Chapter 1

Asset returns

1.1   Returns

Much of financial econometrics concerns about dynamics of the returns and their associated volatilities of assets. Let S_t be the price of an asset (portfolio) at time t.

One-period simple return: $R_t = (S_t - S_{t-1})/S_{t-1}$.

k-period simple return: $R_t[k] = S_t/S_{t-k} - 1$. Note that

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

log-return: $r_t = \log(S_t/S_{t-1}) = \log S_t - \log S_{t-1}$. It is easy to see

$$r_t = \log(1 + R_t) \approx R_t.$$

Note that $r_t \leq R_t$. The k-period log-return is defined as $r_t[k] = \log(S_t/S_{t-k})$, which satisfies

$$r_t[k] = r_t + r_{t-1} + \cdots + r_{t-k+1}.$$

Most of the returns in the class refer to the log-returns.

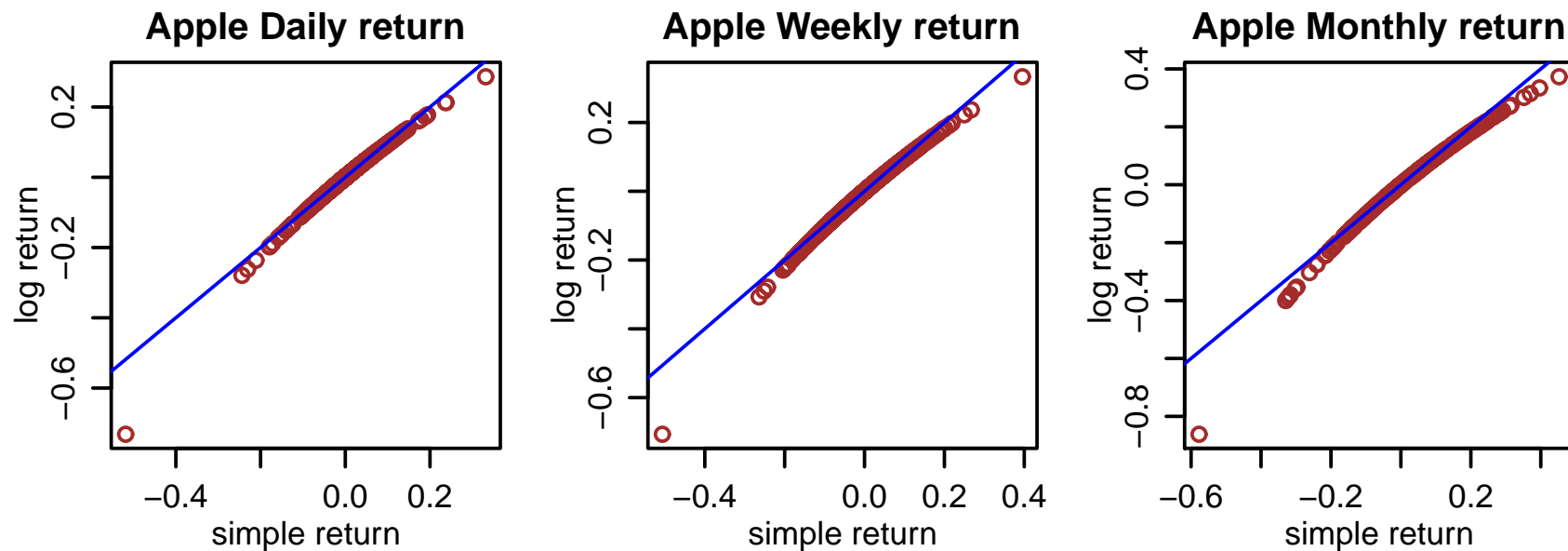**Differences**: are small for simple- and log returns (Fig 1.1).

Figure 1.1: Plots of log returns against simple returns of the Apple Inc share prices in January 1985 – February 2011. The blue straight lines mark the positions where the two returns are identical.

Returns with dividend payments:

$$R_t = \frac{S_t + D_t}{S_{t-1}} - 1; \qquad r_t = \log \frac{S_t + D_t}{S_{t-1}},$$

where D_t is the dividend payment during the period.

