QUANTITATIVE FINANCE
PART I - INTRODUCTION ON ASSET RETURNS

Dr. Luo, Dan

April 1, 2020

Science is measurement.

—- Motto of the Econometric Society

(Source: https://todayinsci.com)

Theory without Practice is Empty;
Practice Without Theory is Blind.

—- Arguably Attributed to Immanuel Kant

What is important is the gradual development of a theory, based on a careful analysis of the ... facts.

- ◇ ... Its first applications are necessarily to elementary problems where the result has never been in doubt and no theory is actually required. At this early stage the application serves to corroborate the theory.
- ◇ The next stage develops when the theory is applied to somewhat more complicated situations in which it may already lead to a certain extent beyond the obvious and familiar. Here theory and application corroborate each other mutually.
- ◇ Beyond lies the field of real success: genuine prediction by theory.

It is well known that all mathematized sciences have gone through these successive stages of evolution.

—- John von Neumann

Looking back . . . over the long and labyrinthine path which finally led to the discovery [of the quantum theory], I am vividly reminded of Goethe's saying that men will always be making mistakes as long as they are striving after something.

—— Max Planck

Why Returns?

Perhaps paradoxically, we primarily analyze returns and the relations between different returns, not prices.

1. Investors have no market power, and hence take the price formation process as given.
   - For the average investor, financial markets may be considered close to being perfectly competitive (he/she is a price-taker).
   - This means the investment technology is constant returns to scale, so return is a scale-free summary of the investment opportunity.
   - Returns are widely examined in studies on portfolio formation and asset pricing (e.g., CAPM).

Why Returns?

2. Technically, returns processes have more attractive statistical properties than prices, such as stationarity and ergodicity.
   - ◇ Sample moments converge to the population moments so that economists can make sensible statistical analyses and economic decisions.
   - ◇ Covariance stationarity and ergodicity are typically assumed to ensure such convergence for time series.
   - ◇ Ludwig Boltzmann coined the term "ergodic" when he was working on a problem in statistical mechanics. Intuitively, the state of an ergodic process after a long time is nearly independent of its initial state. Accordingly, new and useful information to calculate the average and other moments continually arrives as the process evolves.

Simple Returns

The simple return from time $t - 1$ to t is

$$R_t \equiv \frac{P_t}{P_{t-1}} - 1, \qquad (1)$$

where P_t is the asset price at time t.

Returns are scale-free, not unitless! They are always defined w.r.t. some time interval, e.g., a year. Hence R_t is, in economic jargon, a *flow* variable, more properly called a *rate* of return.

Simple Gross Returns

The simple gross return from time $t-1$ to t is

$$1 + R_t.$$

The asset's gross return over the most recent k periods from time $t-k$ to t is

$$\begin{aligned} 1 + R_t(k) &\equiv (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}) \\ &= \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}}, \end{aligned} \tag{2}$$

where the multiperiod returns are also called *compound* returns.

Annualized Returns and Approximations

Among practitioners and in the financial press, returns are usually quoted on an annual basis. Suppose we invested for k years from year $t-k$ to t. Then

$$Annualized\ [R_t(k)] = \left[\prod_{i=0}^{k-1}(1+R_{t-i})\right]^{1/k} - 1. \tag{3}$$

A first-order Taylor expansion provides the following approximation:

$$Annualized\ [R_t(k)] = \frac{1}{k}\sum_{i=0}^{k-1} R_{t-i}. \tag{4}$$

(Hint: For $f(x_1, x_2, ..., x_k) = \left[\prod_{i=1}^{k} x_i\right]^{1/k} - 1$,
$f(1+R_t, 1+R_{t-1}, ..., 1+R_{t-k+1}) \approx f(1,1,...,1) + \sum_{i=1}^{k} f'_{x_i}(1,1,...,1)R_{t-i+1}$.)

A Slight Generalization

Suppose there are k years' data and n compounding periods in a year. Then the total number of periods is $m = nk$. Now,

$$Annualized\ [R_t(k)] = \left[\prod_{i=0}^{m-1}(1 + R_{t-i})\right]^{1/k} - 1. \tag{5}$$

The approximation becomes:

$$Annualized\ [R_t(k)] = \frac{1}{k}\sum_{i=0}^{m-1} R_{t-i}. \tag{6}$$

Numerical Examples

Suppose we invest for 5 years and there are 52 weeks, or 360 days, in a year.

	Annualized Return	Approximation
3% per quarter ($n = 4$)	12.55%	12.00%
1% per month ($n = 12$)	12.68%	12.00%
0.2308% per week ($n = 52$)	12.73%	12.00%
0.0333% per day ($n = 360$)	12.75%	12.00%
Continuously ($n \to \infty$)	12.75%	12.00%

Depending on the volatility of returns, the approximation can be poor.

Continuously Compounded Return

To overcome the difficulty in manipulating the geometric average in (5), the notion of continuous compounding is developed. The continuously compounded return or log return r_t of an asset is defined as

$$r_t \equiv \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1}, \tag{7}$$

where $p_t \equiv \ln P_t$.

The advantages of continuously compounded returns become clear: Conversion of products to sums!

$$\begin{aligned} r_t(k) &\equiv \ln(1 + R_t(k)) = \ln\left((1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1})\right) \\ &= \ln(1 + R_t) + \ln(1 + R_{t-1}) + \cdots + \ln(1 + R_{t-k+1}), \\ &= r_t + r_{t-1} + \cdots + r_{t-k+1}. \end{aligned} \tag{8}$$

Further Discussions on Pros and Cons

Pros for continuously compounded returns:

1. In the modeling of the statistical behavior of asset returns over time, it is far easier to derive the time-series properties of additive processes than of multiplicative processes.
2. Limited liability is imposed in a straightforward way, as price always stays positive under continuously compounded returns.

$$P_t = P_{t-k} exp(r_t(k)) > 0.$$

Further Discussions on Pros and Cons

Cons for continuously compounded returns: The simple return on a portfolio is a value-weighted average of the simple returns on the underlying assets, a convenient property not shared by continuously compounded returns.

- For example, an equally-weighted portfolio consists of two assets with $R_1 = 4\%$ and $R_2 = 20\%$. Then the portfolio simple return is $R_p = 12\%$.
- On the contrary, $r_p = \ln(1 + R_p) = 11.33\%$, but $0.5 \ln(1 + R_1) + 0.5 \ln(1 + R_2) = 11.08\%$.

Usually, we do not differentiate since the problem is minor, especially when returns are measured over short time intervals. However, conventionally, we use simple returns when studying a cross-section of assets, and continuously compounded returns when focusing on the time series behavior of returns.

Dividends

If the asset pays a dividend D_t just before the price P_t is recorded at time t. The net simple return at t is

$$R_t \equiv \frac{P_t + D_t}{P_{t-1}} - 1, \tag{9}$$

where P_t is the *ex-dividend* price at t.

