Chapter 4

Stock Market Investment, as Gambling on a Favorable Game

Lecture 3 mentioned favorable and unfavorable games before focussing on fair games. This lecture will soon focus on the question: how do you best exploit a setting where the odds are in your favor? This scenario is unrealistic for casino games but has historically been realistic for the stock market.

Novice investors are told to view the stock market as a place for long-term investment. This excellent advice is unfortunately rather neglected in most mathematically oriented discussions, but will be emphasized here. Between short-term speculation and long-term investment lies a spectrum of intermediate activities with no clear dividing line, but the ends of the spectrum are quite different to a typical individual.

The stock market is really “a market of stocks”, but for most of this lecture I use a conventional shorthand of representing the U.S. market by the S&P500 index (essentially an actual investment possibility, via an index fund).

4.1 Unfavorable games

There are two things to say about unfavorable games. Mathematicians often say the first thing wrong, for instance by saying

Gambling against the house at a casino is foolish, because the
odds are against you and in the long run you will lose money.

What’s wrong is the *because*. Saying

Spending a day at Disneyland is foolish, because you will leave with less money than you started with

is inane, because people go to Disneyland for entertainment, and know they have to pay for entertainment. And the first quote is equally inane. Casino gamblers may have irrational ideas about chance and luck, but in the U.S. they typically regard it as entertainment with a chance of winning, not as a plan to make money. So it’s worth being more careful and saying

Gambling against the house at a casino and expecting to make money is foolish, because the odds are against you and in the long run you will lose money.

The second thing to say is that buying insurance is mathematically similar to placing an unfavorable bet – your expected gain in negative, because the insurance company needs to cover its costs and make a profit. But buying insurance is often sensible on grounds of *utility theory* (yyy ref discussion in other chapter), which we shall neglect in this lecture.

### 4.2 Favorable games and the long term

Notions of *long term* versus *short term* play an important role in investment, so let’s start with a brief discussion. In everyday language, a job which will only last six months is a short term job; someone who has worked for a company for fifteen years is a long term employee. Joining a softball team for a summer is a short term commitment; raising children is a long term commitment. We judge these matters relative to human lifetime; *long term* means some noticeable fraction of a lifetime.

Turning to money matters, consider the difference between simple interest and compound interest. The Table compares the value, after increasing numbers of years, of an initial $1,000 earning 7% interest.

<table>
<thead>
<tr>
<th>year</th>
<th>0</th>
<th>4</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>simple interest</td>
<td>1000</td>
<td>1,280</td>
<td>1,560</td>
<td>1,840</td>
<td>2,120</td>
<td>2,400</td>
</tr>
<tr>
<td>compound interest</td>
<td>1000</td>
<td>1,311</td>
<td>1,718</td>
<td>2,252</td>
<td>2,952</td>
<td>3,870</td>
</tr>
</tbody>
</table>

*Table 1.* Effect of 7% interest, compounded annually.
4.3. THE IID MODEL AND THE KELLY CRITERION

One of several possible notions of long term in financial matters is “the time span over which compounding has a noticeable effect”. Rather arbitrarily interpreting “noticeable effect” as “10% more” and taking the 7% interest rate, this suggests taking 8 years as the cut-off for long term. Being about 10% of a human lifetime, this fortuitously matches reasonably well the “noticeable fraction of a lifetime” criterion above. And indeed in matters pertaining to individuals, financial or otherwise, most writers use a cut-off between 5 and 10 years for “long term”.

Aside: the one fact from freshman calculus of substantial relevance to your personal life is the inequality

$$1 + \rho(e^{rt} - 1) > e^{prt}.$$  

This shows the value of unit investment, with interest rate $r$ and tax rate $1 - \rho$, is greater when tax is deferred until the sale time $t$ than if tax is paid as the interest is earned.

The theme of this section is the nature of compounding when gains and losses are unpredictable. The relevant arithmetic is multiplication not addition: a 20% gain followed by a 20% loss combine to a 4% loss, because $1.2 \times 0.8 = 0.96$.