Continuously compounding: With initial capital C_0,

$$C_t = C_0 \exp(rt) = C_0 \lim_{m \to \infty} \left(1 + \frac{r}{m}\right)^{mt}.$$

Bond Yields: Bonds are quoted in annualized yields. A zero-coupon bond with yield r_t and duration D has a bond price $B_t = \exp(-r_t D)$. The annualized log-return of the bond is

$$\log(B_t/B_{t-1}) = \log(\exp(-r_t D + r_{t-1} D)) = D(r_{t-1} - r_t)$$

★approx the same as the change of the yield spread times duration

Example 1: We have two baskets of high-yield bonds and investment-grade bonds with average duration 4.4 years. Their yields spread over the Treasury bond with similar maturity are quoted and plotted in Fig 1.2. The daily returns of bonds can then be deduced.
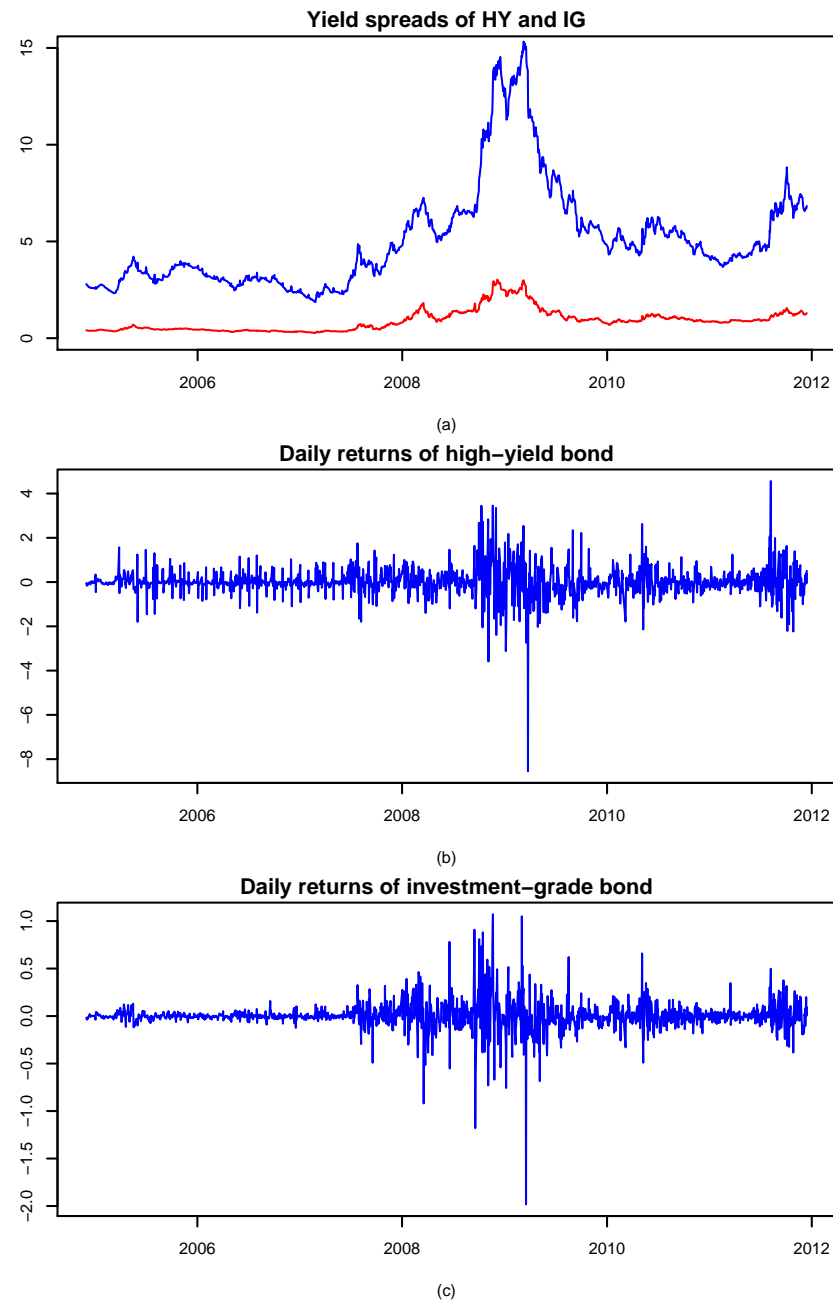
Figure 1.2: Time series plot of the yield spreads of high-yield bonds and investment-grade bonds and their associated daily returns from Nov. 29, 2004 to Dec. 14, 2011.

★ HY has higher yields than IG, but has higher volatility too ($\approx$ 3x).

★ Yield spread widen significantly after financial crisis.

Excess returns: difference between the asset's return and the return on some reference assets. Reference assets include

— **riskless**: 3-month US Treasury bills as a proxy;

— **London interbank offered rate**: LIBOR rate

— **market portfolio**: SP500 index or CRSP (center for research in security prices) value-weighted index as a proxy.

— **Yield spread**: yield differences of a bond over a reference bond (e.g. the US treasury) with similar maturity.