The continuously compounded return on a dividend-paying asset is

$$r_t = \ln \frac{P_t + D_t}{P_{t-1}}, \tag{10}$$

which is a nonlinear function of log price and log dividend. We can only obtain the linearity approximately, when the dividend-price ratio is relatively stable. Mathematically, if $\frac{D_t}{P_t} \approx C$, where C is a constant, then

$r_t = \ln \frac{P_t+D_t}{P_{t-1}} \approx \ln \frac{P_t+CP_t}{P_{t-1}} = \ln \left((1+C)\frac{P_t}{P_{t-1}}\right) = \ln(1+C) + \ln P_t - \ln P_{t-1}.$

Excess Returns

An excess return (Z_{it}) is defined as the difference between an asset i's return (R_{it}) and the return (R_{0t}) on some reference asset (e.g., a riskless asset like the 3-month treasury bill).

$$Z_{it} \equiv R_{it} - R_{0t}. \tag{11}$$

Or, in terms of log returns,

$$z_{it} \equiv r_{it} - r_{0t} \neq \ln(1 + Z_{it}). \tag{12}$$

The excess return can be thought of as the payoff on an arbitrage portfolio, long asset i and short the reference asset so the net initial investment is zero. Note that the return on the portfolio is undefined due to zero investment but the payoff will be proportional to the excess return.

Question 1

A log excess return is commonly defined as the difference in log returns, i.e., $z_{it} \equiv r_{it} - r_{0t}$, not the log of simple excess returns, i.e., $z'_{it} = \ln(1 + Z_{it})$. Compare both definitions.

Question 1

A log excess return is commonly defined as the difference in log returns, i.e., $z_{it} \equiv r_{it} - r_{0t}$, not the log of simple excess returns, i.e., $z'_{it} = \ln(1 + Z_{it})$. Compare both definitions.

Suppose you long amount A of asset i and short the same amount of a riskless asset. Your payoff will be

$$A(1 + R_{it}) - A(1 + R_{0t}) = AZ_{it}.$$

We know that

$$A(exp(z'_{it}) - 1) = A\left(exp(\ln(1 + Z_{it})) - 1\right) = AZ_{it}.$$

Hence in this sense, a log excess return *SHOULD* be defined this way.

However, conventionally,[1] a log excess return is defined by z_{it}. If we mechanically treat it as a log return, your payoff will be

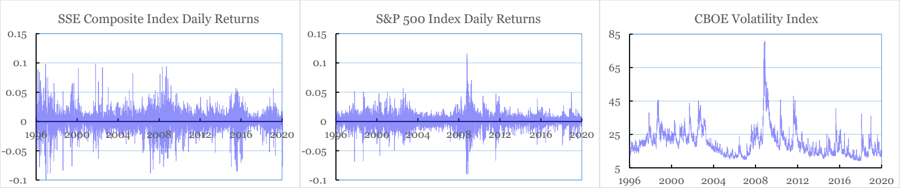
$$\begin{aligned} A(exp(z_{it}) - 1) &= A(exp(r_{it} - r_{0t}) - 1) = A\left(\frac{1 + R_{it}}{1 + R_{0t}} - 1\right) \\ &= A\frac{R_{it} - R_{0t}}{1 + R_{0t}} \approx AZ_{it}, \text{for a small } R_{0t}. \end{aligned} \tag{13}$$

Clearly, such a log excess return only approximates your payoff.

[1] Mainly, this gives the convenience of working with sums, instead of products.

Uncertainty

- Uncertainty lies at the heart of financial economics.
- The theme of financial economics includes risk measurement and management, pricing of risk, and financial decisions under uncertainty.
- Finance without uncertainty would be superfluous.



(Source: https://finance.yahoo.com/)

The Joint Distribution

Consider the returns of N assets for T years. Perhaps the most general model of the collection of returns is its joint distribution function:

$$\boldsymbol{F}(R_{11}, R_{21}, ..., R_{N1}; R_{12}, R_{22}, ..., R_{N2}; R_{1T}, R_{2T}, ..., R_{NT}; \boldsymbol{s}|\theta), \qquad (14)$$

where $\boldsymbol{s}$ is the vector of state variables that capture the relevant economic environment. For instance, we simply take $\boldsymbol{s} \in \{Contraction, Expansion\}$, or $\boldsymbol{s}$ as time-varying return volatility. θ is a vector of fixed parameters that uniquely determines $\boldsymbol{F}$. Financial econometrics focuses on inferring θ given the returns and state variables[2].

[2] In principle, we can estimate the parameters and state variables based solely on the returns.

Return Behavior Across Assets

$$\boldsymbol{F}(R_{11}, R_{21}, ..., R_{N1}; R_{12}, R_{22}, ..., R_{N2}; R_{1T}, R_{2T}, ..., R_{NT}; \boldsymbol{s}|\theta). \quad (15)$$

◇ The above model is too general to be useful. Asset pricing models provide further restrictions on $\boldsymbol{F}$. For example, a model may assume returns are statistically independent through time. Hence the joint distribution of the cross-section of returns is time-invariant. The celebrated CAPM falls into this category.

Return Behavior Over Time

To examine the dynamics of the returns of a single asset or index (a stock index is sometimes called an aggregate stock), we turn to the joint distribution:

$$\boldsymbol{F}(R_1, R_2, ..., R_T). \tag{16}$$

Using the Bayesian rule[3], the above distribution can be written as

$$\boldsymbol{F}_{\boldsymbol{1}}(R_1)\boldsymbol{F}_{\boldsymbol{2|1}}(R_2|R_1)\cdots\boldsymbol{F}_{\boldsymbol{T|T-1,...,1}}(R_T|R_{T-1}, R_{T-2}, R_1). \tag{17}$$

◇ Return predicability relies on the temporal dependencies implicit in R_t shown above.

[3] Repeatedly use $\boldsymbol{F}_{\boldsymbol{2|1}}(R_2|R_1) = \frac{\boldsymbol{F}_{\boldsymbol{1,2}}(R_1, R_2)}{\boldsymbol{F}_{\boldsymbol{1}}(R_1)}$.

The Marginal/Unconditional Distribution

When predictability is a minor issue, the marginal/unconditional distribution of returns is of interest. We most commonly assume returns of an asset follow a temporally independently and identically (*i.i.d.*) normal distribution. It delivers the convenience that sums of normally distributed random variables are normal.

- ◇ If simple returns are normal, then they can be lower than -100% and violate limited-liability.
- ◇ According to (2), multi-period returns cannot be normal because they are products of simple returns.

The Marginal/Unconditional Distribution

An alternative is to let single period simple gross returns be lognormally distributed so that the continuously compounded or log returns are normally distributed.

