \quad \forall t\in(0,\infty)$$

I think one can view this formalization as a constraint on draw-downs from my initial wealth (W_{t-1}^f) for the period $[t-1, t]$, for any possible period, whether that be a day, month, or year. Also note the second formulation is weaker than the first as the first is largest drawdown from largest point in the period and the second is largest drawdown from initial period's wealth.

Namely,

$$W_{t-1}^f \le \sup_{t-1\le r\le s} W_r^f \le \sup_{t-1\le s\le t}\sup_{t-1\le r\le s} W_r^f$$

which implies

$$\{(1-\beta)W_{t-1}^f > \inf_{t-1\le s\le t} W_s^f\} \subset \{(1-\beta)\sup_{t-1\le s\le t}\sup_{t-1\le r\le s} W_r^f > \inf_{t-1\le s\le t} W_s^f\}$$

.

Hence the first formulation is stronger than the second formulation,

$$\mathbf{P}\bigg((1-\beta)W_{t-1}^f > \inf_{t-1\le s\le t} W_s^f\bigg) \le \mathbf{P}\bigg((1-\beta)\sup_{t-1\le s\le t}\sup_{t-1\le r\le s} W_r^f > \inf_{t-1\le s\le t} W_s^f\bigg) \le \epsilon$$

Remark. *I need to justify these two formulations with a practical application story. Also, it seems plausible that given these constraints, f will no longer be a fixed fraction. We should, from the outset, consider a predictable process $f = \{f_t\}$ with $f_t \in (0, 1), \forall t$. To do this we need to consider the discrete time Kelly problem for a predictable fraction. See Lieb's counter-example in (Thorp, 2012). Also consider Bonferroni's inequality to interpret these as confidence interval curves and see the derivation at the back of (Thorp, 2006).*

(Use align command for better appearance)

7.3 Fluctuation and Ruin Theory

See (Feller, 1968), (Feller, 1971), (Asmussen & Albrecher, 2010), (Kyprianou, 2014). Also see Feller's 1948 paper called Fluctuation Theory of Recurrent Events (especially sections 5, 8, and 9).

Part III

Mathematical Formulation of the Generalized Kelly Criterion

Chapter 8

Generalized Kelly Criterion in Discrete Time

This is trying to justify the practical motivation for considering a predictable sequence of fractions $f = \{f_{k-1}\}$, where $f_{k-1} = f(r_1, \cdots, r_{k-1})$ or more generally $f = \{f_{n,k-1}\}$, where $f_{n,k-1} = f(r_{n,1}, \cdots, r_{n,k-1})$. So the betting fraction will be a deterministic function of the previous returns/payoffs/profits. For examples think blackjack, market cycles, or non-stationary gambles/returns. The non-stationary returns/bets could happen because the underlying distribution of our strategy or what we are choosing to invest in is changing, or we could be changing what we're betting in completely to something with a different return profile. It also attempts to answer question six below, posed in (Thorp, 2012).

Assume we're sequentially placing bets/investments/portfolio re-adjustments on a sequence of independent but NOT identically distributed payoffs/returns. Let $\{W_1^{n,f}\}_{n\geq 1}$ be the wealth dynamics achieved from betting the sequence of predictable fractions of $\{f_{n,k-1}\}_{k=1}^{n}$ in the sequence of gambles/returns $\{r_{n,k}\}_{n\geq 1}$, with $r_{n,k} \geq -1$, and compounding my bankroll/wealth. In other words, the wealth dynamics are modeling the process of compounding returns with a varying exposure of (proportion) $f_{n,k-1} \in (0,1)$ to those returns $r_{n,k}$, equivalently, we are betting $f_{n,k-1}W_1^{k-1,f}$ on the return $r_{n,k}$. Note that n is fixed and we will later take the limit as $n \to \infty$. Also, we are writing $f_{n,k-1}$ instead of just f_{k-1} for mathematical purposes to possibly later use the tools of sums of triangular arrays.