In class I introduce the topic as follows. Suppose you invest $1,000 today in the stock market, more precisely in the S&P500 index (via a fund with very low expenses). What do you guess the investment will be worth in 10 years? In 2011 the student guesses\(^1\) were

$500, \$800, \$1,200, \$1,800, \$1,800, \$2,400.$

I have a deck of cards on which are pasted the annual total returns of the S&P500 index over each of the 52 years 1956 through 2007. I say “let’s suppose the annual returns over the next ten years are statistically like random years from the past; we can track our hypothetical investment value over the next ten years by shuffling and dealing ten cards”. Doing this once in the 2008 class, the hypothetical investment grew from $1,000 to $1,839.

4.3 The IID model and the Kelly criterion

Let us make explicit the type of model used implicitly above. A “return” $x = 0.2$ or $x = -0.2$ in a year means a 20% gain or a 20% loss.

\(^1\)on September 6, 2011; the index closed at 1,165. In 2008 (on September 17, 2008; the index closed at 1,156) the guesses were $880, \$1,000, \$1,600, \$1,700, \$1,800, \$2,200, \$2,500.$
The IID model. Write $X_i$ for the return in year $i$. Suppose the $(X_i)$ are IID random variables. Then the value $Y_n$ of your investment at the end of year $n$ is

$$Y_n = Y_0 \prod_{i=1}^{n} (1 + X_i)$$

(4.1)

where $Y_0$ is your initial investment.

To analyze this model we take logs and divide by $n$:

$$n^{-1} \log Y_n = n^{-1} \log Y_0 + n^{-1} \sum_{i=1}^{n} \log(1 + X_i)$$

and the law of large numbers says that as $n \to \infty$ the right side converges to $E \log(1 + X)$. We want to compare this to an investment with a non-random return of $r$. For such an investment (interest rate $r$, compounded annually) we would have $Y_n = Y_0 (1 + r)^n$ and therefore $n^{-1} \log Y_n \to \log(1 + r)$. Matching the two cases gives the conclusion

In the IID model, the long term growth rate is

$$\exp(E \log(1 + X)) - 1.$$  

(4.2)

The formula looks strange, because to compare with the IID annual model we are working with the equivalent “compounded annually” interest rate. It is mathematically nicer to use instead the “compounded instantaneously” interest rate, which becomes just $E \log(1 + X)$.

The main impact of this result is that what matters about the random return $X$ is not precisely the mean $E X$, but rather its “multiplicative” analog $E \log(1 + X)$. Let us note, but set aside for a while, the points

- Is the model realistic for stock market investing?
- The phrase long term here refers to the applicability of the law of averages as an approximation to finite time behavior – this is a third meaning of the phrase, distinct from the two previous meanings.

We can jump to the main mathematical point

The Kelly Criterion. Given a range of possible investment portfolios, that is a range of ways to allocate money to different risky or safe assets, where way $\alpha$ will produce return $X_\alpha$, choose the way $\alpha$ that maximizes the long term growth rate (4.2), that is the way that maximizes $E \log(1 + X_\alpha)$.

Let us show some of the mathematics that can be done using the criterion.
4.4 Mathematics of the Kelly criterion: one risky and one safe asset

Formula (4.2) applies if we invest all our money in “a stock” (meaning an index fund, and assuming the multiplicative model for stock market returns). But suppose there’s a risk-free alternative investment (a “bond”, in the usual terminology) that pays a fixed interest rate \( r \). Suppose we choose some number \( 0 \leq p \leq 1 \) and at the start of each year we invest a proportion \( p \) of our total “investment portfolio” in the stock market, and the remaining proportion \( 1 - p \) in the bond. In this case formula (4.1) becomes

\[
Y_n = Y_0 \prod_{i=1}^{n} (1 + X_i^*)
\]

where \( X_i^* = pX_i + (1-p)r \). The long term growth rate is now a function of \( p \):

\[
growth(p) = \exp(E \log(1 + pX + (1-p)r)) - 1. \tag{4.3}
\]

The Kelly criterion says: choose \( p \) to maximize \( growth(p) \). Let’s see two examples. In the first \( X \) is large, and we end up with \( p \) small; in the second \( X \) will be small, and we end up with large \( p \). In these two examples we take the time unit to be 1 day instead of 1 year (which doesn’t affect math formulas).