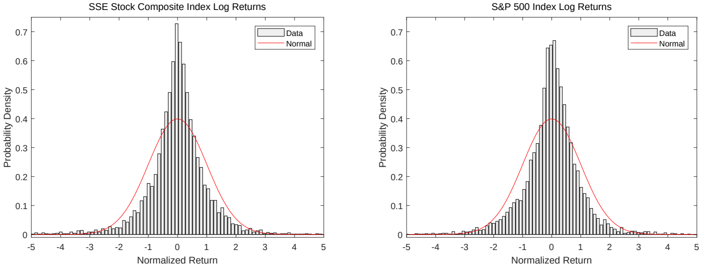
- For asset i, $r_{it} = \ln(1+R_{it}) \sim N(\mu_i, \sigma_i)$, with the density function $f(r_{it}) = \frac{1}{\sqrt{2\pi}\sigma_i} exp\left(-\frac{(r_{it}-\mu_i)^2}{2\sigma_i^2}\right)$.
- $1 + R_{it}$ is lognormally distributed and thus has a minimum realization of zero. The density function of $1 + R_{it}$ is $f(1+R_{it}) = \frac{1}{\sqrt{2\pi}\sigma_i(1+R_{it})} exp\left(-\frac{(\ln(1+R_{it})-\mu_i)^2}{2\sigma_i^2}\right)$.
- We can compute that

$$\boldsymbol{E}(R_{it}) = exp\left(\mu_i + \frac{1}{2}\sigma_i^2\right) - 1,$$

$$\boldsymbol{Var}(R_{it}) = (exp(\sigma_i^2) - 1)exp(2\mu_i + \sigma_i^2).$$

The Lognormal Distribution

The lognormal model becomes the workhorse model of financial economics.[4] How does it perform empirically? We do tests using the SSE Stock Composite Index and S&P 500 Index daily log returns. The sample period is Jan. 1996 to Dec. 2019.



The actual return distribution is more peaked, and has fatter tails.

[4] Especially after the ground-breaking work of Black and Scholes (1973).

Diagnostics

Two measures for deviations from normal distributions:

$$Skewness = \mathbf{E}\left[\frac{(r_{it}-\mu_i)^3}{\sigma_i^3}\right] \quad \left(or\ \frac{1}{T}\sum_{t=1}^{T}\frac{(r_{it}-\hat{\mu}_i)^3}{\hat{\sigma}_i^3}\right),$$

$$Excess\ Kurtosis = \mathbf{E}\left[\frac{(r_{it}-\mu_i)^4}{\sigma_i^4}\right] - 3 \quad \left(or\ \frac{1}{T}\sum_{t=1}^{T}\frac{(r_{it}-\hat{\mu}_i)^4}{\hat{\sigma}_i^4} - 3\right).$$

Useful Formulas for Normal Distributions

For normal distributions, we calculate directly with the normal density to get:

$$
\begin{aligned}
Mean &= \boldsymbol{E}\left[r_{it}\right] = \mu_i, \\
Standard\ Deviation &= \sqrt{\boldsymbol{E}\left[(r_{it} - \mu_i)^2\right]} = \sigma_i, \\
Skewness &= \boldsymbol{E}\left[\frac{(r_{it} - \mu_i)^3}{\sigma_i^3}\right] = 0, \\
Excess\ Kurtosis &= \boldsymbol{E}\left[\frac{(r_{it} - \mu_i)^4}{\sigma_i^4}\right] - 3 = 0.
\end{aligned}
$$

Diagnostics Results

	Normal	SSE SCI	S&P 500
Skewness	0	-0.3793	-0.2649
Excess Kurtosis	0	5.2796	8.1547

At short (e.g., daily) horizons, we obtain:

- Weak evidence of skewness
- Strong evidence of excess kurtosis

A Generalized Model

One idea is to use a mixture of normal distributions, perhaps because of time-varying volatility. Let consider a simple case of only two volatility levels, one corresponding to high volatility during market turbulence and one corresponding to low volatility during “quiet” periods. Then the density function is

$$f(r_{it}) = p_L \frac{1}{\sqrt{2\pi}\sigma_{iL}} exp\left(-\frac{(r_{it}-\mu_i)^2}{2\sigma_{iL}^2}\right) + (1-p_L)\frac{1}{\sqrt{2\pi}\sigma_{iH}} exp\left(-\frac{(r_{it}-\mu_i)^2}{2\sigma_{iH}^2}\right), \tag{18}$$

where p_L is the probability of low volatility and it mixes the two normal densities. Obviously, $0 \leq p_L \leq 1$.

To get further insights, we need to estimate the new parameters p_L, σ_{iL}, and σ_{iH}. How? One approach is the Maximum Likelihood Estimation (MLE).

We do not go into deep theory, but illustrate the steps in MLE.

1. Derive the log likelihood function $log\mathcal{L}$ (why log?). In our case,

$$log\mathcal{L} = \ln \prod_{t=1}^{T} f(r_{it}) = \sum_{t=1}^{T} \ln f(r_{it}).$$

2. Choose some sensible initial parameter values for p_L, μ_i, σ_{iL}, and σ_{iH}. Solve

$$\max_{p_L,\mu_i,\sigma_{iL},\sigma_{iH}} log\mathcal{L}.$$

See the spreadsheet *mnormal.xlsx* for implementation using excel solver.

Useful Formulas

We use $\phi(r_{it}, \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} exp\left(-\frac{(r_{it}-\mu_i)^2}{2\sigma_i^2}\right)$ to denote the normal density with mean μ_i and standard deviation σ_i.

Then the density function of a normal mixture distribution is

$$f(r_{it}) = p_L\phi(r_{it}, \mu_i, \sigma_{iL}) + (1 - p_L)\phi(r_{it}, \mu_i, \sigma_{iH}). \tag{19}$$

And

$$\begin{aligned} Mean = \boldsymbol{E}\left[r_{it}\right] &= \int_{-\infty}^{+\infty} r_{it} f(r_{it}) dr_{it} \\ &= p_L \int_{-\infty}^{+\infty} r_{it}\phi(r_{it}, \mu_i, \sigma_{iL}) dr_{it} + (1 - p_L) \int_{-\infty}^{+\infty} r_{it}\phi(r_{it}, \mu_i, \sigma_{iH}) dr_{it} \\ &= p_L\mu_i + (1 - p_L)\mu_i \\ &= \mu_i. \end{aligned} \tag{20}$$

Useful Formulas - Continued

$$\begin{aligned} \textit{Variance } \sigma_i^2 = \boldsymbol{E}\left[(r_{it}-\mu_i)^2\right] &= \int_{-\infty}^{+\infty}(r_{it}-\mu_i)^2 f(r_{it})dr_{it}, \\ &= p_L\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^2\phi(r_{it},\mu_i,\sigma_{iL})dr_{it} \\ &+ (1-p_L)\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^2\phi(r_{it},\mu_i,\sigma_{iH})dr_{it} \\ &= p_L\sigma_{iL}^2+(1-p_L)\sigma_{iH}^2. \end{aligned} \tag{21}$$