1.2   Behavior of financial returns

Stylized features:

★ **Random Walks** with some drifts (Fig. 1.3-1.4).

★ **Volatility clustering**: large price changes in clusters. Increase with time horizon. (Fig. 1.3-1.4). Time-varying volatility (conditional) can be seen. The SD of SP500 returns in 11/29/04–12/31/07 is 0.78%, and afterwards (01/02/08–02/28/11) 1.83%. The volatility in 05-06 is 0.64%, 3.44% at the height of the 2008 financial crisis (8/15/08-3/16/09).
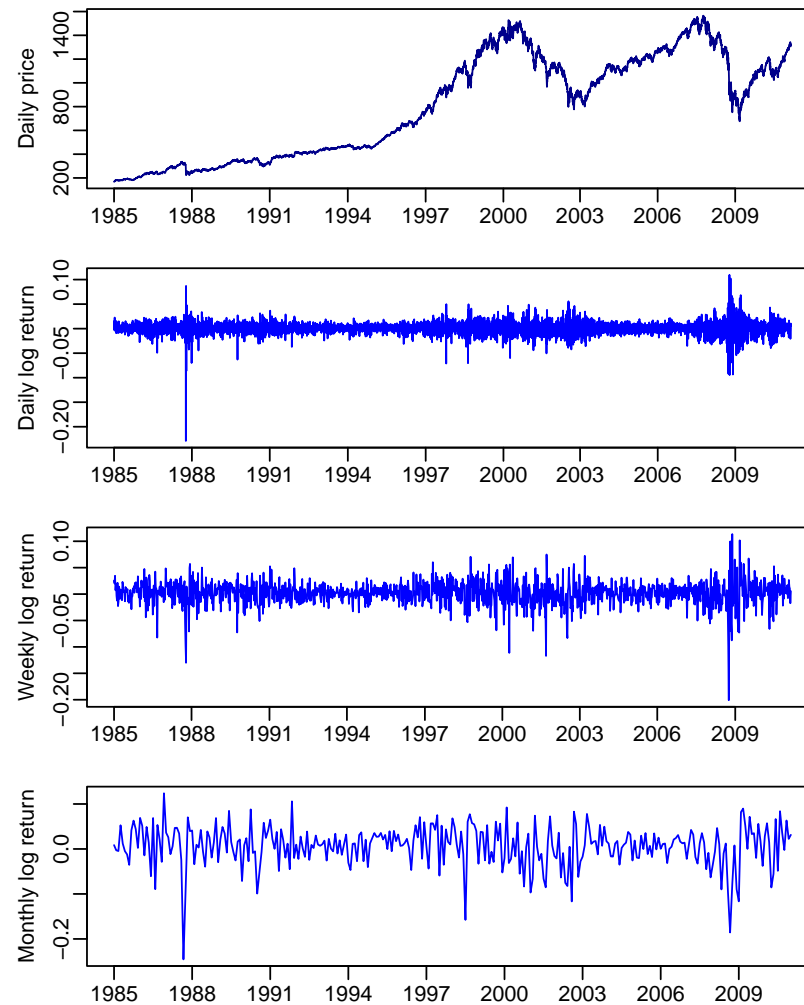