Remark. *Note that we interpret "discrete time" the same as discretely compounding over the unit interval $[0,1]$, which explains the notation $W_1^{n,f}$. This allows for an easy generalization to continuous time. The important part is that time is regarded as a point where we place a bet. So when we let $n \to \infty$, it can be interpreted as time (number of bets) increasing, or equivalently the frequency of bets per unit interval increasing.*

(Need to fix notation and clarify ideas here, along with properly set up the filtrations $\{\mathcal{F}_{n,k}\}$)

8.1 Discrete Time Wealth Dynamics

$\mathrm{W}_0 = w =$ initial wealth/bankroll/portfolio size

$W_1^{1,f} = W_0 + r_1(f_0 W_0) = W_0(1 + f_0 r_1)$

$W_1^{2,f} = W_1^{1,f} + r_2(f_1 W_1^{1,f}) = W_1^{1,f}(1 + f_1 r_2) = W_0(1 + f_0 r_1)(1 + f_1 r_2)$

$\vdots$

$W_1^{n,f} = W_1^{n-1,f} + r_n(f_{n-1} W_1^{n-1,f}) = W_1^{n-1,f}(1 + f_{n-1} r_n) = W_0 \prod_{k=1}^{n}(1 + f_{k-1} r_k)$

8.2 Kelly Criterion

$$\frac{W_1^{n,f}}{W_0} = \mathrm{e}^{n \ln(\frac{W_1^{n,f}}{W_0})^{\frac{1}{n}}} = \mathrm{e}^{n \frac{1}{n} \sum_{k=1}^{n} \ln(1+f_{k-1} r_k)} \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{\infty ??}$$

I think we need some sort or ergodic theorem here, possible ideas are in (Algoet & Cover, 1988)?

We need some sort of conditional law of large numbers.

I think the idea shown in (Algoet & Cover, 1988) and generalized in (Algoet, 1994), will help to establish a "Kelly-type criterion." What I mean by a "Kelly-type criterion" is a method of choosing a fraction at each stage that leads to doing the best in the long run, i.e. that leads to an optimal asymptotic growth rate. The claim is that optimizing the conditional expectation at each time we compound (or place a bet) is the strategy that turns out to do the best in the long run. So previously, we would find the long run growth rate for any possible fraction, and the find the optimal fraction. Now, we find the optimal fraction when we place a bet (today, as opposed to in the long run) and show that it is in fact optimal in the long run. This method would work also in the i.i.d. case, but is not how we previously found the Kelly criterion. More specifically, $f_{k-1}^* = argmax_{f_{k-1}} \mathbf{E}[\frac{1}{k} \ln(\frac{W_1^{k,f}}{W_0})|\mathcal{F}_{k-1}] = f^*(r_1, \cdots, r_{k-1})$ and show this betting sequence maximizes $\lim_{k\to\infty} \frac{1}{k} \ln(\frac{W_1^{k,f}}{W_0})$. This isn't precisely correct, but I believe it is the first step stragegy to making it more precise.

Chapter 9

Continuous Time Wealth Dynamics under Additive (non-stationary increment) Processes

Need to replace the computations in chapter with a predictable "fraction/proportion" process and I think the computations should follow through similarly and the main change is anytime you see an f times any process, e.g. fR_t, it would be replaced by an integral, e.g. $\displaystyle\int_0^t f_{s-}dR_s$

Chapter 10

Kelly Criterion in Continuous Time Under Additive Processes

As stochastic integrals are also Levy Process, I think there should be a long run limit but have yet to work out the details.

Chapter 11

Other Questions Vaguely Posed by Thorp Worth Investigating

Some additional research questions (to be investigated after the loose ends of the above results are formally proved) posed by Thorp in (Thorp, 2012) are in regards to an investment scenario faced by fund manager Mohnish Pabrai (discussed in The Dhandho Investor: The Low-Risk Value Method to High Return). Thorp points out why, in this particular situation, he wouldn't necessarily use the Kelly Criterion without particular adaptations. His discussion of the possible problems with the Kelly motivate many interesting and practical questions that should be resolved. They are the following:

11.1 Opportunity Costs

This is a problem of a gambler placing simultaneous bets at a given instant of time, such as holding a portfolio with assets that have some dependence structure between them. We could attempt to solve this issue by finding, either explicitly or numerically, the optimal $\vec{f}^*$ when we have a (log) return vector $\vec{R}_t = (R_t^1, \cdots, R_t^d)^\top \in \mathbf{R}^d$ that has a particular dependence structure. I think the mathematics of heavy tailed copulas could help answer this question. See (Cont & Tankov, 2004), chapter 5. Also, for intuition on how this can be used and why it would be important, see section 10.2 of (Ethier, 2010). At the end of this section, Ethier has an example where he has two betting scenarios, one which is more diversified than the other, with an edge, but you would choose the one with lower edge per trial since the lower edge per trial strategy leads to more diversification and hence a higher long term growth rate with lower volatility.