Useful Formulas - Continued

$$
\begin{aligned}
\boldsymbol{E}\left[(r_{it}-\mu_i)^3\right] &= \int_{-\infty}^{+\infty} (r_{it}-\mu_i)^3 f(r_{it}) dr_{it}, \\
&= p_L \int_{-\infty}^{+\infty} (r_{it}-\mu_i)^3 \phi(r_{it},\mu_i,\sigma_{iL}) dr_{it} \\
&+ (1-p_L) \int_{-\infty}^{+\infty} (r_{it}-\mu_i)^3 \phi(r_{it},\mu_i,\sigma_{iH}) dr_{it} \\
&= p_L \times 0 + (1-p_L) \times 0 \\
&= 0. \\
Skewness &= \boldsymbol{E}\left[\frac{(r_{it}-\mu_i)^3}{\sigma_i^3}\right] = 0.
\end{aligned} \tag{22}
$$

Useful Formulas - Continued

$$
\begin{aligned}
\boldsymbol{E}\left[(r_{it}-\mu_i)^4\right] &= \int_{-\infty}^{+\infty}(r_{it}-\mu_i)^4 f(r_{it})dr_{it}, \\
&= p_L\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^4\phi(r_{it},\mu_i,\sigma_{iL})dr_{it} \\
&+ (1-p_L)\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^4\phi(r_{it},\mu_i,\sigma_{iH})dr_{it} \\
&= p_L\times 3\sigma_{iL}^4+(1-p_L)\times 3\sigma_{iH}^4. \\
\textit{Excess Kurtosis} &= \boldsymbol{E}\left[\frac{(r_{it}-\mu_i)^4}{\sigma_i^4}\right]-3 \\
&= \frac{3(p_L\sigma_{iL}^4+(1-p_L)\sigma_{iH}^4)}{(p_L\sigma_{iL}^2+(1-p_L)\sigma_{iH}^2)^2}-3.
\end{aligned} \tag{23}
$$

Discussion on the Kurtosis Formula

$$\textit{Excess Kurtosis} = 3\left(\frac{p_L\sigma_{iL}^4 + (1-p_L)\sigma_{iH}^4}{(p_L\sigma_{iL}^2 + (1-p_L)\sigma_{iH}^2)^2} - 1\right). \quad (24)$$

For the non-trivial case of $0 < p_L < 1$, then

- *Excess Kurtosis* is 0 when $\sigma_{iL} = \sigma_{iH}$.
- Furthermore, *Excess Kurtosis* is minimized when $\sigma_{iL} = \sigma_{iH}$.
- We can show that

$$\begin{aligned} & p_L\sigma_{iL}^4 + (1-p_L)\sigma_{iH}^4 - (p_L\sigma_{iL}^2 + (1-p_L)\sigma_{iH}^2)^2 \\ & = p_L(1-p_L)(\sigma_{iL}^2 - \sigma_{iH}^2)^2 \geq 0. \end{aligned} \quad (25)$$

Hence, other than the trivial case of degeneration ($\sigma_{iL} = \sigma_{iH}$), we always have *Excess Kurtosis* > 0.

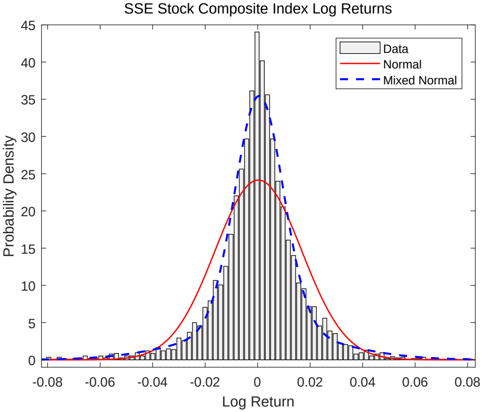
Results

Here are the results for SSE Stock Composite Index.

$$p_L = 0.726778,\ \mu_i = 0.000593,$$
$$\sigma_{iL} = 0.009196,\ \sigma_{iH} = 0.027834.$$

	Data	Normal	Mixed Normal
Mean	0.00029	0.00029	0.00059
St. Dev.	0.01653	0.01653	0.01653
Skewness	-0.3793	0	0.0000
Excess Kurtosis	5.2796	0	3.8036

Normal v.s. Mixed Normal



Further Extending the Analysis

Now we allow for different μ_{iL} and μ_{iH}. This could make the model asymmetric thus we can also fit the skewness. Then the density function of the full model is

$$
\begin{aligned}
f(r_{it}) = p_L \frac{1}{\sqrt{2\pi}\sigma_{iL}} exp\left(-\frac{(r_{it}-\mu_{iL})^2}{2\sigma_{iL}^2}\right) \\
+ (1-p_L)\frac{1}{\sqrt{2\pi}\sigma_{iH}} exp\left(-\frac{(r_{it}-\mu_{iH})^2}{2\sigma_{iH}^2}\right)
\end{aligned}
\tag{26}
$$

See the spreadsheet *mnormal_ex.xlsx* for implementation using excel solver.

Useful Formulas

The density function of a full normal mixture distribution is

$$f(r_{it}) = p_L \phi(r_{it}, \mu_{iL}, \sigma_{iL}) + (1 - p_L)\phi(r_{it}, \mu_{iH}, \sigma_{iH}). \tag{27}$$

And

$$\begin{aligned} \mu_i = \boldsymbol{E}[r_{it}] &= \int_{-\infty}^{+\infty} r_{it} f(r_{it}) dr_{it} \\ &= p_L \int_{-\infty}^{+\infty} r_{it}\phi(r_{it}, \mu_{iL}, \sigma_{iL}) dr_{it} \\ &+ (1 - p_L) \int_{-\infty}^{+\infty} r_{it}\phi(r_{it}, \mu_{iH}, \sigma_{iH}) dr_{it} \\ &= p_L \mu_{iL} + (1 - p_L)\mu_{iH}. \end{aligned} \tag{28}$$

Useful Formulas - Continued

$$
\begin{aligned}
\sigma_i^2 &= \boldsymbol{E}\left[(r_{it}-\mu_i)^2\right] = \int_{-\infty}^{+\infty}(r_{it}-\mu_i)^2 f(r_{it})dr_{it}, \\
&= p_L\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^2\phi(r_{it},\mu_{iL},\sigma_{iL})dr_{it} \\
&+ (1-p_L)\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^2\phi(r_{it},\mu_{iH},\sigma_{iH})dr_{it} \\
&= p_L\int_{-\infty}^{+\infty}((r_{it}-\mu_{iL})+(\mu_{iL}-\mu_i))^2\phi(r_{it},\mu_{iL},\sigma_{iL})dr_{it} \\
&+ (1-p_L)\int_{-\infty}^{+\infty}((r_{it}-\mu_{iH})+(\mu_{iH}-\mu_i))^2\phi(r_{it},\mu_{iH},\sigma_{iH})dr_{it} \\
&= p_L(\sigma_{iL}^2+(\mu_{iL}-\mu_i)^2)+(1-p_L)(\sigma_{iH}^2+(\mu_{iH}-\mu_i)^2) \\
&= p_L\sigma_{iL}^2+(1-p_L)\sigma_{iH}^2+\underbrace{p_L(1-p_L)(\mu_{iL}-\mu_{iH})^2}_{\text{An additional term due to difference in means, }\geq 0}.
\end{aligned}
\tag{29}
$$

