1 Financial Time Series and Their Characteristics

Financial time series analysis is concerned with theory and practice of asset valuation over time. For example, there are various definitions of asset volatility, and for a stock return series, the volatility is not directly observable. As a result of the added uncertainty, statistical theory and methods play an important role in financial time series analysis.

The objective of this course is to provide some knowledge of financial time series, introduce some statistical tools useful for analyzing these series, and gain experience in financial applications of various econometric methods.

1.1 Asset Returns

Let $P_t$ be the price of an asset at time $t$. We discuss some definitions of returns that are used throughout the course. Assume for the moment that the asset pays no dividends.
One-Period Simple Return

Holding the asset for one period from date \( t - 1 \) to date \( t \) would result in a simple gross return

\[
1 + R_t = \frac{P_t}{P_{t-1}} \quad \text{or} \quad P_t = P_{t-1}(1 + R_t)
\]

The corresponding one-period simple net return or simple return is

\[
R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}.
\]

Multiperiod Simple Return

Holding the asset for \( k \) periods between dates \( t - k \) and \( t \) gives a \( k \)-period simple gross return

\[
1 + R_t[k] = \frac{P_t}{P_{t-k}}
\]

\[
= \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}}
\]

\[
= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1})
\]

\[
= \prod_{j=0}^{k-1} (1 + R_{t-j}).
\]

Thus, the \( k \)-period simple gross return is just the product of the \( k \) one-period simple gross returns involved. This is called a compound return. The \( k \)-period simple net return is

\[
R_t[k] = \frac{P_t - P_{t-k}}{P_{t-k}}.
\]

In practice, the actual time interval is important in discussing and comparing returns (e.g., monthly return or annual return).
If the time interval is not given, then it is implicitly assumed to be one year. If the asset was held for $k$ years, then the annualized (average) return is defined as

$$\text{Annualized}\{R_t[k]\} = \left[\prod_{j=0}^{k-1} (1 + R_{t-j})\right]^{\frac{1}{k}} - 1.$$ 

This is a geometric mean of the $k$ one-period simple gross returns involved and can be computed by

$$\text{Annualized}\{R_t[k]\} = \exp\left[\frac{1}{k} \sum_{j=0}^{k-1} \ln(1 + R_{t-j})\right] - 1,$$

where $\exp(x)$ denotes the exponential function and $\ln(x)$ is the natural logarithm of the positive number $x$. Because it is easier to compute arithmetic average than geometric mean and the one-period returns tend to be small, one can use a first-order Taylor expansion to approximate the annualized return and obtain

$$\text{Annualied}\{R_t[k]\} \approx \frac{1}{k} \sum_{j=0}^{k-1} R_{t-j}. \quad (1)$$

Accuracy of the approximation in (1) may not be sufficient in some applications, however.

**Continuous Compounding**

Before introducing continuously compounded return, we discuss the effect of compounding. Assume that the interest rate of a bank deposit is 10% per annum and the initial deposit is $1.00.
If the bank pays interest once a year, then the net value of the deposit becomes $1(1 + 0.1) = $1.1 one year later. If the bank pays interest semi-annually, the 6-month interest rate is $10%/2 = 5\%$ and the net value is $1(1 + 0.1/2)^2 = $1.1025 after the first year. In general, if the bank pays interest $m$ times a year, then the interest rate for each payment is $10%/m$ and the net value of the deposit becomes $1(1 + 0.1/m)^m$ one year later. Table 1.1 gives the results for some commonly used time intervals on a deposit of $1.00 with interest rate 10% per annum. In particular, the net value approaches $1.1052$, which is obtained by $\exp(0.1)$ and referred to as the result of continuous compounding. The effect of compounding is clearly seen.

Table 1.1. Illustration of the Effects of Compounding: The Time Interval Is 1 Year and the Interest Rate is 10% per Annum.

<table>
<thead>
<tr>
<th>Type</th>
<th>Number of payments</th>
<th>Interest rate per period</th>
<th>Net Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual</td>
<td>1</td>
<td>0.1</td>
<td>$1.10000</td>
</tr>
<tr>
<td>Semiannual</td>
<td>2</td>
<td>0.05</td>
<td>$1.10250</td>
</tr>
<tr>
<td>Quarterly</td>
<td>4</td>
<td>0.025</td>
<td>$1.10381</td>
</tr>
<tr>
<td>Monthly</td>
<td>12</td>
<td>0.0083</td>
<td>$1.10471</td>
</tr>
<tr>
<td>Weekly</td>
<td>52</td>
<td>$0.1/52</td>
<td>$1.10506</td>
</tr>
<tr>
<td>Daily</td>
<td>365</td>
<td>$0.1/365</td>
<td>$1.10516</td>
</tr>
<tr>
<td>Continuously</td>
<td>$\infty$</td>
<td>$0.1/365$</td>
<td>$1.10517</td>
</tr>
</tbody>
</table>

In general, the net asset value $A$ of continuous compounding
is

\[ A = C \exp(r \times n), \]

where \( r \) is the interest rate per annum, \( C \) is the initial capital, and \( n \) is the number of years. From the above equation, we have

\[ C' = A \exp(-r \times n), \]

which is referred to as the present value of an asset that is worth \( A \) dollars \( n \) years from now, assuming that the continuously compounded interest rate is \( r \) per annum.