Figure 1.3: The time series plots of the daily indices, the daily log returns, the weekly log returns, and the monthly log returns of the **S&P500 index in January 1985 – February 2011**.
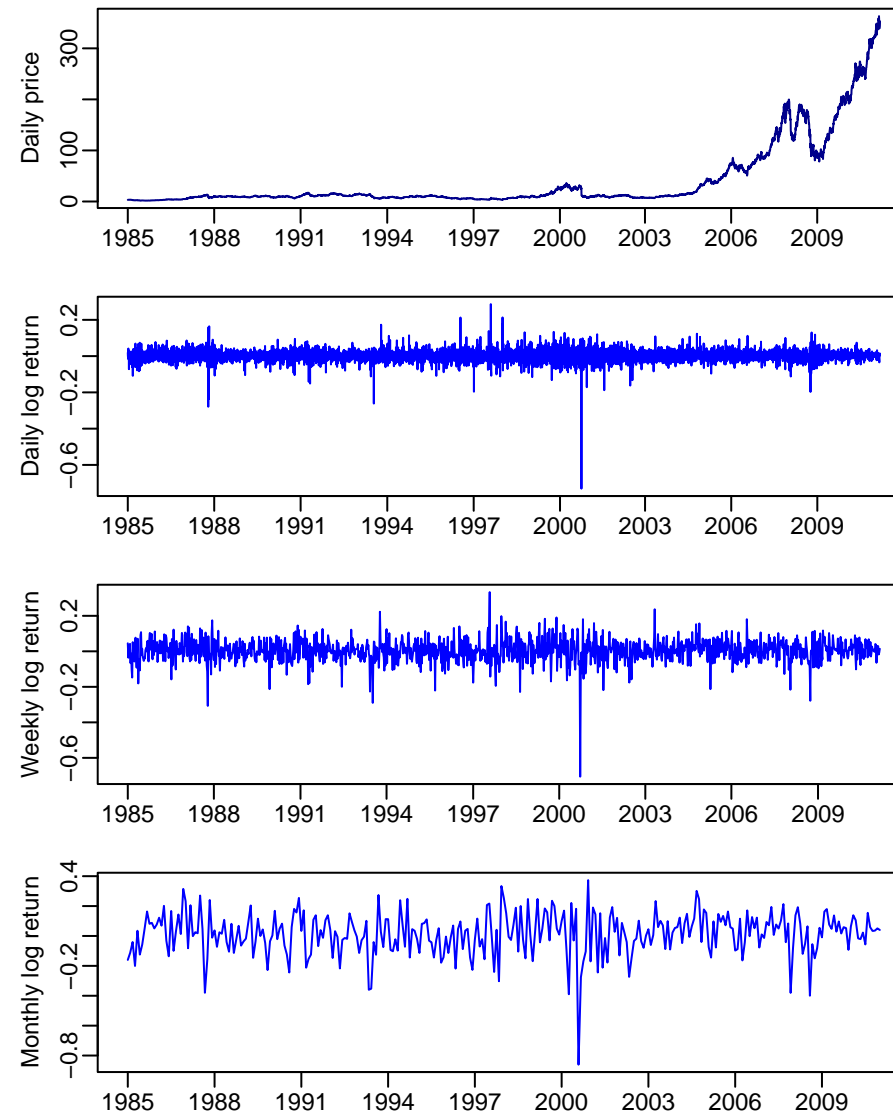
Figure 1.4: The time series plots of the daily prices, the daily log returns, the weekly log returns, and the monthly log returns of the **Apple stock in January 1985 – February 2011**.
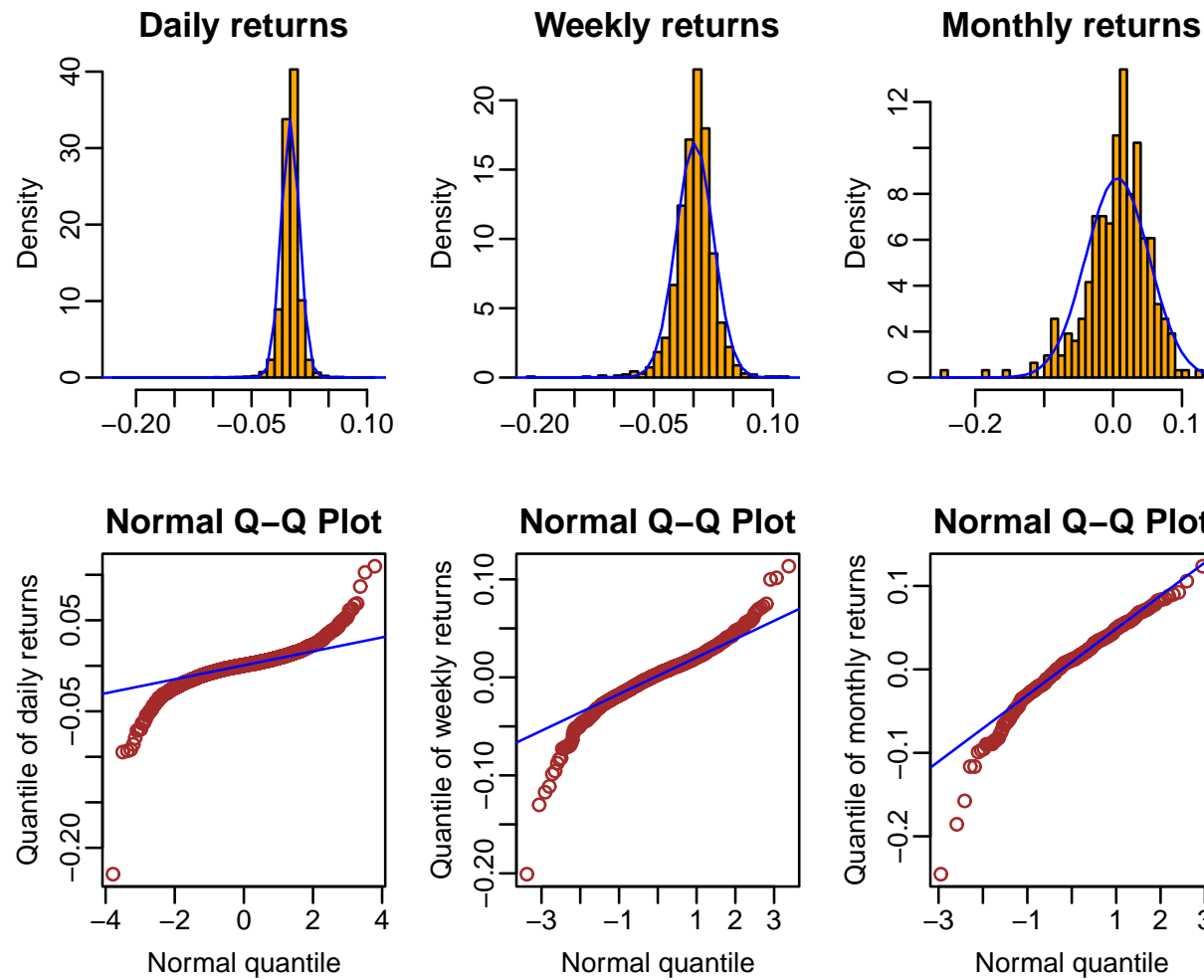
