6. The econometrics of Financial Markets:

Empirical Analysis of Financial Time Series

MA6622, Ernesto Mordecki, CityU, HK, 2006.

References for Lecture 5:

Quantitative Risk Management. A. McNeil, R. Frey, P. Embrechts, Princeton University Press (2005)

Purpose: Given a stock price, index or exchange-rate series

$$S(0), S(1), \ldots, S(n),$$

find an adequate statistical model for the stochastic process of returns

$$X(1) = \log \frac{S(1)}{S(0)}, \ldots, X(n) = \log \frac{S(n)}{S(n-1)}.$$

Main interest: predict future values, and estimate the risk involved in holding the asset.

We assume that our data has no seasonal components, and has no trend, i.e the usual procedures of data transformation have been used in case of need. We also assume, in principle, that the data is stationary.

Depending on the time intervals between data, one distinguishes:

- Large time intervals: weekly, monthly, quarterly, or yearly data. Methods developed in the 1970s and before use linear models (ARMA, ARIMA).
- Daily financial data: Main methods, developed in the 1980s, are nonlinear: ARCH, GARCH, etc. models.
- Intraday data, or tick data, beginning in the 1990s, can reach intevals of seconds. Also named high-frequency data. New statistical phenomena appears.

We concentrate our analysis on the first two types of data, with emphasis on the second: daily data.

One must also distinguish between the type of financial data. For instance: Forex data is continuosly quoted (including week-ends), while stock prices are available only on trading days and hours.

6a. Stylized Facts of Financial time series.

Empirical observations on daily returns of financial time series led to the following 6 stylized facts, widely understood to be empirical truths, to which theories must fit.

(1) Return series show little serial correlation.

This supports the random walk hypothesis, as increments of a random walk are independent, but...

(2) Series of absolute or squared returns show profound serial correlation.

This contradicts the random walk hyphotesis, as if X, Y are independent, then $|X|, |Y|$ are independent, and X^2, Y^2 are also independent, and in consequence they should be non-correlated. The conclussion is that financial time series are uncorrelated but not independent.

(3) Conditional expected returns are close to zero.

More problems: this gives that the daily returns form a martingale difference, i.e. the accumulated process (sum of daily returns) is a martingale. This supports the martingale hypothesis, so it is very difficult to predict future values, based on historical data.

(4) Volatility appears to vary over time.

Defining the volatility as the conditional standard deviation of the returns given the past information, it is observed that if recent returns have been large, it is expected to have large returns.

(5) Return series are leptokurtic or heavy-tailed.

The kurtosis of a random variable X is defined as

$$\kappa_X = \frac{\mathbf{E}\left(X - \mathbf{E}\,X\right)^4}{(\mathbf{var}\,X)^2} - 3,$$

and $3 = \frac{\mathbf{E}(Z-\mu)^4}{\sigma^4}$ for a gaussian random variable $Z \sim N(\mu, \sigma^2)$.

A random variable with positive kurtosis is named leptokurtic.

It models a large amount of small movements, plus some relatively large movements, in other terms, its density is more narrow around its expectation, and has heavier tails than a gaussian random variable.

This contradicts the gaussian returns hypothesis in the Black-Scholes model.

(6) Extreme returns apperar in clusters: volatility clustering phenomena

Is a tendency for large returns to be followed by large returns (positive or negative, as they are uncorrelated)

Remarks:

- The longer the time intervals considered, the less pronunciated the reported facts appear.
- For intraday data new statistical phenomena, not discussed here, appears.

6b. Notations and definitions for Time Series

- $\mu(t) = \mathbf{E}\, X(t)$ is the expectation of the t-th return,
- $\mathbf{cov}(s,t) = \mathbf{E}\big(X(s) - \mu(s)\big)\big(X(t) - \mu(t)\big)$ is the covariance between the returns at times s and t,
- $\mathbf{var}(t) = \mathbf{cov}(t,t)$ is the variance of the t-th return.
- $\rho(s,t) = \dfrac{\mathbf{cov}(s,t)}{\sqrt{\mathbf{var}(s)\,\mathbf{var}(t)}}$ is the correlation between the returns at times s and t.

