Financial Modeling and Data Analysis

Equity Returns and the Random Walk Model

CHEUNG Ying Lun

Capital University of Economics and Business

Outline

Compounding and Log Returns

Effect of compounding

Suppose you are going to deposit \$10,000 in a bank, which offers you a 10% per annum interest rate and the following compounding scheme:

1. Compounding every year, where the one-year interest rate is 10%;
2. Compounding every 6 months, where the 6-month interest rate is $10\%/2 = 5\%$.

Which one should you choose?

Effect of compounding

Frequency	**No. of payments**	**Interest rate per period**	**Total value**
Annual	1	10%	$11000.00
Semiannual	2	5%	$11025.00
Quarterly	4	2.5%	$11038.13
Monthly	12	0.833%	$11047.13
Weekly	52	0.192%	$11050.65
Daily	365	0.027%	$11051.56

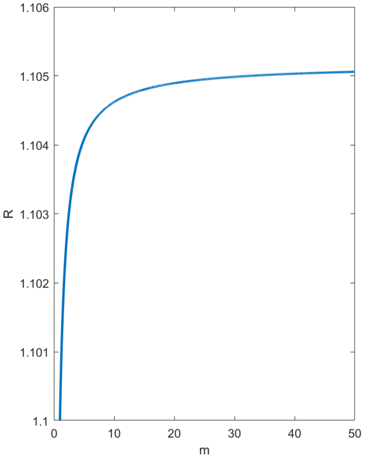
Table: Values of a loan with 10% per annum interest rate

Effect of compounding

In general, if the bank gives interest m times a year, you get

$$\$10,000 \times \left(1 + \frac{10\%}{m}\right)^m.$$

What if $m \to \infty$?



Continuous compounding

Suppose the continuously compounded interest rate is r, the simple gross return, or the *effective annual interest rate*, is

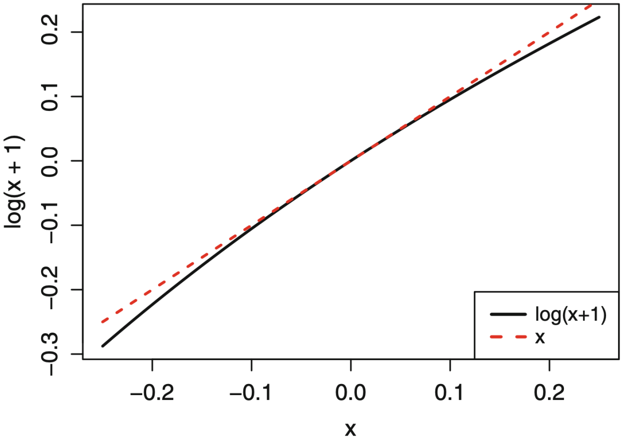
$$1 + R = \lim_{m \to \infty} \left(1 + \frac{r}{m}\right)^m .$$

Taking logarithm, and by L'Hopital's Rule

$$\lim_{m \to \infty} m \ln\left(1 + \frac{r}{m}\right) = r.$$

Therefore, $1 + R = \mathrm{e}^r$, or $r = \ln(1 + R)$, where r is also called the *log return*.

Log returns



- The difference between simple returns and log returns is small.
- One advantage of using log returns is simplicity of multi-period returns, which can be written as

$$
\begin{aligned}
1 + R_t(k) = \frac{P_t}{P_{t-k}} &= \left(\frac{P_t}{P_{t-1}}\right) \cdots \left(\frac{P_{t-k+1}}{P_{t-k}}\right) \\
&= (1 + R_t) \cdots (1 + R_{t-k+1}) \\
&= \exp(r_t) \cdots \exp(r_{t-k+1}).
\end{aligned}
$$

Taking logarithm of both sides,

$$
r_t(k) = \ln(1 + R_t(k)) = r_t + \cdots + r_{t-k+1}
$$

Efficient Market Hypothesis

Efficient markets

An efficient market is one where:

- important current information is almost freely available to all participants, and
- where there are large numbers of rational, profit-maximizers actively competing, with each trying to predict future market values of individual securities.

Forms of efficient markets

Weak Today's stock prices reflect all the information of past prices.

- No form of technical analysis can be effectively utilized to aid investors in making trading decisions.
- Fundamental analysis can be used to determine undervalued and overvalued stocks through research on companies' financial statements.

Forms of efficient markets

Semi-srtong All information that is public is used in the calculation of a stock's current price.

- Investors cannot utilize either technical or fundamental analysis to gain higher returns in the market.
- Only information that is not readily available to the public can help investors beat the market.

Strong All information—both the information available to the public and any information not publicly known—is completely accounted for in current stock prices.

- There is no type of information that can give an investor an advantage on the market.
- Investors cannot beat the market, regardless of information retrieved or research conducted.

