Introduction

Statistical Methods in Finance

Price, returns, and other terminologies

- Price

$$P_t - \text{price of the asset at time t}$$

- Return (single-period net return)

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

- k-period (net) return

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

Price, returns, and other terminologies

- Gross return (single-period)

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

- k-period (gross) return

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

Multiplicative relationship of gross returns

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

Statistical models

- Single asset models (longitudinal)
  1. R_t, $t = 1, 2, \ldots$ are iid (μ, σ^2).
  2. R_t, $t = 1, 2, \ldots$ are iid $N(\mu, \sigma^2)$.
  3. $\{R_t\}$ is a time series, e.g. AR, MA, GARCH.
- Log returns

  It is sometimes mathematically convenient to consider the logarithm of (gross) return

  $$\tilde{R}_t = log(1 + R_t).$$

  It follows that the k-period gross return has the following expression

  $$1 + R_t(k) = \exp\{\tilde{R}_t + \tilde{R}_{t-1} + \cdots + \tilde{R}_{t-k+1}\}.$$

Statistical models

- Joint modeling of n asset returns
  For a market with n assets, denote their returns by $R_{i,t}, i = 1, 2, \cdots, n, t = 1, 2, \cdots$. We need both cross sectional and longitudinal modeling of the returns to capture the dynamics of the market behavior over time. For example, we may assume that $\mathbf{R}_t = (R_{1,t}, R_{2,t} ..., R_{n,t}), t = 1, ..., m$ are iid random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ.
- Adjusting for dividends
  Assets like stocks sometimes pay dividends. Let D_t denote the dividend paid by the asset at time t. Then the return becomes

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

- Example 1 (Wheel of fortune; Luenberger, 98)

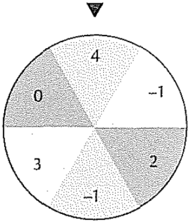

Figure: **Wheel of fortune**. If you bet \$1 on the wheel, you will receive the amount equal to the value shown in the segment under the marker after the wheel is spun

- Example 2 (IBM stock price data)
  The data are daily adjusted closing prices of the IBM stock from Jan 02 to Dec 31, 2013 (downloaded from YAHOO Finance).

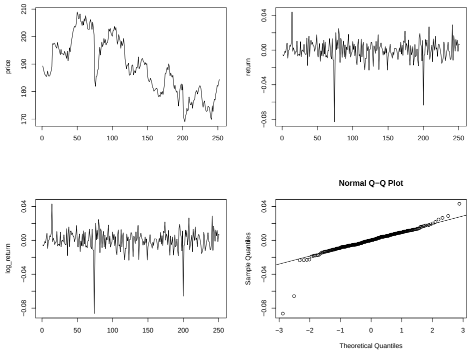

Figure: IBM price, return, and log return

Examples

Checking for normality.

- normal Q-Q plot
- Kolmogorov Smirnov test. Output: $D = 0.088$, p-value $= 0.040$
- Shapiro-Wilk test. Output: $W = 0.859$, p-value $= 2.279\text{e-}14$
- The p-values are less than 0.05; we reject the null hypothesis of normality.

Continuous time formulation and geometric Brownian motion

The discrete time asset prices may be viewed as observed values of an underlying continuous time process on an equally spaced grid: $t = 0, \Delta, 2\Delta, \cdots$. Thus,

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

Taking the infinitesimal limit ($\Delta \to 0$), with $\Delta X_t = R_t(\Delta)$, we have

$$dX_t = \frac{dP_t}{P_t}.$$

Suppose X_t is modeled as a Wiener process, i.e. $X_t = \mu t + \sigma B_t$, where B_t is the Brownian motion (standard Wiener process). Letting P_0 denote the initial asset price, we can solve (Ito's formula) for P_t to get

$$P_t = P_0 \exp\{X_t - \frac{1}{2}\sigma^2 t\},$$

which is known as the geometric Brownian motion.