**Example: pure gambling.** Imagine a hypothetical bet which is slightly favorable. Suppose each day we can place a bet of any size \( s \); we will either gain \( s \) (with probability \( 0.5 + \delta \)) or lose \( s \) (with probability \( 0.5 - \delta \)), independently for different days (here \( \delta \) is assumed small). Take \( r = 0 \) for the moment. What proportion \( p \) of our portfolio do we want to bet each day? Here, for small \( \delta \),

\[
E \log(1 + pX) = (\frac{1}{2} + \delta) \log(1 + p) + (\frac{1}{2} - \delta) \log(1 - p)
\]

\[
\approx (\frac{1}{2} + \delta)(p - p^2/2) + (\frac{1}{2} - \delta)(-p - p^2/2)
\]

\[
= 2\delta p - p^2/2.
\]

Thus the asymptotic growth rate is approximately the quadratic function of \( p \)

\[
G(p) = 2\delta p - p^2/2 \tag{4.4}
\]

shown in the Figure. The Kelly criterion says to choose \( p \approx 2\delta \) and then your long term growth rate will be \( \approx 2\delta^2 \).
Now recall that we simplified by taking $r = 0$; when $r > 0$, the fact that a proportion $1 - p \approx 1$ of the portfolio not put at risk each day can earn interest, brings up the optimal growth rate to $r + 2\delta^2$; the quantity $2\delta^2$ represents the extra growth one can get by exploiting the favorable gambling opportunity.

To give a more concrete mental picture, suppose $\delta = 1\%$. The model matches either of the two following hypothetical scenarios.

(a) To attract customers, a casino offers (once a day) an opportunity to make a roulette-type bet with a $51\%$ chance of winning.

(b) You have done a statistical analysis of day-to-day correlations in some corner of the stock market and have convinced yourself that a certain strategy (buying a portfolio at the start of a day, and selling it at the end) replicates the kind of favorable bet in (a).

In either scenario, the quantity $2\delta^2 = 2/10,000$ is the “extra” long term growth rate available by taking advantage of the risky opportunity. Note this growth rate is much smaller than $2\%$ “expected gain” on one bet. On the other hand we are working “per day”, and in the stock market case there are about 250 days in a year, so the growth rate becomes about $5\%$ per year; recalling this is “5\% above the risk-free interest rate”, it seems a rewarding outcome. But if $\delta$ were instead $0.5\%$ then the extra growth rate becomes $1\frac{1}{2}\%$, and (taking into account transaction costs and our work) the strategy hardly seems worth the effort.

Implicit in the Figure is a fact that at first strikes everyone as counterintuitive. The curve goes negative when $p$ increases above approximately $4\delta$. So even though it is a favorable game, if you are too greedy then you
will lose in the long run!

**Example: small** $X$  I first give a nicer algebraic way of dealing with the interest rate $r$. Set $$X = r + (1 + r)X^*$$ and interpret $X^* = (X - r)/(1 + r)$ as “return relative to interest rate”. Then a couple of lines of algebra let us rewrite (4.3) as
growth($p$) = $(1 + r)\exp(E\log(1 + pX^*)) - 1$ \hspace{1cm} (4.5)
and the optimization problem now doesn’t involve any $r$. If we imagine the stock market on a daily time-scale and suppose changes $X^*$ are small, with mean $\mu$ and variance $\sigma^2$, then we can use the series approximation
$$\log(1 + pX^*) \approx pX^* - \frac{1}{2}(pX^*)^2$$
to calculate $$E\log(1 + pX^*) \approx p\mu - \frac{1}{2}p^2(\mu^2 + \sigma^2) \approx p\mu - \frac{1}{2}p^2\sigma^2$$ (the latter because $\mu$ and $\sigma^2$ are in practice of the same order, so $\mu^2$ is of smaller order than $\sigma^2$). So the Kelly criterion says: choose $p$ to maximize $p\mu - \frac{1}{2}p^2\sigma^2$, that is choose
$$p = \mu/\sigma^2.$$ This is another remarkable formula, and let us discuss some of its mathematical implications.