The opportunity cost refers to over-diversifying in the presence of large edges as this can lead to missed opportunities for compounding. Considering available alternative investments at any given time will help avoid this problem, see chapter 5 of (Brown, 2011).

11.2 Risk (volatility and ruin) Tolerance

As discussed in (Thorp, 2006), and in many articles in the main reference book, betting your full optimal fraction f^* can lead to large swings in your wealth in the short run, i.e. "there are drawdowns which are too large for the comfort of many investors" in the short run. So a heuristic that Thorp implemented was to use a percentage of the optimal fraction, called "fractional Kelly," which tamed the wild drawdowns while also reducing the long term growth rate.

I believe this can be added as a constraint to the long term grown rate optimization problem. We consider (or something like this) $\max_f g_R(f)$ subject to, either minimizing or ensuring that it is small enough, the constraint $\mathbf{P}(\frac{W_t^f - \max_{s\leq t} W_s^f}{\max_{s\leq t} W_s^f} > \beta | W_0 = w) =: \Psi(t, \beta, w, f)$. Where we're formalizing "drawdown" as $\frac{W_t^f - \max_{s\leq t} W_s^f}{\max_{s\leq t} W_s^f} \in [-1, 0]$.

Additionally, we are interested in the most important type of risk which is risk of ruin for a given initial wealth and betting fraction, $\mathbf{P}(\inf_{t\geq 0} W_t^f \leq 0 | W_0 = w) =: \Psi(w, f)$.

It appears that these formalizations can be handled with the tools of fluctuation theory for Levy processes, see chapter 11 especially in (Asmussen & Albrecher, 2010). In addition to answering questions about ruin and drawdown probabilities, it appears this theory can answer questions (involving asymptotic and non-asymptotic superior properties) like the chance your wealth will be larger than any goal on or before a particular time, $\mathbf{P}(\max_{s\leq t} W_s^f > B | W_0 = w)$. Moreover, chapter 8 (Level Dependent Risk Processes) might be able to answer questions about transaction costs and taxes. We need to include these into our wealth model as they really affect compounding, again see $(Thorp, 2006)$.

(Comment: We claim that one of the worst things that can happen when trying to compound gains is having "large" draw-downs. As a simple example, it is worse to have a ten percent drop in my wealth today than to have 10 one percent drops in my wealth spread out over time, say daily or monthly, due to the non-linearity of the exponential. Note we need to formally demonstrate this claim.) Thus the moral is that steep (implied by the slope) draw-downs destroy compounding!

11.3 Model Uncertainty

Another reason for using a "fractional Kelly" strategy (besides minimizing short term drawdown risk) is when we are unsure about our return/payoff distribution. If we get the distribution wrong then our optimal fraction could still lead to over-betting.

See (Browne & Whitt, 1996) and papers by S. Peng on model uncertainty.

11.4 Black Swans

This is what the above work is trying to solve. It is solving the Kelly Criterion under more general, including fat-tailed, Levy processes.

11.5 The "Long Run"

Here one would be concerned with convergence rates of the Strong Law of Large Numbers and its variants. In other words, how long do I need to keep compounding my wealth (or about how many bets do I need to place/frequency of bets) to achieve my long term growth rate. Note that not every gambler can make enough bets for the long run to kick in, at least with a high probability. These asymptotic properties will not hold true if the gambler/investor doesn't have enough opportunities to make it into the long run. As of now, I am not too sure how to go about finding these convergence rates, yet it would be a very important question to have answered.

11.6 Independent Increment (Additive) Processes

It is of interest to investigate processes that have independent increments and non-stationary increments, see (Thorp, 2006). Note these could help an investor formalize economic cycles or environments and change their incremental distributions in a deterministic way. It also makes sense in the case of card counting, where your payoff distribution is changing and is being measured by the count. (See Bill Gross's quote about knowing where you are in the cycle is like knowing the count). Additionally, in this framework we wouldn't have a reason for our optimal fraction to be a constant as every increment has a different environment, so we would be solving for an optimal fraction process $\{f_t^*\}$ and we leave the setting of constant proportion/myopic strategies. And in this framework, it may be possible to find a strategy that dominates the Kelly strategy and is "essentially different," see "Leib's Paradox" in (Thorp, 2012). This should be mathematically formalized using additive processes.