Figure 1.5: Histograms (the top panel) and Q-Q plots (the bottom panel) of the daily, weekly and monthly log returns of the **S&P500 index in January 1985 – February 2011**. The normal density with the same mean and variance are superimposed on the histogram plots.
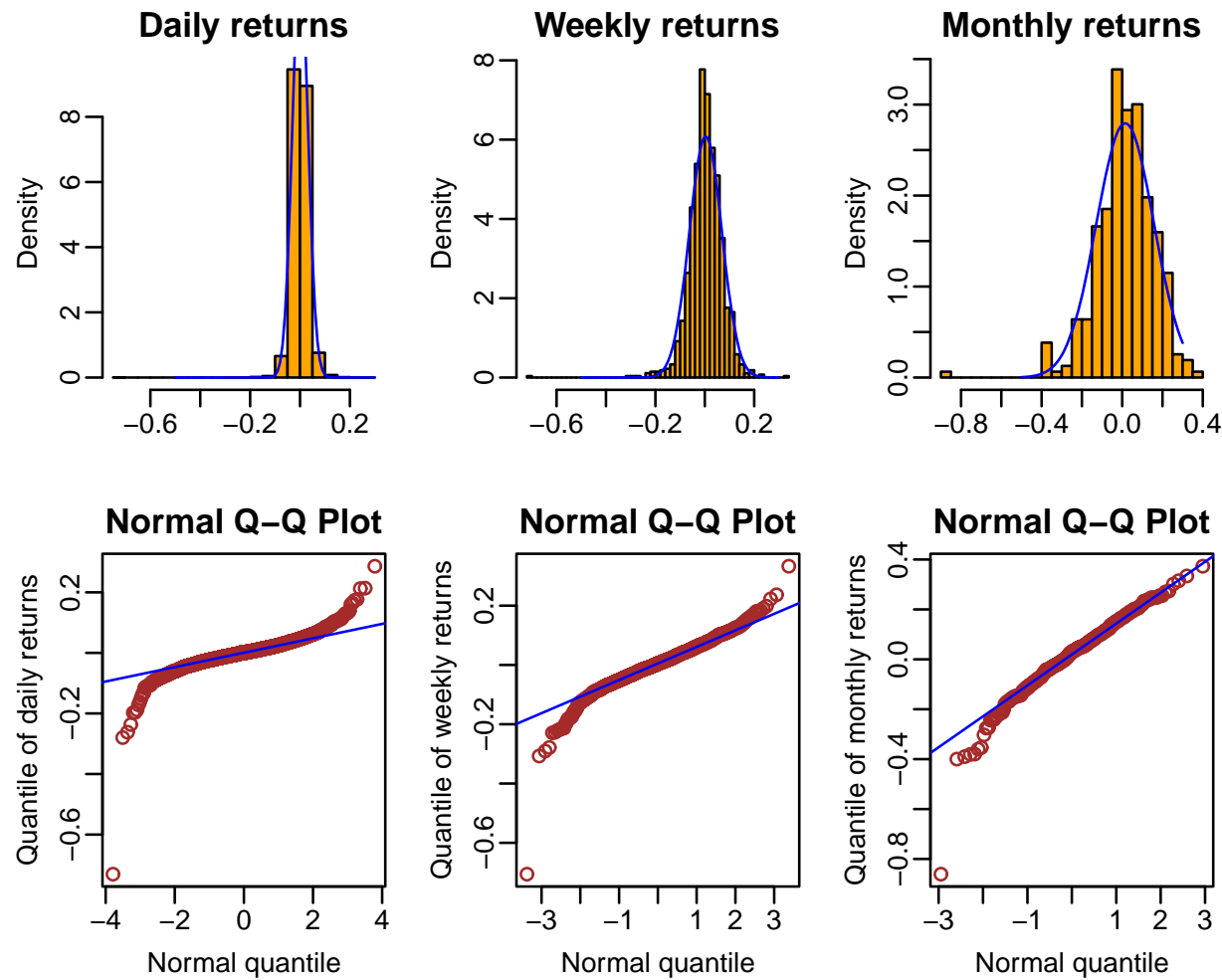
Figure 1.6: Histograms (the top panel) and Q-Q plots (the bottom panel) of the daily, weekly and monthly log returns of the **Apple stock in January 1985 – February 2011**. The normal density with the same mean and variance are superimposed on the histogram plots.