1. The formula is (as it should be) time-scale free. That is, writing $\mu_{\text{day}}, \mu_{\text{year}}, \sigma^2_{\text{day}}, \sigma^2_{\text{year}}$ for the means and variances over a day and a $N$-day year, then (because compounding has negligible effect over a year) $\mu_{\text{year}} \approx N\mu_{\text{day}}$ and $\sigma^2_{\text{year}} \approx N\sigma^2_{\text{day}}$, so we get the same value for $\mu/\sigma^2$ whether we work in days or years.

2. Even though we introduced the setup by stating that $0 \leq p \leq 1$, the model and its analysis make sense outside that range. Economic theory and experience both say that the case $\mu < 0$ doesn’t happen (investors are risk averse and so would buy no stock; this would cause the current price of stock to drop), but if it did then the formula $p = \mu/\sigma^2 < 0$ says that no only should be invest 100% of our wealth in the bond, but also we should “sell short” (i.e. borrow) stock and invest the proceeds in the bond.

3. More interesting is the case $p > 1$. Typical values given for the S&P500 index (as noted later, stating meaningful historical values is much
harder than one might think) are $\mu = 5.6\%$ and $\sigma = 20\%$, in which case the Kelly criterion says to invest a proportion $p = 140\%$ of your wealth in the stock market, i.e. to borrow money (at fixed interest rate $r$) and invest your own and the borrowed money in the stock market.

### 4.5 What about the not-so-long term?

We started with the multiplicative model, which assumes that returns in different time periods are IID. This is not too realistic, but the general idea behind the Kelly criterion works without any such assumption, as we now explain.

Going back to basics, the idea

> to invest successfully in the stock market, you need to know whether the market is going to go up or go down

is just wrong. Theory says you just need to know the probability distribution of a future return. So suppose (a very big **SUPPOSE**, in practice!) at the beginning of each year you could correctly assess the probability distribution of the stock return over the coming year, then you can use the Kelly criterion (4.3) to make your asset allocation. The fact that the distribution, and hence your asset allocation, would be different in different years doesn’t make any difference – this strategy is still optimal for long-term growth.

The numbers for growth rates that come out of the formula of course depend on the distributions of each next year’s returns, but there’s one aspect which is “universal”. In any situation where there are sensible risky investments, following the Kelly strategy means that you accept a short-term risk which is always of the same format:

> 40% chance that at some time your wealth will drop to only 40% of what it you started with.

The magical feature of this formula is that the percents always match: so there is a 10% chance that at some time your wealth will drop to only 10% of what you started with.

For an individual investor, it is perfectly OK to be uncomfortable with this level of medium-term risk and to be less **aggressive** (in investment jargon) by using a partial Kelly strategy, that is using some smaller value of

$$p = \text{proportion of your assets invested in stocks}$$
than given by the Kelly criterion. Theory predicts you will thereby get slower long-term growth but with less short-term volatility.

How one might expect this theory (based on assuming known true probability distributions for the future, and on seeking to optimize long-term growth rate) to relate to the actual stock market is not obvious, but one can certainly look at what the actual percentages have been.

![Figure 2](image)

**Figure 2.** Historical distribution\(^2\) of the minimum future value of a 100 investment.

This is obviously very different from the “Kelly” prediction of a flat histogram over \([0, 100]\). This data is not adjusted for inflation or for comparison with a risk-free investment, and such adjustments (a possible small project) would make the histogram flatter, but still not close to the Kelly prediction. We mentioned before that over the historical long term it has been more profitable to borrow to invest more than 100% of your assets in the market. Both observations reflect the fact that the stock market fluctuates *less* than would the fortune of a Kelly-optimizing speculator.

### 4.6 Can one execute this theory?

Relating the mathematical theory to actual stock market investing raises many issues.