Of course the practical question of estimating the characteristics of a Levy process from data and numerically computing all of these different quantities is also a question needing to be answered.

11.7 Appendix: Results on Levy Processes and Convergence of Discrete Models

1. **Defn: Levy Process**
2. **Theorem: Levy-Ito Decomposition**

Bibliography

Algoet, P. H. (1994). The strong law of large numbers for sequential decisions under uncertainty. *IEEE Trans. Inform. Theory*, *40*(3), 609–633. doi:10.1109/18.335876

Algoet, P. H. & Cover, T. M. (1988). Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. *Ann. Probab. 16*(2), 876–898. Retrieved from http://links.jstor.org/sici?sici=0091-1798(198804)16:2%3C876:AOAAEP%3E2.0.CO;2-D&origin=MSN

Asmussen, S. & Albrecher, H. (2010). *Ruin probabilities* (Second). Advanced Series on Statistical Science & Applied Probability. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ. doi:10.1142/9789814282536

Barndorff-Nielsen, O. E. & Shiryaev, A. (2010). *Change of time and change of measure*. Advanced Series on Statistical Science & Applied Probability. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ. doi:10.1142/7928

Brown, A. (2011). *Red-blooded risk: The secret history of wall street*. Wiley.

Browne, S. & Whitt, W. (1996). Portfolio choice and the Bayesian Kelly criterion. *Adv. in Appl. Probab. 28*(4), 1145–1176. doi:10.2307/1428168

Cont, R. & Tankov, P. (2004). *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL.

Eberlein, E. (2009). Jump type levy processes. In *Handbook of financial time series* (pp. 439–456). Springer.

Ethier, S. N. (2010). *The doctrine of chances*. Probability and its Applications (New York). Probabilistic aspects of gambling. Springer-Verlag, Berlin. doi:10.1007/978-3-540-78783-9

Feller, W. (1968). *An introduction to probability theory and its applications. vol. i*. Third edition. John Wiley & Sons, Inc., New York-London-Sydney.

Feller, W. (1971). *An introduction to probability theory and its applications. vol. ii*. Second edition. John Wiley & Sons, Inc., New York-London-Sydney.

Jacod, J. & Shiryaev, A. N. (2003). *Limit theorems for stochastic processes* (Second). Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin. doi:10.1007/978-3-662-05265-5

Kardaras, C. & Platen, E. (2013). Multiplicative approximation of wealth processes involving no-short-sales strategies via simple trading. *Math. Finance*, *23*(3), 579–590. doi:10.1111/j.1467-9965.2011.00511.x

Kelly, J. (1956). A new interpretation of information rate. In *The kelly capital growth investment criterion* (Chap. 2, pp. 25–34). World Scientific.

Kyprianou, A. E. (2014). *Fluctuations of Lévy processes with applications* (Second). Universitext. Introductory lectures. Springer, Heidelberg. doi:10.1007/978-3-642-37632-0

Poundstone, W. (2006). *Fortune's formula: The untold story of the scientific betting system that beat the casinos and wall street.* Hill and Wang.

Rachev, S. T. & Rushendorf, L. (1994). Models for option prices. *Teor. Veroyatnost. i Primenen. 39*(1), 150–190. doi:10.1137/1139005

Thorp, E. O. (1966). *Beat the dealer: A winning strategy for the game of twenty-one.* Vintage.

Thorp, E. O. (2006). The kelly criterion in blackjack, sports betting, and the stock market. In *The kelly capital growth investment criterion* (Chap. 54, pp. 791–834). World Scientific.

Thorp, E. O. (2012). Understanding the kelly criterion. In *The kelly capital growth investment criterion* (Chap. 36, pp. 511–525). World Scientific.

Thorp, E. O. (2017). *A man for all markets: From las vegas to wall street, how i beat the dealer and the market.* Random House.

Thorp, E. O., MacLean, L. C., & Ziemba, W. T. (2011). *The kelly capital growth investment criterion.* World Scientific.

Wilson, A. N. (1970). *The casino gambler's guide.* Harper Collins.