Useful Formulas - Continued

$$
\begin{aligned}
\boldsymbol{E}\left[(r_{it}-\mu_i)^3\right] &= \int_{-\infty}^{+\infty}(r_{it}-\mu_i)^3 f(r_{it})dr_{it},\\
&= p_L\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^3\phi(r_{it},\mu_{iL},\sigma_{iL})dr_{it}\\
&+ (1-p_L)\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^3\phi(r_{it},\mu_{iH},\sigma_{iH})dr_{it}\\
&= p_L\int_{-\infty}^{+\infty}((r_{it}-\mu_{iL})+(\mu_{iL}-\mu_i))^3\phi(r_{it},\mu_{iL},\sigma_{iL})dr_{it}\\
&+ (1-p_L)\int_{-\infty}^{+\infty}((r_{it}-\mu_{iH})+(\mu_{iH}-\mu_i))^3\phi(r_{it},\mu_{iH},\sigma_{iH})dr_{it}\\
&= p_L(3(\mu_{iL}-\mu_i)\sigma_{iL}^2+(\mu_{iL}-\mu_i)^3)\\
&+ (1-p_L)(3(\mu_{iH}-\mu_i)\sigma_{iH}^2+(\mu_{iH}-\mu_i)^3).
\end{aligned}
\tag{30}
$$

Discussion on the Skewness Formula

$$\boldsymbol{E}\left[(r_{it}-\mu_i)^3\right] = 3p_L(1-p_L)(\sigma_{iL}^2-\sigma_{iH}^2)(\mu_{iL}-\mu_{iH}) + p_L(1-p_L)(1-2p_L)(\mu_{iL}-\mu_{iH})^3. \quad (31)$$

For non-trivial $p_L \in (0,1)$, there are three special cases.

1. Equal mean: If $\mu_{iL} = \mu_{iH}$, $Skewness = 0$.
2. Equal standard deviation: If $\sigma_{iL} = \sigma_{iH}$, $Skewness = p_L(1-p_L)(1-2p_L)(\mu_{iL}-\mu_{iH})^3$ and depends on which normal distribution has a higher weight.
3. Equal weight: If $p_L = 0.5$, $Skewness = 3p_L(1-p_L)(\mu_{iL}-\mu_{iH})(\sigma_{iL}^2-\sigma_{iH}^2)$ and depends on which one has a higher dispersion (standard deviation).

Generally, the *Skewness* is flexible. It can be positive or negative or zero, depending on all the parameters.

Useful Formulas - Continued

$$
\begin{aligned}
\boldsymbol{E}\left[(r_{it}-\mu_i)^4\right] &= \int_{-\infty}^{+\infty}(r_{it}-\mu_i)^4 f(r_{it})dr_{it}, \\
&= p_L\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^4\phi(r_{it},\mu_{iL},\sigma_{iL})dr_{it} \\
&+ (1-p_L)\int_{-\infty}^{+\infty}(r_{it}-\mu_i)^4\phi(r_{it},\mu_{iH},\sigma_{iH})dr_{it} \\
&= p_L\int_{-\infty}^{+\infty}((r_{it}-\mu_{iL})+(\mu_{iL}-\mu_i))^4\phi(r_{it},\mu_{iL},\sigma_{iL})dr_{it} \\
&+ (1-p_L)\int_{-\infty}^{+\infty}((r_{it}-\mu_{iH})+(\mu_{iH}-\mu_i))^4\phi(r_{it},\mu_{iH},\sigma_{iH})dr_{it} \\
&= p_L(3\sigma_{iL}^4+6(\mu_{iL}-\mu_i)^2\sigma_{iL}^2+(\mu_{iL}-\mu_i)^4) \\
&+ (1-p_L)(3\sigma_{iH}^4+6(\mu_{iH}-\mu_i)^2\sigma_{iH}^2+(\mu_{iH}-\mu_i)^4).
\end{aligned}
\tag{32}
$$

Discussion on the Kurtosis Formula

$$\boldsymbol{E}\left[(r_{it}-\mu_i)^4\right] = \underbrace{3(p_L\sigma_{iL}^4 + (1-p_L)\sigma_{iH}^4)}_{\text{The old term due to difference in standard deviations}} + \underbrace{6p_L(1-p_L)((1-p_L)\sigma_{iL}^2 + p_L\sigma_{iH}^2)(\mu_{iL}-\mu_{iH})^2}_{\text{A new term from difference in means, } \geq 0} + \underbrace{p_L(1-p_L)(p_L^3 + (1-p_L)^3)(\mu_{iL}-\mu_{iH})^4}_{\text{A new term from difference in means, } \geq 0}. \quad (33)$$

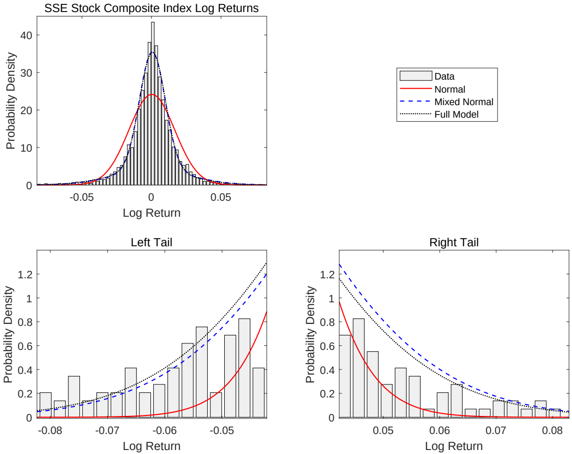
For the non-trivial case of $0 < p_L < 1$, this numerator further increases as long as $\mu_{iL} \neq \mu_{iH}$.

Results for the Extended Analysis

$$p_L = 0.730568,$$
$$\mu_{iL} = 0.000774,\ \mu_{iH} = -0.001021,$$
$$\sigma_{iL} = 0.009248,\ \sigma_{iH} = 0.027913.$$

	Data	Normal	Mixed Normal	Full Model
Mean	0.00029	0.00029	0.00059	0.00029
St. Dev.	0.01653	0.01653	0.01653	0.01652
Skewness	-0.3793	0	0.0000	-0.1631
Excess Kurtosis	5.2796	0	3.8036	3.8270

Normal v.s. Mixed Normal



Exercise 1

Repeat the analysis using the S&P 500 index daily returns data in the spreadsheet *SP500.xlsx*.