---

\(^2\)based on hypothetical purchase of S&P500 index on first day of each year 1950-2009 and on subsequent monthly closing data.
CHAPTER 4. STOCK MARKET INVESTMENT, AS GAMBLING ON A FAVORABLE (...

Understanding the past. The first is an issue you might not have expected:

it is very hard to pin down a credible and useful number for the historical long-term average growth rate of stock market investments.

Over the 60 years 1950-2009 the S&P500 index rose at (geometric) average rate\(^3\) 7.2%. Aside from the (rather minor) point that we are using a particular index to represent the market what could possibly be wrong with using this figure? Well,

- it ignores dividends
- it ignores expenses
- it is sensitive to choice of start and end dates; starting in 1960 would make the figure noticeably lower, whereas ending in 1999 would make it noticeably higher.
- to interpret the figure we need to compare it to some alternative investment, by convention some “risk-free” investment.
- it ignores inflation
- it ignores taxes.

So one can get very different numbers, depending on which of these factors is taken into account. For instance, taking two of these factors into account, the same site gives, for the same period 1950-2009, annual averages

\[
\text{inflation-adjusted total return} \\
= \text{price change (7.2\%) + dividends (3.6\%) - inflation (3.8\%)} \\
= 7.0\%.
\]

The past as a guide to the future  

Forests have been felled for paper for philosophers to discuss the problem of induction (\(W\)). In the context of investment, my take is

As a default, assume the future will be statistically similar to the past. Not because this is true in any Platonic sense, but because anyone who says different is trying to sell you something.

\(^3\)Data here and below from www.simplestockinvesting.com/SP500-historical-real-total-returns.htm
4.7. GEOMETRIC BROWNIAN MOTION AND THE BLACK-SCHOLEs FORMULA

As often remarked, the four most dangerous words in finance are *this time it’s different*.

Asking how long of a past time period to use for statistical purposes as a guide for the future is a question with no right answer. Asking how far into the future one should care about does have an answer – until you’re age 80 or so (that is, 60 years ahead for my students).

**Psychology in executing Kelly**  So if you do set up a historically-based model for the stock market (and at least one alternative investment) and assume the future will be statistically similar to the past, then the Kelly criterion tells you how to divide your money between these investments for maximum long term growth, assuming the model were correct. But some practical issues still remain. What is your personal trade-off between long term reward and short-term risk? How should this change with age (typically one is advised to become more risk-averse as one ages). (In principle one could introspect your utility function and then calculate your optimal trade-off, but I suspect few people have ever done so.) In a different vein, long term strategies can only work if you avoid changing your mind partway through, so how does one plan to avoid changing one’s mind later?

**Bottom line**  Going through the procedures above by oneself will have little appeal to a typical individual, so what is the closest practical option? There are many internet sites that provide combinations of historical data, hypothetical data and (not so mathematical) theory relevant to this topic (presented in the more conventional mean-variance format of section 4.8). In class I exhibit material from the IFA (Index Funds Advisors: www.ifa.com) site which has a wide range of interesting graphical data displays.\(^4\) IFA and other sites start out by trying to assess the individual’s subjective risk tolerance using a questionnaire. The site then suggests one of a range of 21 portfolios, roughly 0% to 100% Kelly in 5% increments, based on historic data.

4.7 Geometric Brownian motion and the Black-Scholes formula

Most introductory accounts of mathematics related to the stock market do not focus on the Kelly criterion, but instead on the (related) topics of this

\(^4\)Of course I am not endorsing this particular outfit.
and the next section. The IID model involved choosing a time unit – we chose one year – and the resulting strategy involves rebalancing your portfolio at the end of each year but not more frequently. But the choice of one year is arbitrary – why not a month or a week or a day or an hour? It is perhaps more natural to seek to model as a random process what you see in a stock chart, the fluctuation of prices with time, regarding time as a continuous variable. It turns out, as a remarkable mathematical fact, that the only possible model one can use as the continuous analog of the IID model (and have prices move continuously, without substantial instantaneous jumps) is geometric Brownian motion (W). Instead of using data from the past directly (as in my “deck of cards” class demonstration) this theoretical setup allows and encourages you to assume that the probability distribution of a 1-year return has the log-Normal distribution (W), then to estimate the two parameters of that distribution, and then to use the associated geometric Brownian motion model for making predictions about the future.