Forms of efficient markets

- **Weak** Today's stock prices reflect all the information of past prices.
- **Semi-srtong** All information that is public is used in the calculation of a stock's current price.
- **Strong** All information—both the information available to the public and any information not publicly known—is completely accounted for in current stock prices.

Efficient markets

An efficient market is one where:

- important current information is almost freely available to all participants, and
- where there are a large number of rational profit-maximizers, actively competing with each trying to predict future market values of individual securities.

Implications of market efficiency

- Competition will cause the full effects of new information on intrinsic values to be reflected *instantaneously* in actual prices.
- Due to the vagueness or uncertainty surrounding new information,
  - actual prices will initially over-adjust to changes in intrinsic values as often as they will under-adjust;
  - the lags in the complete adjustment of actual prices to successive new intrinsic values will be independent.
- The "*instantaneous adjustment*" property of an efficient market implies that successive price changes in individual securities will be independent.

The Random Walk Model

Recall that the multi-period returns can be written as

$$1 + R_t(k) = \frac{P_t}{P_{t-k}} = \left(\frac{P_t}{P_{t-1}}\right) \cdots \left(\frac{P_{t-k+1}}{P_{t-k}}\right)$$
$$= (1 + R_t) \cdots (1 + R_{t-k+1})$$
$$= \exp(r_t) \cdots \exp(r_{t-k+1}).$$

Taking logarithm of both sides,

$$r_t(k) = \ln(1 + R_t(k)) = r_t + \cdots + r_{t-k+1}$$

where r_t are independent over t if the market is efficient and $r_t(k)$ follows a *random walk model*.

Let ε_t be a white noise process with zero mean and variance σ^2, or $\varepsilon_t \sim WN(0, \sigma^2)$, i.e., for all s, t,

$$\mathbb{E}\left[\varepsilon_t\right] = 0, \qquad \text{cov}(\varepsilon_s, \varepsilon_t) = \begin{cases} 0, & s \neq t \\ \sigma^2, & s = t \end{cases}$$

The cumulation of ε_t is called a random walk,

$$x_t = x_{t-1} + \varepsilon_t = x_0 + \sum_{s=1}^{t} \varepsilon_s, \qquad t = 1, 2, \ldots$$

Since ε_t is a white noise process, the moments of x_t can be obtained easily:

$$\mu = \mathbb{E}\left[x_t\right] = \mathbb{E}\left[x_0 + \sum_{s=1}^{t} \varepsilon_s\right] = x_0$$

$$\sigma_t^2 = \operatorname{var}(x_t) = \operatorname{var}\left(x_0 + \sum_{s=1}^{t} \varepsilon_s\right) = t\sigma^2$$

Since σ_t^2 increases with t, x_t is not stationary.

Wiener Process

Define an equidistant, disjoint partition of the continuous time interval $[0,1]$

$$[0,1) = \bigcup_{i=1}^{n} \left[\frac{i-1}{n}, \frac{i}{n}\right)$$

Interval-by-interval, define a scaled random walk as a continuous-time process of step function:

$$X_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} \varepsilon_j \qquad \text{for } t \in \left[\frac{i-1}{n}, \frac{i}{n}\right), \ i = 1, \ldots, n.$$

Define in addition for $t = 0$ and 1

$$X_n(0) = 0, \qquad X_n(1) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \varepsilon_j.$$

Suppose that $\sigma^2 = 1$, then by the Central Limit Theorem,

$$X_n(1) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \varepsilon_j \xrightarrow{d} \mathcal{N}(0,1).$$

Similarly, for any fixed $t \in (0,1)$, let $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x,

$$X_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nt \rfloor} \varepsilon_j = \frac{\sqrt{t}}{\sqrt{nt}} \sum_{j=1}^{\lfloor tn \rfloor} \varepsilon_j \xrightarrow{d} \mathcal{N}(0,t).$$

Theorem

Let $S_n = \sum_{i=1}^{n} \varepsilon_i,\ \varepsilon_i \overset{iid}{\sim} (0,1)$ *be a random walk. Define the re-scaled partial-sum process*

$$X_n(t) = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}, \qquad t \in [0,1].$$

Then, the Donsker's Theorem, or the functional central limit theorem, states that $X_n(t) \Rightarrow W(t), t \in [0,1]$*, where* $W(t)$ *is the Wiener process.*

Definition

A stochastic process $W(t),\ t \in [0,T]$ is said to be a *Wiener process*, or a *standard Brownian motion*, if:

1. Zero starting value: $P(W(0)=0)=1$;
2. Independent increments: for any $0 \le t_0 \le t_1 \le \cdots \le t_n$, $W(t_1)-W(t_0), \ldots, W(t_n)-W(t_{n-1})$ are independent;
3. Stationary increments: $W(t+s)-W(s) \sim \mathcal{N}(0,t)$ for any $s, t > 0$.