**Continuously Compounded Return**

The natural logarithm of the simple gross return of an asset is called the continuously compounded return or log return:

\[ r_t = \ln(1 + R_t) = \ln\frac{P_t}{P_{t-1}} = p_t - p_{t-1}, \]

where \( p_t = \ln(P_t) \). Continuously compounded returns \( r_t \) enjoy some advantages over the simple net returns \( R_t \). First, consider multiperiod returns. We have

\[ r_t[k] = \ln(1 + R_t[k]) = \ln[(1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1})] = \ln(1 + R_t) + \ln(1 + R_{t-1}) + \cdots + \ln(1 + R_{t-k+1}) = r_t + r_{t-1} + \cdots + r_{t-k+1}. \]
Thus, the continuously compounded multiperiod return is simply the sum of continuously compounded one-period returns involved. Second, statistical properties of log returns are more tractable.

Portfolio Return

The simple net return of a portfolio consisting of $N$ assets is a weighted average of the simple net returns of the assets involved, where the weight on each asset is the percentage of the portfolio’s value invested in that asset. Let $p$ be a portfolio that places weight $w_i$ on asset $i$, then the simple return of $p$ at time $t$ is $R_{p,t} = \sum_{i=1}^{N} w_i R_{it}$, where $R_{it}$ is the simple return of asset $i$.

The continuously compounded returns of a portfolio, however, do not have the above convenient property. If the simple returns $R_{it}$ are all small in magnitude, then we have $r_{p,t} \approx \sum_{i=1}^{N} w_i r_{it}$, where $r_{p,t}$ is the continuously compounded return of the portfolio at time $t$. This approximation is often used to study portfolio returns.

Dividend Payment

If an asset pays dividends periodically, we must modify the definitions of asset returns. Let $D_t$ be the dividend payment of an asset between dates $t - 1$ and $t$ and $P_t$ be the price of the
asset at the end of period $t$. Thus, dividend is not included in $P_t$. Then the simple net return and continuously compounded return at time $t$ become

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1, \quad r_t = \ln(P_t + D_t) - \ln(P_{t-1}).$$

1.2 Distributional Properties of Returns

To study asset returns, it is best to begin with their distributional properties. The objective here is to understand the behavior of the returns across assets and over time. Consider a collection of $N$ assets held for $T$ time periods, say $t = 1, \cdots, T$. For each asset $i$, let $r_{it}$ be its log return at time $t$. The log returns under study are $\{r_{it}; i = 1, \cdots, N; t = 1, \cdots, T\}$. One can also consider the simple returns $\{R_{it}; i = 1, \cdots, N; t = 1, \cdots, T\}$ and the log excess returns $\{z_{it}; i = 1, \cdots, N; t = 1, \cdots, T\}$.

1.2.1 Review of Statistical Distributions and Their Moments

We briefly review some basic properties of statistical distributions and the moment equations of a random variable. Let $R^k$ be the $k$-dimensional Euclidean space. A point in $R^k$ is denoted by $x \in R^k$. Consider two random vectors $X = (X_1, \cdots, X_k)'$ and $Y = (Y_1, \cdots, Y_q)'$. Let $P(X \in A, Y \in B)$ be the probabil-
ity that $X$ is in the subspace $A \subset \mathbb{R}^k$ and $Y$ is in the subspace $B \subset \mathbb{R}^q$. For most of the cases considered in this course, both random vectors are assumed to be continuous.

**Joint Distribution**

The function

$$F_{X,Y}(x, y; \theta) = P(X \leq x, Y \leq y),$$

where $x \in \mathbb{R}^p$, $y \in \mathbb{R}^q$, and the inequality “$\leq$” is a component-by-component operation, is a joint distribution function of $X$ and $Y$ with parameter $\theta$. Behavior of $X$ and $Y$ is characterized by $F_{X,Y}(x, y; \theta)$. If the joint probability density function $f_{x,y}(x, y; \theta)$ of $X$ and $Y$ exists, then

$$F_{X,Y}(x, y; \theta) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{x,y}(w, z; \theta) \, dz \, dw.$$ 

In this case, $X$ and $Y$ are continuous random vectors.