This geometric Brownian motion model captured the attention of mathematically inclined researchers in the early 1970s as a setting where it is possible to write down formulas for aspects of the future price fluctuations (assuming the model is correct). The prototype is the Black-Scholes formula, which we now state, following (W). If a stock price is currently $S_0$ and its future prices $S_t$ are assumed to follow geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{4.7}$$

then the fair price of an option to buy at time $t$ in the future at price $K$ is

$$S_0 \Phi(d_1) - Ke^{-rt} \Phi(d_2) \tag{4.8}$$

where $r$ is the risk-free interest rate, $\Phi(\cdot)$ is the standard Normal distribution function,

$$d_1 = \frac{\log(S_0/K) + (r + \sigma^2/2)t}{\sigma \sqrt{t}}, \quad d_2 = \frac{\log(S_0/K) + (r - \sigma^2/2)t}{\sigma \sqrt{t}}.$$

This certainly belongs on our list (yyy cross-ref) of the top ten formulas in applied probability, but its interpretation is more subtle than the others. The fair price in question is by definition

$$e^{-rt} \mathbb{E} \max(S_t - K, 0).$$

The formula for this quantity, applied to geometric Brownian motion (4.7), depends on the drift $\mu$, but the Black-Scholes formula (4.8) does not. The
4.8 MEAN-VARIANCE ANALYSIS

key issue is that an extra “no arbitrage” assumption, loosely analogous to the “martingale” assumption for prediction market prices in Lecture 3, is used to establish (4.8), by imagining a portfolio consisting of time-varying amounts of the stock and the option but with a total value that does not vary with the fluctuations of the stock price.

4.8 Mean-variance analysis

We earlier discussed diversification in the very simple context of one safe and one risky asset. The general case of diversifying over many assets is usually presented as mean-variance analysis, or (more pretentiously) as modern portfolio theory (W). This should be viewed as a medium-term theory – it doesn’t explicitly involve compounding – a viewpoint perhaps influenced by traditional economics focus on medium term issues like the business cycle.

Here’s the conceptual setup. A portfolio has a reward, which is mean annual return, and a risk, which is the s.d. of annual return. Different investors have different risk tolerances; the goal of theory is to produce a spectrum of portfolios which provide maximum reward for each given level of risk.

The mathematics of mean-variance analysis involves the kind of matrix algebra that one learns as an undergraduate Statistics or Economics major. Rather than give the algebra I will illustrate with simple hypothetical numbers and then jump to the bottom line.

Suppose there were only two stocks, and historical data for annual returns showed

<table>
<thead>
<tr>
<th>stock</th>
<th>mean</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8%</td>
<td>20%</td>
</tr>
<tr>
<td>B</td>
<td>8%</td>
<td>20%</td>
</tr>
</tbody>
</table>

One’s first thought is that there’s no difference between investing in A or in B. But the point is that one can invest 50% of one’s portfolio in each; this preserves the mean of 8% but reduces the s.d., that is reduces the risk. If the stock price fluctuations were independent then the s.d. of the portfolio would be reduced to around 14%, In practice stock returns are typically positively correlated, so the reduction in s.d. is smaller, but still desirable.

For the next simplest hypothetical example, some the s.d.s are unequal

<table>
<thead>
<tr>
<th>stock</th>
<th>mean</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8%</td>
<td>15%  = σA</td>
</tr>
<tr>
<td>B</td>
<td>8%</td>
<td>20%  = σB</td>
</tr>
</tbody>
</table>
Again one might first think that one should invest entirely in A. But a mixture – proportion $p$ in A and proportion $1-p$ in B – has s.d. $\sigma$ given (in the independent case) by

$$\sigma^2 = p^2 \sigma_A^2 + (1-p)^2 \sigma_B^2$$

and this is minimized by taking $p = \sigma_B^2/\left(\sigma_A^2 + \sigma_B^2\right)$. With the numbers shown, take $p = 0.64$ to get $\sigma = 12\%$.

