Chapter 1: Market Indexes, Financial Time Series and their Characteristics

What is financial time series (FTS) analysis?

Theory and practice of asset valuation over time.

Different from other T.S. analysis?

Not exactly, but with an added uncertainty.

For example, FTS must deal with the changing business and economic environment and the fact that volatility is not directly observed.

Objective of the course

to provide some basic knowledge of financial time series data

to introduce some statistical tools and econometric models useful for analyzing these series.

to gain empirical experience in analyzing FTS

to study methods for assessing market risk

to analyze high-dimensional asset returns.

Examples of financial time series

1. Daily log returns of Hang Sang Index .
2. Monthly log return of exchange rates of Japan-USA.
3. China life daily stock data.
4. HSBC daily stock data.

1.1 **Asset Returns**

Let P_t be the price of an asset at time t, and assume no dividend. One-period simple return or simple net return:

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

Gross return

$$1 + R_t = \frac{P_t}{P_{t-1}} \quad or \quad P_t = P_{t-1}(1 + R_t).$$

Multi-period simple return or the $k-$period simple net return:

$$R_t(k) = \frac{P_t}{P_{t-k}} - 1.$$

Gross return

$$\begin{aligned} 1 + R_t(k) &= \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}} \\ &= (1 + R_t)(1 + R_{t-1}) \times \cdots \times (1 + R_{t-k+1}) \\ &= \prod_{j=0}^{k-1} (1 + R_{t-j}). \end{aligned}$$

Example: Suppose the daily closing prices of a stock are

Day	1	2	3	4	5
Price	37.84	38.49	37.12	37.60	36.30

1. What is the simple return from day 1 to day 2?

Ans: $R_2 = \frac{38.49 - 37.84}{37.84} = 0.017$.

2. What is the simple return from day 1 to day 5?

Ans: $R_5(4) = \frac{36.30 - 37.84}{37.84} = -0.041$.

3. Verify that

$1 + R_5(4) = (1 + R_2)(1 + R_3)(1 + R_4)(1 + R_5)$.

Time interval is important! Default is one year.

Annualized (average) return:

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

An approximation:

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

Continuous 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 m times a year, then the interest rate for each payment is $10\%/m$, and the net value of the deposit become

$$\$1 \times \left(1 + \frac{0.1}{m}\right)^m.$$

Illustration of the power of compounding (int. rate 10% per annum):

Type	m(payment)	Int.	Net
Annual	1	0.1	\$1.10000
Semi-Annual	2	0.05	\$1.10250
Quarterly	4	0.025	\$1.10381
Monthly	12	0.0083	\$1.10471
Weekly	52	0.1/52	\$1.10506
Daily	365	0.1/365	\$1.10516
Continuously	∞		\$1.10517

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

$$A = C \exp(r \times n),$$

r is the interest rate per annum, C is the initial capital, n is the number of years, and and exp is the exponential function.

Present value:

$$C = A \exp[-r \times n].$$

Continuously compounded (or log) return

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

where $p_t = \ln(P_t)$

Multi-period log return:

$$\begin{aligned} r_t(k) &= \ln[1 + R_t(k)] \\ &= \ln[(1 + R_t)(1 + R_{t-1})(1 + R_{t-k+1})] \\ &= \ln(1 + R_t) + \ln(1 + R_{t-1}) + \cdots + \ln(1 + R_{t-k+1}) \\ &= r_t + r_{t-1} + \cdots + r_{t-k+1}. \end{aligned}$$

Example (continued). Use the previous daily prices.

1. What is the log return from day 1 to day 2?

A: $r_2 = \ln(38.49) - \ln(37.84) = 0.017$.

2. What is the log return from day 1 to day 5?

A: $r_5(4) = \ln(36.3) - \ln(37.84) = -0.042$.

3. It is easy to verify $r_5(4) = r_2 + \cdots + r_5$.

Portfolio return: N assets

$$R_{p,t} = \sum_{i=1}^{N} w_i R_{it},$$

where R_{it} is the simple return of asset i.

Example: An investor holds stocks of IBM, Microsoft and Citi- Group. Assume that her capital allocation is 30%, 30% and 40%. The monthly simple returns of these three stocks are 1.42%, 3.37% and 2.20%, respectively. What is the mean simple return of her stock portfolio in percentage?

Answer:

$E(R_t) = 0.3 \times 1.42 + 0.3 \times 3.37 + 0.4 \times 2.20 = 2.32$.

The continuously compounded returns of a portfolio do not have the previous convenient property. When R_{it} is small in absolute value, we have

$$r_{p,t} \approx \sum_{i=1}^{N} w_i r_{it},$$

where r_{it} is the simple return of asset i.

Dividend payment: 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.

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

Excess return: (adjusting for risk)

$$Z_t = R_t - R_{0t}, z_t = r_t - r_{0t},$$

where r_{0t} denotes the log return of a reference asset (e.g. risk-free interest rate) such as short-term U.S. Treasury bill return, etc..