Definition A process $\{X(t)\}$ is weakly stationary when

- $\mu(t) \equiv \mu$, for all t,
- $\mathbf{cov}(s+k, t+k) = \mathbf{cov}(s,t)$ for all s, t, k.

In this case the autocovariance function $\mathbf{cov}(s,t)$ depends only the difference $|s-t|$ (take $k=-t$), the variance is constant (as $t-t=0$) and consequently the autocorrelation function $\rho(s,t)$ also depends on the differnce $|s-t|$. This is why we write

$$\rho(h)=\rho(h,0)=\frac{\mathbf{cov}(h,0)}{\mathbf{var}(0)}.$$

The autocorrelation $\rho(h)$ is also termed as serial correlation at lag h.

Definition We say that $\{X(t)\}$ is a weak white noise process if it is weakly stationary with $\mu=0$, and autocorrelation function

$$\rho(h)=\begin{cases}1, & \text{when } h=0\\ 0, & \text{when } h\neq 0\end{cases}$$

Definition A process $\{X(t)\}$ is strictly stationary when

$$\mathbf{P}\left(X(1) \in I_1, \ldots, X(t) \in I_h\right) = \mathbf{P}\left(X(1+h) \in I_1, \ldots, X(t+h) \in I_h\right),$$

for all t, h. That is, the probability distribution of the vector

$$(X(1), \ldots, X(t))$$

is invariant under translations of the time.

Some Theoretical Facts:

- It can be proved that a strictly stationary process with finite variance is weakly stationary.
- There exists then strictly stationary processes that are not weakly stationary.
- In general, a weakly stationary process is not strictly stationary.

Definition We say that $\{X(t)\}$ is a strict white noise process if the random variables

- are centered, i.e. $\mathbf{E}\, X(t) = 0$ for all t
- are independent
- are identically distributed
- have finite variance $\sigma_X^2 = \mathbf{E}\left(X(t)^2\right)$.

Definition We say that $\{X(t)\}$ is a normal or gaussian white noise process if the random variables

- form a strict white noise,
- have gaussian distribution with parameters $(0, \sigma^2)$.

6c. Plotting the Empirical Correlogram

We now assume that $\{X(t)\}$ is weakly stationary, and compute the correlogram, from the observed data in order to make inference about the serial dependence structure of the process, i.e. try to find an adequate model for the returns.

STEP 1. We compute the sample mean

$$\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X(t).$$

STEP 2. We compute the sample variance

$$\bar{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n} \big(X(t) - \bar{X}\big)^2,$$

STEP 3. We compute the sample autocovariances at different lags $h = 1, 2, \ldots, n_0$

$$\overline{\mathbf{cov}}(h) = \frac{1}{n} \sum_{t=1}^{n-h} \big(X(t+h) - \bar{X}\big)\big(X(t) - \bar{X}\big),$$

n_0 should be significatively smaller than n, in daily data it is desirable to have $n_0 = 30$.

STEP 4. We compute the serial correlations at lags $h = 1, 2, \ldots, n/2$

$$\bar{\rho}(h) = \frac{\overline{\mathbf{cov}}(h)}{\bar{\sigma}^2}$$

STEP 5. We plot the map $(h, \bar{\rho}(h))$, called the correlogram.

6d. Testing for white noise

Purpose: Given a time series,

$$\varepsilon(0), \varepsilon(1), \ldots, \varepsilon(n).$$

that is either the result of observations, or was obtained as residuals in a statistical procedure, we study methods to determine whether the series can be modelled by a white noise.

We use two complementary approachs:

- Visual analysis of the correlograms
- A Statistical test

Visual Analysis

The visual analysis consist in plotting the correspondent correlogram, and compare it with the white noise correlogram.