Growth of Mean and Variance Over Time

Suppose monthly (log) returns R_m follow an identical and independent normal distribution with mean μ_m and standard deviation σ_m. Let R_y be the corresponding yearly return thus $R_y = \sum_{t=1}^{12} R_{mt}$.
Then

$$\mu_y \equiv \boldsymbol{E}[R_y] = \boldsymbol{E}\left[\sum_{t=1}^{12} R_{mt}\right] = \sum_{t=1}^{12} \boldsymbol{E}[R_{mt}] = 12\mu_m. \tag{34}$$

And

$$\sigma_y^2 \equiv \boldsymbol{Var}[R_y] = \boldsymbol{E}\left[\left(\sum_{t=1}^{12} R_{mt} - \mu_y\right)^2\right] = \sum_{t=1}^{12} \boldsymbol{E}\left[(R_{mt} - \mu_m)^2\right] = 12\sigma_m^2. \tag{35}$$

Equivalently,

$$\mu_m = \frac{1}{12}\mu_y,$$
$$\sigma_m = \frac{1}{\sqrt{12}}\sigma_y.$$

Generally, if we divide a year into p subperiods, then

$$\mu_p = \frac{1}{p}\mu_y,$$
$$\sigma_p = \frac{1}{\sqrt{p}}\sigma_y.$$

The square root lies at the heart of our discussions in the following.

The Effect of Data Frequency

$$\frac{\mu_p}{\sigma_p} = \frac{1}{\sqrt{p}} \frac{\mu_y}{\sigma_y}. \tag{36}$$

As we measure asset returns more frequently ($\frac{\mu_y}{\sigma_y}$ is held constant), we reduce the time interval and increase p. This has an effect on the above mean-standard deviation ratio. The mean falls quickly while the standard deviation relatively slowly thanks to the square root. Hence the ratio decreases.

- Over a short time interval, noise dominates signal. That is, standard deviation dominates the expected value.

Numeric Examples

Take $\mu_y = 12\%$ and $\sigma_y = 20\%$. And there are 4 quarters, 12 months, 52 weeks, or 252 trading days in a year. To the extreme, the highest frequency of returns data used in market microstructure studies is one per minute.[5] The stock market opens for 4 hours on a normal trading day. Then there are 60480 trading minutes in a year.

[5] At an even higher frequency, market microstructural effects, such as price discreteness and bid-ask bounce, will set in and returns will measure things besides changes in valuation.

Numeric Examples

	Yearly	Quarterly	Monthly	Weekly	Daily	Per Minute
μ	12%	3%	1%	0.23%	0.048%	0.00020%
σ	20%	10%	5.77%	2.77%	1.26%	0.081%
μ/σ	0.6	0.3	0.17	0.083	0.038	0.0024
σ/μ	1.67	3.33	5.77	12.02	26.46	409.88

- The mean is miserable compared to the standard deviation at the 1-minute level.[6]
- Our common sense also has that the stock price moves around from minute to minute. (Like a shark?)
- At the daily level, the standard deviation is still 26 times of the mean.

[6] This is the reason the market microstructure studies usually neglect the mean.

Another implication is that estimation of means are very difficult. First consider yearly returns.

- Returns are *i.i.d.* normal thus stationary.
- We have T years of data.
- Then our estimate of the mean is

$$\hat{\mu}_y = \frac{1}{T}\sum_{t=1}^{T} R_{yt}.$$

  The estimate is unbiased as $\boldsymbol{E}\hat{\mu}_y = \mu_y$.
- The standard deviation of our estimate is

$$\sigma_{\hat{\mu}_y} = \sqrt{\boldsymbol{E}\left[(\hat{\mu}_y - \mu_y)^2\right]} = \sqrt{\boldsymbol{E}\left[\left(\frac{1}{T}\sum_{t=1}^{T}(R_{yt} - \mu_y)\right)^2\right]} = \frac{\sigma_y}{\sqrt{T}},$$

  which is the estimation error in the mean.

An Illustration

Again, Take $\mu_y = 12\%$ and $\sigma_y = 20\%$. Suppose we have 4 years' data. Recall

$$\sigma_{\hat{\mu}_y} = \frac{\sigma_y}{\sqrt{T}} = \frac{20\%}{\sqrt{4}} = 10\%.$$

The estimation is close to the mean. A 95% confidence interval for the mean is (-8%, 32%). The is too coarse to be of use.

A good estimate may have a standard deviation one-tenth of the mean. To reach this criterion, we can compute that the length of data should be no less than $T = \left(\frac{10\sigma_y}{\mu_y}\right)^2 = 278$ years. We hardly have one stock market with such a long history. To make things worse, the mean and standard deviation vary a lot in the long run. Hence the stationarity assumption is likely to be true only for a much shorter time.

Does High Frequency Data Help?

We have illustrated the mean blur using yearly returns. However, we do have higher frequency returns data readily available in the market.

Question 2

What if we switch to daily data?

Does High Frequency Data Help?

Question 2

What if we switch to daily data?

At daily frequency,

$$\mu_d = \frac{12\%}{252} = 0.048\%, \ \sigma_d = \frac{20\%}{\sqrt{252}} = 1.26\%.$$

Similarly, we have $\boldsymbol{E}[\hat{\mu}_d] = 0.048\%$. If we have T years' data, $\sigma_{\hat{\mu}_d} = \frac{1.26\%}{\sqrt{252T}}$. To get an error that is one-tenth of the mean, we need again $T = 278$ years' data. Increasing the data frequency does not help at all! (Show it for any data frequency by yourself.)

- ◇ The only way to obtain a more accurate estimate for the mean is to have a longer time series. Most of the times, there is nothing to do for this other than to wait.

Higher Moments

Higher moments are a lot easier. Let's instance this with variance. Consider a sample of returns $R_t \sim i.i.d.\ N(\mu, \sigma)$ with T observations. Then

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t,$$

$$\hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=1}^{T}(R_t - \hat{\mu})^2.$$

$\boldsymbol{E}\left[\hat{\sigma}^2\right] = \sigma^2$. The variance of this variance estimate[7] is

$$\sigma^2_{\hat{\sigma}^2} = \frac{2\sigma^4}{T-1}. \Rightarrow \sigma_{\hat{\sigma}^2} = \frac{\sqrt{2}\sigma^2}{\sqrt{T-1}}.$$

[7] Here, $\hat{\sigma}^2$ follows a χ^2 distribution with $(T-1)$ degrees of freedom. Hence the variance needs to be understood with this in mind.