If the log returns $r_t \overset{\text{iid}}{\sim} \mathcal{N}\left(\mu, \sigma^2\right)$, then

$$r_t(t) = r_1 + \cdots + r_t \sim \mathcal{N}(\mu t, \sigma^2 t).$$

Moreover,

$$P_t = P_0 \exp(r_1 + \cdots + r_t)$$

is *lognormal* since its logarithm is normally distributed.

R Lab

Use the data set `Stock_bond.csv` to answer the following questions:

1. Compute the returns and log returns for GM and plot them against each other.
2. Compute the mean, standard deviation, skewness and kurtosis of the log returns.
3. Create a QQ plot against (i) the standard normal distribution, (ii) the t-distribution with degrees of freedom 4, 10 and 30.

Mean

The first moment is the *mean*, which measures the (average) location of X.

$$\mu_X = \mathbb{E}\left[X\right]$$

The sample mean is

$$\widehat{\mu}_X = \frac{1}{T}\sum_{t=1}^{T} x_t$$

The second centered moment is the *variance*, which measures the dispersion of X around its mean.

$$\sigma_X^2 = \mathbb{E}\left[(X - \mu_X)^2\right]$$

The sample variance is

$$\widehat{\sigma}_X^2 = \frac{1}{T-1} \sum_{t=1}^{T} (x_t - \widehat{\mu}_X)^2$$

Skewness

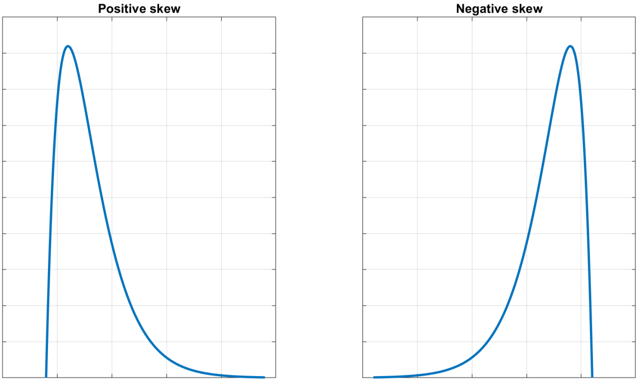
The third centered moment is *skewness*, which measures the degree of asymmetry in the distribution of X.

$$S(X) = \mathbb{E}\left[\frac{(X-\mu_X)^3}{\sigma_X^3}\right].$$

The sample skewness is

$$\widehat{S}(X) = \frac{1}{T\widehat{\sigma}_X^3}\sum_{t=1}^{T}(x_t-\widehat{\mu}_X)^3.$$

Skewness



Kurtosis

The fourth centered moment is *kurtosis*, which measures the fatness of the tails of the distribution of X.

$$K(X) = \mathbb{E}\left[\frac{(X-\mu_X)^4}{\sigma_X^4}\right]$$

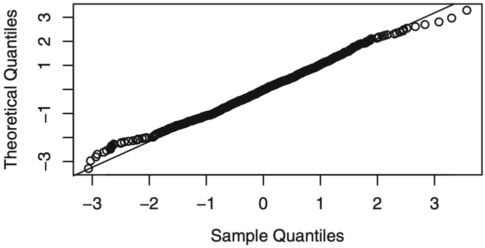
The sample kurtosis is

$$\widehat{K}(X) = \frac{1}{T\widehat{\sigma}_X^4}\sum_{t=1}^{T}(x_t - \widehat{\mu}_X)^4$$

Since the kurtosis of the normal distribution is 3, sometimes we report the *excess* kurtosis $\widehat{K}(X) - 3$ instead.

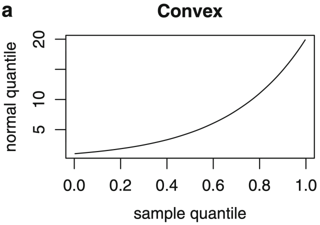
QQ plots

- A quantile-quantile plot, or a QQ plot, is a plot of the quantiles of one sample or distribution against the quantiles of a second sample or distribution.
- The QQ plot is *linear* if the samples (x-axis) and the reference (y-axis) share the same distribution, up to a shift and scaling.

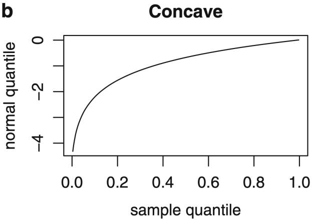


QQ plots

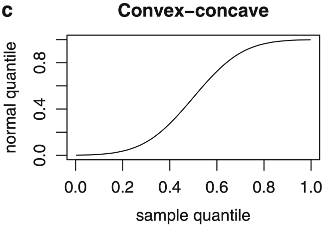
Left skewed



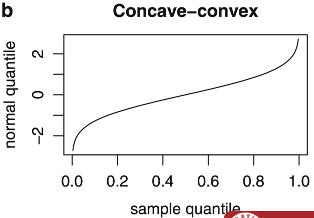
Right skewed



Heavy tail



Light tail



t-distribution