Remember that the withe noise correlogram has the values

$$\begin{cases} \rho(0) = 1 \\ \rho(h) = 0 \quad \text{for all } h \neq 0 \end{cases}$$

An important result in order to give statistical sustent to the visual analysis is the following

Theorem [KEY] Assume that $\{X(t)\}$ is an ARMA(p,q) centered time series driven by a strict white noise.

Then, for big values of n and fixed h, the estimated correlation vector

$$\big(\bar{\rho}(1), \ldots, \bar{\rho}(h)\big)$$

is approximately a normal random vector with mean value vector

$$\big(\rho(1),\ldots,\rho(h)\big)$$

i.e. the true correlations, and variance covariance matrix $\mathbf{W}/n = ((W_{ij}/n))_{i,j=1,\ldots,h}$ given by Bartlett's formula:

$$W_{ij} = \sum_{k=1}^{\infty} \big([\rho(k+i) + \rho(k-i) - 2\,\rho(i)\,\rho(j)] \times [\rho(k+j) + \rho(k-j) - 2\,\rho(i)\,\rho(j)] \big)$$

In particular, for the variances, we have:

$$W_{ii} = \sum_{k=1}^{\infty} [\rho(k+i) + \rho(k-i) - 2\,\rho(i)^2]^2$$

In several cases it is simple to compute W_{ij} and we have a way to construct confidence intervals for our estimated correlations.

For instance, if we have a white noise, it is direct to see that

$$W_{ij} = \delta_{ij} = \begin{cases} 1 & \text{when } i = j \\ 0 & \text{when } i \neq j \end{cases}$$

Then under the null hypothesis of the sequence

$$\varepsilon(0), \ldots, \varepsilon(n)$$

being a strict white noise, the estimated correlations are (for big values of n) approximately normally distributed, centered, and with variance $1/n$.

In other terms, the random vector

$$\big(\bar{\rho}(1), \ldots, \bar{\rho}(h)\big) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_h/n),$$

where the second term is a centered gaussian vector with covariance matrix equal to the $h \times h$ identity matrix $\mathbf{I}_h$ divided by n.

This means that the coordinates of the vector $\sqrt{n}\bar{\rho}$ are independent standard normal random variables.

This implies that 95% of the estimated correlation values should lie in the interval

$$(-1.96/\sqrt{n}, 1.96/\sqrt{n}),$$

and this is the reason why correlograms shoud indicate these two levels.

If more that 5% of the estimated correlations lie outside of this bound, it is an evidence against the null hypothesis that the data are strict white noise.

Statistical Test

A popular Pormanteau Test was proposed by Ljung and Box (1978). It uses the statistic

$$Q_{LB} = n(n+2) \sum_{j=1}^{h} \frac{\bar{\rho}(j)}{n-j}$$

This statistic has (for large values of n) a

- χ^2_h distribution when directly observing the data,
- χ^2_{h-p-q} distribution when the residuals have been obtained after the estimation of an ARMA(p,q) process.

i.e. a Chi squared distribution with h or $h-p-q$ degrees of freedom.

Small values of Q_{LB} indicate that there is no statistical evidence to reject the null hypothesis of $\{\varepsilon(t)\}$ being a strict white noise. In consequence, the test is preformed constructing a critical region of the form

$$Q_{LB} > t_{1-\alpha,h}$$

where $t_{1-\alpha,h}$ satisfies the condition

$$\mathbf{P}\left(J_h^2 > t_{1-\alpha,h}\right) = 1 - \alpha,$$

where J_h^2 is a random variable with χ_h^2 distribution.

For instance, if $\alpha = 0.05$ and $h = 10$, we get from the statistical tables

$$t_{0.95,10} = 18.30704$$

For the same α with $h = 30$ we have

$$t_{0.95,30} = 43.77297$$