As a third hypothetical example consider

<table>
<thead>
<tr>
<th>stock</th>
<th>mean</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8%</td>
<td>15% = \sigma_A</td>
</tr>
<tr>
<td>B</td>
<td>6%</td>
<td>20% = \sigma_B</td>
</tr>
</tbody>
</table>

Here a 50-50 mixture has $\mu = 7\%$ and $\sigma = 12.5\%$, and one might prefer that trade-off to the stock A parameters.

(In class I continue, discussing the efficient frontier, following Capital asset pricing model (W).)

Analogous to the Kelly criterion one can identify one of these as the mean-variance optimal portfolio. By typing several stocks, e.g. apple exxon coca-cola, into WolframAlpha one can see this portfolio on these stocks plus S&P500, bonds and T-bills.

**Linking mean-variance analysis and the Kelly criterion**  Writing $X$ as before for return in one year, and writing $\log(1+X) \approx X - X^2/2$ so that $E\log(1+X) \approx \mu - (\mu^2 + \sigma^2)/2$, we see that the Kelly criterion corresponds (approximately) to choosing, over possible portfolio combinations, the one whose $(\mu, \sigma)$ value maximizes $\mu - (\mu^2 + \sigma^2)/2$. The approximation here is that we are ignoring the possibility of unusually large changes.

### 4.9 On investment advice

A search on “diet” in amazon.com books produces\(^5\) a claim of 63,932 results, though the listings actually stop at number 1201 (*The Raw Secrets: The Raw Vegan Diet in the Real World* by Patenaude). Similarly, a search on “investment” produces a claim of 78,686 results, while these listings also stop at number 1201 (evidently an Amazon policy). A naive visiting Martian might think the former reflected a vast diversity of human dietary requirements for some genetic, environmental or occupational reasons, and wonder what analogous factors might require such a wide variety of investment strategies.

---

\(^5\)May 2011
4.9. **ON INVESTMENT ADVICE**

The reader, as a skeptical human, may suspect such books exist merely because enough people are willing to buy them. Here are two of the many strategies.

**Fundamental analysis of an individual stock** seeks to assess the “intrinsic value” of the business by analyzing its likely future profits, and thereby find stocks which the current market price overvalues or undervalues.

**Market timing** seeks to analyze aggregate economic data (business cycle, inflation, unemployment, corporate profits) as well as sentiment (opinions about the near future) to decide when to switch between stocks/bonds/cash or between subsectors of the markets.

Almost all investment passes through some kind of “professional advisor” (including mutual fund managers, analysts etc). So as a simple matter of arithmetic, the average return to investors must equal the average return of the markets, minus expenses and fees paid to the advisors. I shall not go into details about the *efficient-market hypothesis* (EMH) (W). Treating EMH as an ideology or a law of physics strikes me as silly, but simply asserting that it is at least *very difficult* to consistently beat the market – just as it is very difficult to be one of the best people in the world at anything substantial – is much more plausible. Concretely, what the EMH predicts is that if you take any well-defined strategy that has been used by a group of advisors, then over the long term, the average return to investors must again equal the average return of the markets, minus expenses and fees. Numerous academic studies of this prediction gave been done, and generally confirm the prediction (reviewing such studies is a course project).

Even people who accept the simple logic and experience that most professional advisors can’t beat the market are frequently seduced by the notion that a few can. Here’s my take on this. Suppose someone comes up to me, takes a coin out of their pocket, says “I’m going to toss the coin 10 times and make it land heads every time”, and then does so. What’s my reaction? Well, there are three possibilities. They might just have been very lucky. They might be cheating (a two-headed coin, or a second concealed coin). Or they might be exhibiting a rare and difficult skill, the ability to toss a coin which in fact only rotates a specific few times in the air and lands predictably. Analogously, if you look at the 5 best-performing advisors over the last 5 or 10 or 15 years, then they might just have been lucky (*some* 5 people must be best), they might be exhibiting a rare and difficult skill of actually being able to consistently beat the market (as does Warren Buffett, in the opinion of many people) or they might be another Bernie Madoff. Whether or not a few advisors will in the future be able to consistently beat the market is an interesting intellectual question *but it doesn’t matter to you*
– you can’t reliably pick one of these few out of the pack, any better than you can pick one the few future-best-performing stocks out of the pack.