★ **Heavy tails** (heavier than normal, finite moments; Fig. 1.5-1.9). Modeled by a t-distribution with degree of freedom ν that has density

$$f_\nu(x) = d_\nu^{-1}\left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2} \asymp |x|^{-(\nu+1)}, \text{ as } |x| \to \infty$$

where $d_\nu = B(0.5, 0.5\nu)\sqrt{\nu}$. t_ν-distribution has mean 0 and variance $\nu/(\nu-2)$.

★t-dist. with d.f. 10 is indistinguishable from normal based on 10-years daily data using the Kolmogrov-Smirov test (`ks.test` in R)[1]. However their tail behaviors are very different. A 5%-event under normal distribution occurs in 14000 years ($= 1/pnorm(-5)/252$)and

---

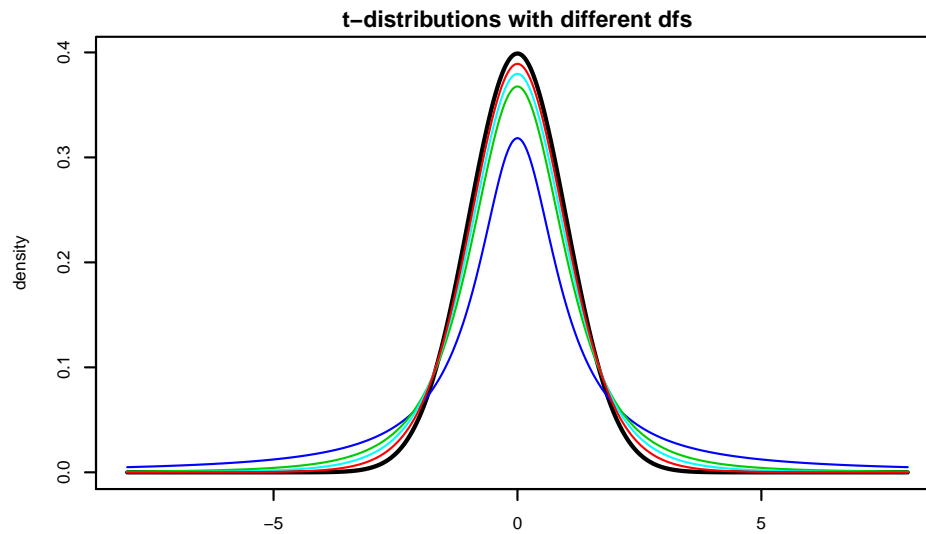[1]ks.test(rt(2520,10), "pnorm") gives P-values 0.1616, 0.0705, 0.0568, 0.1838, 0.3303

Figure 1.7: Left: t-distributions with various degrees of freedom. $t_\infty = N(0,1)$. Right: A black swan event means rare but devastating.

under t_{10}-distribution in 15 years and under $t_{4.5}$ is 1.5 years

$$(= 1/pt(-5, 4.5)/252).\ ^2$$

Quantile-Quantile plot: Theoretical quantiles plot against empirical quantiles to compare the tails. Let F be the target (refer-