Derivations for $\sigma^2_{\hat{\sigma}^2}$

We repeatedly use the following three results:

1. For any normally distributed variable x with mean μ and variance σ^2, $\boldsymbol{E}\left[(x-\mu)^n\right] = 0$ if n is odd, and $\boldsymbol{E}\left[(x-\mu)^n\right] = \sigma^n (n-1)(n-3)\cdots 1$ if n is even.
2. $\left(\sum_{i=1}^T A_i\right)^2 = \sum_{i=1}^T \sum_{j=1}^T A_i A_j = \sum_{i=1}^T \left(A_i \sum_{j=1}^T A_j\right) = \sum_{i=1}^T A_i \sum_{j=1}^T A_j$.
3. $$\begin{aligned}\boldsymbol{E}\left[(R_t-\hat{\mu})^2\right] &= \boldsymbol{E}\left[((R_t-\mu)-(\hat{\mu}-\mu))^2\right] \\ &= \boldsymbol{E}\left[(R_t-\mu)^2 - 2(R_t-\mu)(\hat{\mu}-\mu) + (\hat{\mu}-\mu)^2\right] \\ &= \sigma^2 - \frac{2}{T}\sigma^2 + \frac{1}{T}\sigma^2 = \frac{T-1}{T}\sigma^2.\end{aligned}$$

Derivations for $\sigma^2_{\hat{\sigma}^2}$ - Continued

$$
\begin{aligned}
\sigma^2_{\hat{\sigma}^2} &= \boldsymbol{E}\left[\frac{1}{T-1}\sum_{t=1}^{T}(R_t-\hat{\mu})^2-\sigma^2\right]^2 \\
&= \frac{1}{(T-1)^2}\boldsymbol{E}\left[\sum_{t=1}^{T}\left((R_t-\hat{\mu})^2-\frac{T-1}{T}\sigma^2\right)\right]^2 \\
&= \frac{1}{(T-1)^2}\boldsymbol{E}\left[\sum_{i=1}^{T}\sum_{j=1}^{T}\left((R_i-\hat{\mu})^2-\frac{T-1}{T}\sigma^2\right)\left((R_j-\hat{\mu})^2-\frac{T-1}{T}\sigma^2\right)\right] \\
&= \frac{1}{(T-1)^2}\boldsymbol{E}\left[\sum_{i=1}^{T}\sum_{j=1}^{T}\left((R_i-\hat{\mu})^2(R_j-\hat{\mu})^2-\frac{T-1}{T}\sigma^2((R_i-\hat{\mu})^2+(R_j-\hat{\mu})^2)+\frac{(T-1)^2}{T^2}\sigma^4\right)\right] \\
&= \frac{1}{(T-1)^2}\boldsymbol{E}\left[\underbrace{\sum_{i=1}^{T}\sum_{j=1}^{T}(R_i-\hat{\mu})^2(R_j-\hat{\mu})^2}_{③}\underbrace{-\frac{T-1}{T}\sigma^2\sum_{i=1}^{T}\sum_{j=1}^{T}((R_i-\hat{\mu})^2+(R_j-\hat{\mu})^2)}_{②}+\underbrace{\frac{(T-1)^2}{T^2}\sigma^4\sum_{i=1}^{T}\sum_{j=1}^{T}1}_{①}\right].
\end{aligned}
$$

Derivations for $\sigma^2_{\hat{\sigma}^2}$ - Continued

$$\sigma^2_{\hat{\sigma}^2} = \frac{1}{(T-1)^2} \boldsymbol{E}\left[③ + ② + ① \right].$$

$$\begin{aligned}
\boldsymbol{E}\left[①\right] &= (T-1)^2 \sigma^4. \\
\boldsymbol{E}\left[②\right] &= -\frac{T-1}{T}\sigma^2 \left(\sum_{i=1}^{T}\sum_{j=1}^{T} \boldsymbol{E}(R_i - \hat{\mu})^2 + \sum_{i=1}^{T}\sum_{j=1}^{T} \boldsymbol{E}(R_j - \hat{\mu})^2 \right) \\
&= -\frac{T-1}{T}\sigma^2 \times 2T^2 \times \frac{T-1}{T}\sigma^2 \\
&= -2(T-1)^2\sigma^4.
\end{aligned}$$

Derivations for $\sigma^2_{\hat{\sigma}^2}$ - Continued

$$
\begin{aligned}
\boldsymbol{E}[③] &= \boldsymbol{E}\left[\sum_{i=1}^{T}\left((R_i-\hat{\mu})^2\sum_{j=1}^{T}((R_j-\mu)-(\hat{\mu}-\mu))^2\right)\right] \\
&= \boldsymbol{E}\left[\sum_{i=1}^{T}\left((R_i-\hat{\mu})^2\left(\sum_{j=1}^{T}(R_j-\mu)^2-2(\hat{\mu}-\mu)\sum_{j=1}^{T}(R_j-\mu)+\sum_{j=1}^{T}(\hat{\mu}-\mu)^2\right)\right)\right] \\
&= \boldsymbol{E}\left[\sum_{i=1}^{T}\left((R_i-\hat{\mu})^2\left(\sum_{j=1}^{T}(R_j-\mu)^2-T(\hat{\mu}-\mu)^2\right)\right)\right] \\
&= \boldsymbol{E}\left[\left(\sum_{i=1}^{T}(R_i-\mu)^2-T(\hat{\mu}-\mu)^2\right)\left(\sum_{j=1}^{T}(R_j-\mu)^2-T(\hat{\mu}-\mu)^2\right)\right] \\
&= \boldsymbol{E}\left[\underbrace{\sum_{i=1}^{T}(R_i-\mu)^2\sum_{j=1}^{T}(R_j-\mu)^2}_{⑤}\underbrace{-T(\hat{\mu}-\mu)^2\left(\sum_{i=1}^{T}(R_i-\mu)^2+\sum_{j=1}^{T}(R_j-\mu)^2\right)}_{⑥}\underbrace{+T^2(\hat{\mu}-\mu)^4}_{④}\right].
\end{aligned}
$$

Derivations for $\sigma^2_{\hat{\sigma}^2}$ - Continued

$$\boldsymbol{E}[③] = \boldsymbol{E}[⑤ + ⑥ + ④].$$

$$\boldsymbol{E}[④] = (T^2) \times 3 \left(\frac{\sigma}{\sqrt{T}}\right)^4 = 3\sigma^4.$$

$$\begin{aligned}
\boldsymbol{E}[⑤] &= \sum_{i=1}^{T}\sum_{j=1}^{T} \boldsymbol{E}\left[(R_i - \mu)^2 (R_j - \mu)^2\right] \\
&= T \times 3\sigma^4 + T(T-1) \times \sigma^4 = T(T+2)\sigma^4.
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{E}[⑥] &= -\frac{2}{T}\boldsymbol{E}\left[\left(\sum_{j=1}^{T}(R_j - \mu)\right)^2 \left(\sum_{i=1}^{T}(R_i - \mu)^2\right)\right] \\
&= -\frac{2}{T}\boldsymbol{E}\left[\left(\sum_{j=1}^{T}\sum_{k=1}^{T}(R_j - \mu)(R_k - \mu)\right)\left(\sum_{i=1}^{T}(R_i - \mu)^2\right)\right] \\
&= -\frac{2}{T}\boldsymbol{E}\left[\sum_{j=1}^{T}\sum_{k=1}^{T}\sum_{i=1}^{T}(R_j - \mu)(R_k - \mu)(R_i - \mu)^2\right] \\
&= \underbrace{T \times 3\sigma^4}_{j=k=i} + \underbrace{T(T-1) \times \sigma^4}_{j=k\neq i} + \underbrace{(T^3 - T - T(T-1)) \times 0}_{j \neq k} = -2(T+2)\sigma^4.
\end{aligned}$$