**Marginal Distribution**

The marginal distribution of $X$ is given by

$$F_X(x; \theta) = F_{X,Y}(x, \infty, \cdots, \infty; \theta).$$

Thus, the marginal distribution of $X$ is obtained by integrating out $Y$. A similar definition applies to the marginal distribution of $Y$. 
If \( k = 1 \), \( X \) is a scalar random variable and the distribution function becomes

\[
F_X(x) = P(X \leq x; \theta),
\]

which is known as the cumulative distribution function (CDF) of \( X \). The CDF of a random variable is nondecreasing [i.e., \( F_X(x_1) \leq F_X(x_2) \) if \( x_1 \leq x_2 \), and satisfies \( F_X(-\infty) = 0 \) and \( F_X(\infty) = 1 \)]. For a given probability \( p \), the smallest real number \( x_p \) such that \( p \leq F_X(x_p) \) is called the \( p \)-th quantile of the random variable \( X \). More specifically,

\[
x_p = \inf_x \{ x \mid p \leq F_X(x) \}.
\]

We use CDF to compute the \( p \) value of a test statistic in the course.

**Conditional Distribution**

The conditional distribution of \( X \) given \( Y \leq y \) is given by

\[
F_{X|Y \leq y}(x; \theta) = \frac{P(X \leq x, Y \leq y)}{P(Y \leq y)}.
\]

If the probability density functions involved exist, then the conditional density of \( X \) given \( Y = y \) is

\[
f_{x|y}(x; \theta) = \frac{f_{x,y}(x, y; \theta)}{f_y(y; \theta)}, \quad (2)
\]
where the marginal density function $f_y(y; \theta)$ is obtained by

$$f_y(y; \theta) = \int_{-\infty}^{\infty} f_{x,y}(x, y; \theta) dx.$$ 

From (2), the relation among joint, marginal, and conditional distributions is

$$f_{x,y}(x, y; \theta) = f_{x|y}(x; \theta) f_y(y; \theta).$$

This identity is used extensively in time series analysis (e.g., in maximum likelihood estimation). Finally, $X$ and $Y$ are independent random vectors if and only if $f_{x|y}(x; \theta) = f_x(x; \theta)$. In this case, $f_{x,y}(x, y; \theta) = f_x(x; \theta) f_y(y; \theta)$.

### Moments of a Random Variable

The $\ell$-th moment of a continuous random variable $X$ is defined as

$$m'_\ell = E(X^\ell) = \int_{-\infty}^{\infty} x^\ell f(x) dx,$$

where “$E$” stands for expectation and $f(x)$ is the probability density function of $X$. The first moment is called the mean or expectation of $X$. It measures the central location of the distribution. We denote the mean of $X$ by $\mu_x$. The $\ell$-th central moment of $X$ is defined as

$$m_\ell = E[(X - \mu_x)^\ell] = \int_{-\infty}^{\infty} (x - \mu_x)^\ell f(x) dx,$$
provided that the integral exists. The second central moment, denoted by $\sigma_x^2$, measures the variability of $X$ and is called the \textit{variance} of $X$. The positive square root, $\sigma_x$ of variance is the \textit{standard deviation} of $X$. The first two moments of a random variable uniquely determine a normal distribution. For other distributions, higher order moments are also of interest.

The third central moment measures the symmetry of $X$ with respect to its mean, whereas the 4th central moment measures the tail behavior of $X$. In statistics, \textit{skewness} and \textit{kurtosis}, which are normalized 3rd and 4th central moments of $X$, are often used to summarize the extent of asymmetry and tail thickness. Specifically, the skewness and kurtosis of $X$ are defined as

\[ S(x) = E \left[ \frac{(X - \mu_x)^3}{\sigma_x^3} \right], \quad K(x) = E \left[ \frac{(X - \mu_x)^4}{\sigma_x^4} \right]. \]

The quantity $K(x) - 3$ is called the \textit{excess kurtosis} because $K(x) = 3$ for a normal distribution. Thus, the excess kurtosis of a normal random variable is zero. A distribution with positive excess kurtosis is said to have heavy tails, implying that the distribution puts more mass on the tails of its support than a normal distribution does. In practice, this means that a random sample from such a distribution tends to contain more extreme values.

In application, skewness and kurtosis can be estimated by their
sample counterparts. Let \( \{x_1, \cdots, x_T\} \) be a random sample of \( X \) with \( T \) observations. The sample mean is
\[
\hat{\mu}_x = \frac{1}{T} \sum_{t=1}^{T} x_t,
\]
the sample variance is
\[
\hat{\sigma}_x^2 = \frac{1}{T - 1} \sum_{t=1}^{T} (x_t - \hat{\mu}_x)^2,
\]
the sample skewness is
\[
\hat{S}(x) = \frac{1}{(T - 1)\hat{\sigma}_x^3} \sum_{t=1}^{T} (x_t - \hat{\mu}_x)^3,
\]
and the sample kurtosis is
\[
\hat{K}(x) = \frac{1}{(T - 1)\hat{\sigma}_x^4} \sum_{t=1}^{T} (x_t - \hat{\mu}_x)^4.
\]
Under normality assumption, \( \hat{S}(x) \) and \( \hat{K}(x) \) are distributed asymptotically as normal with zero mean and variances \( 6/T \) and \( 24/T \), respectively.