4.10 Wrap-up

I have only touched upon a corner of a large topic, but within this corner let me reiterate some key points.

1. Common sense says objects can be stationary or move slowly or move fast or move very fast, and that there should be no theoretical limit to speed – but physics says in fact you can’t go faster than the speed of light. And that’s a very non-obvious fact. Similarly, we know there are risk-free investments with low return; by taking a little risk (risk here equals short-term fluctuations) we can get higher long-term reward. Common sense says this risk-reward trade-off spectrum continues forever. But it doesn’t. As a math fact, you can’t get a higher long-term growth rate than you get from the "100% Kelly strategy". You’ve free to take more risk if you like excitement but you don’t benefit from it.

2. Mathematics (section 4.8) not only confirms that diversification is good but also shows it is even better than you might intuitively expect.

3. Any mathematics one can do, involving the stock market or wider aspects of finance and risk, either assumes (as we have) that the future will be statistically like the past, or makes an explicit assumption of some ways in which it will be different. Now of course the rules of the game do change with time, but the consensus view of such trends is already reflected in current prices. To profit one would need to adopt some minority view of the future, and be correct. Surely the majority of people who try to do so get it wrong.

4. Popular opinion often says that stock market fluctuations are larger than they “should be”, whereas mathematics says that the stock market has historically fluctuated less than would the fortune of a Kelly-optimizing speculator. In fact no-one knows how large the short-term fluctuations “should be”, under either a “rational” or a “psychological” theory of the market, and a testable theoretical prediction relating market fluctuations to measurable aspects of the real economy would surely win you a Nobel prize or enable you to make a fortune (by speculating on volatility).

5. Geometric Brownian motion is a mathematically natural, and reasonably accurate, model for the short term fluctuations of stocks. After Bachelier pointed this out in 1900, the model was mostly underused until the 1970s, but subsequently overused and treated as more accurate than it
6. Unlikely things are going to happen over your 50-year investment lifetime, and thinking in terms of Kelly rather than mean-variance encourages you to pay attention to small chances of dramatic losses.

7. To adapt Churchill’s comment on democracy,

The EMH is the worst way of thinking about the stock market, except for all those other ways that have been tried from time to time.

4.11 Further reading

Poundstone’s *Fortune’s Formula* is a non-technical book on the Kelly criterion. Much of the book is entertaining episodic anecdotal history of characters like Shannon, Kelly, Thorp, Milken, Boesky and Long Term Capital Management. It has an interesting account of the dispute between the proponents of Kelly (math types) and economists led by Samuelson who viewed it as too risky even in the long run. Memorable slogan, in place of my “speed of light” analogy,

100% Kelly strategy marks the boundary between aggressive and insane investing.

Malkiel’s classic *A Random Walk Down Wall Street* sets out in plain language the academic view that seeking to beat the market via fundamental analysis is a mug’s game. Taleb’s recent best-seller *The Black Swan* is hard to describe in a few words but one of its main points is that the geometric Brownian motion model, and formulas such as Black-Scholes based on it, ignore small chances of unforeseen events that might have substantial effect. Reiner and Rogoff’s *This Time Is Different: Eight Centuries of Financial Folly* is a quantitative academic study, pointing out (to quote one reviewer) “the highly repetitive nature of financial crises resulted from a dangerous mix of hubris, euphoria and amnesia”. Zingales discusses the EMH in the light of the 2008 financial crisis.

---

*See www.stat.berkeley.edu/~aldous/157/Books/taleb.html for my lengthy review.*

*faculty.chicagobooth.edu/luigi.zingales/papers/LearningtoLivewithmotsoEfficientMarkets.pdf*