---

[2]They have different standard deviations. The unconditional SD of S&P 500 is approximately 1%. Thus, its returns should be modeled as $\sqrt{(\nu - 2)/\nu} t_\nu$ in order to have the same SD. Thus, a -5%-event has the probability $P(\sqrt{(\nu - 2)/\nu} t_\nu < -5) = P\{t_\nu < -5\sqrt{\nu/(\nu - 2)}\}$. Thus, a -5-SD event under t_{10}-distribution is $-5 * \sqrt{10/8} = -5.59\%$ and occurs once every 34.4 years (1/pt(-5.59,10)/252). Similarly, a 5-SD event under $t_{4.5}$ is $-5 * \sqrt{4.5/2.5} = -6.71\%$ and happens once $1/pt(-6.71, 4.5)/252 = -4.76$ years.
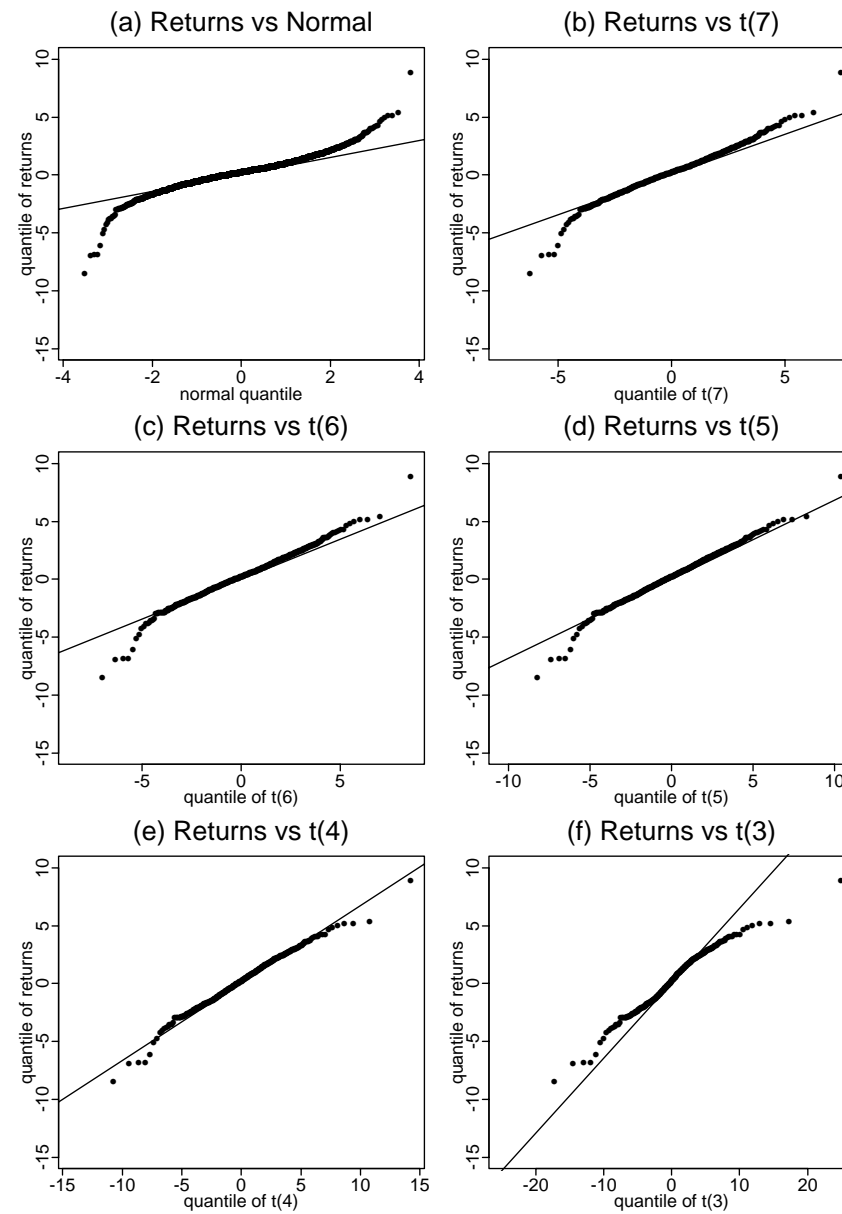
Figure 1.8: Distribution of the returns of **S&P500 from Jan. 1972 to Dec. 1999** vs t-distribution.
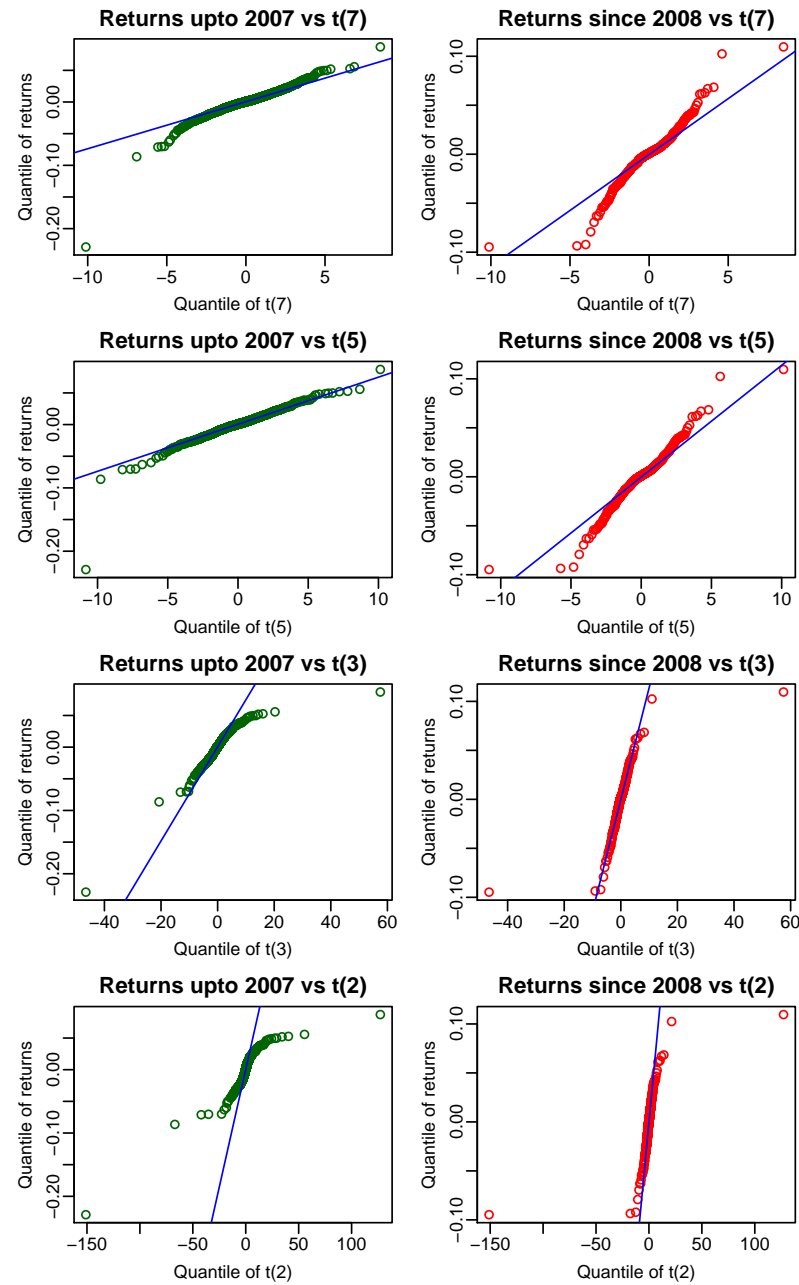
Figure 1.9: Quantiles of the returns of the **SP500 in January 1985 − February 2011** against the quantiles of the *t*-distributions with the degrees of freedom between 7 and 2.