Relationship:

$$r_t = \ln(1 + R_t), R_t = e^{r_t} - 1.$$

If the returns are in percentage, then

$$r_t = 100 \times \ln(1 + \frac{R_t}{100}), R_t = [\exp(r_t/100) - 1] \times 100.$$

Temporal aggregation of the returns produces

$$1 + R_t(k) = (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}),$$

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

These two relations are important in practice, e.g. obtain annual returns from monthly returns.

Example: If the monthly log returns of an asset are 4.46%, -7.34% and 10.77%, then what is the corresponding quarterly log return?

Answer: (4.46 - 7.34 + 10.77)% = 7.89%.

Example: If the monthly simple returns of an asset are 4.46%, -7.34% and 10.77%, then what is the corresponding quarterly simple return?

Answer: $R = (1 + 0.0446)(1 - 0.0734)(1 + 0.1077) - 1 = 1.0721 - 1 = 0.0721 = 7.21\%$.

1.2 Distributional properties of returns

Is r_t a data or random variable?

What is the difference?

Key: What is the distribution of

$$(r_t: \ t = 1, \cdots, T)?$$

Review of theoretical statistics:

Moments of a random variable X with density $f(x)$: $l-$th moment

$$m_l' = E(X^l) = \int_{-\infty}^{\infty} x^l f(x) dx.$$

First moment: mean or expectation of X. $l-$th central moment

$$m_l = E(X - \mu_x)^l = \int_{-\infty}^{\infty} (x - \mu_x)^l f(x) dx,$$

2nd c.m.: $\sigma_x^2 = E(X - \mu_x)^2$— Variance of X.

Skewness (symmetry) and kurtosis (fat-tails)

$$S(x) = E\frac{(X - \mu_x)^3}{\sigma_x^3}, \ K(x) = E\frac{(X - \mu_x)^4}{\sigma_x^4}.$$

$K(x) - 3$: Excess kurtosis.

Why are mean and variance of returns important?

They are concerned with long-term return and risk, respectively.

Why is symmetry of interest in financial study?

Symmetry has important implications in holding short or long financial positions and in risk management.

Why is kurtosis important?

Related to volatility forecasting, efficiency in estimation and tests, etc.

High kurtosis implies heavy (or long) tails in distribution.

Estimation:

Data: $\{x_1, \cdots, x_T\}$.

sample mean:

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

sample variance:

$$\hat{\sigma}_x^2 = \frac{1}{T-1}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^2,$$

sample skewness:

$$\hat{S}(x) = \frac{1}{(T-1)\hat{\sigma}_x^3}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^3,$$

sample kurtosis:

$$\hat{K}(x) = \frac{1}{(T-1)\hat{\sigma}_x^4}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^4.$$

Random sample: $\{x_1, \cdots, x_T\}$.

$\hat{\mu}_x$, $\hat{\sigma}_x^2$, $\hat{S}(x)$ and $\hat{K}(x)$ are random.

Under normality assumption,

$$\hat{S}(x) \sim N(0, \frac{6}{T}), \;\; \hat{K}(x) - 3 \sim N(0, \frac{24}{T}).$$

Some simple tests for normality (for large T).

1. Test for symmetry:

$$S^* = \frac{\hat{S}(x)}{\sqrt{6/T}} \sim N(0,1)$$

if normality holds.

Decision rule: Reject H_0 of a symmetric distribution if $|S^*| > Z_{\alpha/2}$ or p-value is less than α.

2. Test for tail thickness:

$$K^* = \frac{\hat{K}(x) - 3}{\sqrt{24/T}} \sim N(0,1)$$

if normality holds.

Decision rule: Reject H_0 of normal tails if $|K^*| > Z_{\alpha/2}$ or p-value is less than α

Goodness-of-Fit Tests in SAS

The empirical distribution function is defined for a set of n independent observations $r_1, \cdots, r_n$ with a common distribution function F(x). Denote the observations ordered from smallest to largest as $r_{(1)}, ..., r_{(n)}$. The empirical distribution function, $F_n(x)$, is defined as

$$F_n(x) = \begin{cases} 0 & \text{if } x < r_{(1)} \\ \frac{i}{n} & \text{if } r_{(i)} \le x < r_{i+1},\ i = 1, \cdots, n-1 \\ 1 & \text{if } x \ge r_{(n)}. \end{cases}$$

PROC UNIVARIATE provides three EDF tests:

Kolmogorov-Smirnov (D)

Anderson-Darling ($A - sq$)

Cramr-von Mises ($W - sq$)

Kolmogorov D Statistic:

$$D_n = \sup_x |F_n(x) - F(x)|.$$

1.3 Main objective of TS analysis:

Past data $\Longrightarrow$TS r.v. $Z_t \Longrightarrow$ future of TS.

$$\begin{array}{ll} (a) & E\left(Z_{n+l} | Z_1, \cdots, Z_n\right), \\ (b) & P(a \le Z_{n+l} \le b | Z_1, \cdots, Z_n) \text{ for some } a < b. \end{array}$$