Derivations for $\sigma^2_{\hat{\sigma}^2}$ - Continued

Finally,

$$\begin{aligned}\sigma^2_{\hat{\sigma}^2} &= \frac{1}{(T-1)^2}\boldsymbol{E}\left[① + ② + ④ + ⑤ + ⑥\right] \\ &= \frac{2\sigma^4}{T-1}.\end{aligned}$$

The proportion of error in variance estimate to variance is

$$\frac{\sigma_{\hat{\sigma}^2}}{\sigma^2} = \frac{\sqrt{2}}{\sqrt{T-1}}, \tag{37}$$

which is determined solely by T.
For yearly data, if we impose the one-tenth rule again, we get $T = 201$ years. Better. Not that impressive? Fortunately, this is *not* the end of story!

- ◇ We can always increase the accuracy by increasing the data frequency.
- ◇ We need to generalize (37) a little to account for sampling frequency.

A Comparison

Now we compare estimation of first and second moments in detail for an arbitrary data frequency. Suppose we have p (equally spaced) subperiods in a year and T years of data.

$$\sigma_{\hat{\mu}_p} = \frac{\sigma_p}{\sqrt{pT}}, \qquad \frac{\sigma_{\hat{\mu}_p}}{\mu_p} = \frac{\sigma_p}{\mu_p\sqrt{pT}} = \frac{\sigma_y/\sqrt{p}}{(\mu_y/p)\sqrt{pT}} = \frac{\sigma_y}{\mu_y}\cdot\frac{1}{\sqrt{T}},$$

$$\sigma_{\hat{\sigma}_p^2} = \frac{\sqrt{2}\sigma_p^2}{\sqrt{pT-1}}, \quad \frac{\sigma_{\hat{\sigma}_p^2}}{\sigma_p^2} = \frac{\sqrt{2}}{\sqrt{pT-1}}.$$

- Clearly, the error ratio of the mean only depends on T, while that of the variance also depends on p.

Numerical Examples

Take the previous numbers. There are 252 trading days in a year. Daily returns are *i.i.d*. normal with $\mu_d = 0.048\%$ and $\sigma_d = 1.26\%$. Simulate 10 years of daily returns. Then calculate $\hat{\mu}_d$ and $\hat{\sigma}_d$ for each year. Here is the results.

	Year1	Year2	Year3	Year4	Year5	Year6	Year7	Year8	Year9	Year10	Mean	St.Dev.
$\hat{\mu}_d$	0.021	-0.085	0.005	-0.041	0.071	0.067	0.021	0.024	-0.024	0.085	0.014	0.054
$\hat{\sigma}_d$	1.24	1.29	1.27	1.30	1.21	1.25	1.20	1.26	1.22	1.26	1.25	0.035

Other Side of the Coin

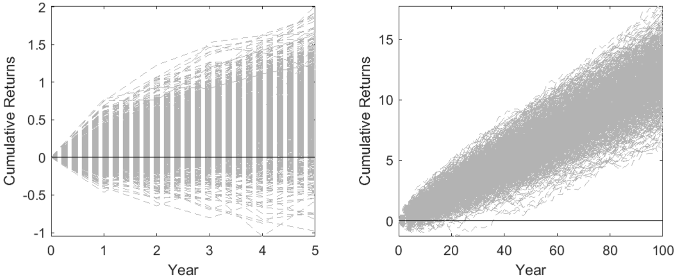
We have shown that during a short time interval, mean return is much smaller than the variance. How about long term returns?
We repeat (36) here.

$$\frac{\mu_p}{\sigma_p} = \frac{1}{\sqrt{p}} \frac{\mu_y}{\sigma_y}.$$

As we measure asset returns less frequently ($\frac{\mu_y}{\sigma_y}$ is held constant), we increase the time interval and decrease p. ($p < 1$ means returns over multiple years.)
The ratio could be very high for a very small p. That is, in the very long term, the mean will in fact dominate the variance. The signal will show up no matter how feeble it is or how rampant the noise is, as long as we give it enough time!

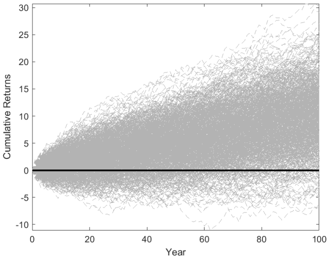
Numeric Example

Take $\mu_y = 12\%$ and $\sigma_y = 20\%$. Simulate returns for 100 years. We repeat for 1000 times to get 1000 alternative histories.



Diversification

Let's take another asset with $\mu_y = 10\%$ and $\sigma_y = 60\%$. Investing in it for 100 years, you expect to make a lot of money (to be exact, a log return of 1000% = a simple return of 2202500%). Run a similar simulation.



You still have a significant chance of losing money.

Question 3

What is the probability of losing money?

Question 3

What is the probability of losing money?

Investing for 100 years, the gross log return r_{100} is normally distributed with $\mu_{100} = 100\mu_y = 1000\%$ and $\sigma_{100} = \sqrt{100}\sigma_y = 600\%$.
Then

$$\begin{aligned} Prob.(losing\ money) &= Prob.(r_{100} < 0) \\ &= Prob.\left(\frac{r_{100} - \mu_{100}}{\sigma_{100}} < -\frac{\mu_{100}}{\sigma_{100}}\right) \\ &= \Phi(-\frac{\mu_{100}}{\sigma_{100}}) = \Phi(-1.67) \approx 5\%, \end{aligned}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

We can also find the loss probability by simulation. See the spreadsheets *SimuStock.xlsm* and *SingleAsset.xlsm*.

Diversification - Continued

Now take an equally-weight portfolio of two of the assets, each with $\mu_y = 10\%$ and $\sigma_y = 60\%$. The returns of the asset are independent. Investing in the portfolio for 100 years, you expect to make a similar amount of a lot of money.

Remember that the portfolio return is not the weighted average of the underlying returns, when we are using log returns.

Exercise 2

What is the probability of losing money when investing in the portfolio?

(See the spreadsheets *Portfolio.xlsm*.)

If you find a very slim chance of losing money, then try $\mu_y = -3.49\%$ and $\sigma_y = 51.04\%$ such that the simple returns have $\mu_y = 10\%$ and $\sigma_y = 60\%$.
What is the probability of losing money now?