ence) distribution and $\{x_{(i)}\}_{i=1}^n$ be the ordered data. Plot

$$\left\{\left(F^{-1}\left(\frac{i-0.5}{n}\right), x_{(i)}\right)\right\}.$$

<u>Jarque-Bera test</u>: The JB-test against normality of data $\{x_i\}_{i=1}^n$

is a statistic based on skewness and kurtosis

$$JB = \frac{n}{6}[S^2 + (K-3)^2/4] \overset{a}{\sim}_{H_0} \chi_2^2$$

under the normality assumption, where

$\blacksquare S = \frac{\widehat{\mu}_3}{\widehat{\sigma}^3} = \frac{\sum_{i=1}^n(x_i-\bar{x})^3/n}{(\sum_{i=1}^n(x_i-\bar{x})^2/n)^{3/2}},$ $\qquad \blacksquare K = \frac{\widehat{\mu}_4}{\widehat{\sigma}^4} = \frac{\sum_{i=1}^n(x_i-\bar{x})^4/n}{(\sum_{i=1}^n(x_i-\bar{x})^2/n)^2}.$

★ Asymmetry: The distributions of returns are negatively skewed. $(Er_t^3 < 0$, Fig. 1.5-1.9).

★ Aggregational Gaussianity: The returns over a long-time horizon

tend to a normal distribution (Fig. 1.5–1.6)

$$r_t[k] = r_t + r_{t-1} + \cdots + r_{t-k+1}, \quad r_t = \log \frac{S_t}{S_{t-1}}.$$

★ Leverage effect: asset returns are negatively correlated their volatilities (Black 1976, Christie 1982). (As asset prices declines, companies become more leveraged and riskier, and hence more volatile. Volatilities caused by price decline are larger than the appreciation due to declined volatilities.) See Fig 1.10.

VIX: A proxy of the implied volatility of a basket of at-money SP500 options maturity in one month.

★ Dependence. Both squared and absolute returns exhibit non-vanishing autocorrelations (Fig 1.11 and 1.12).
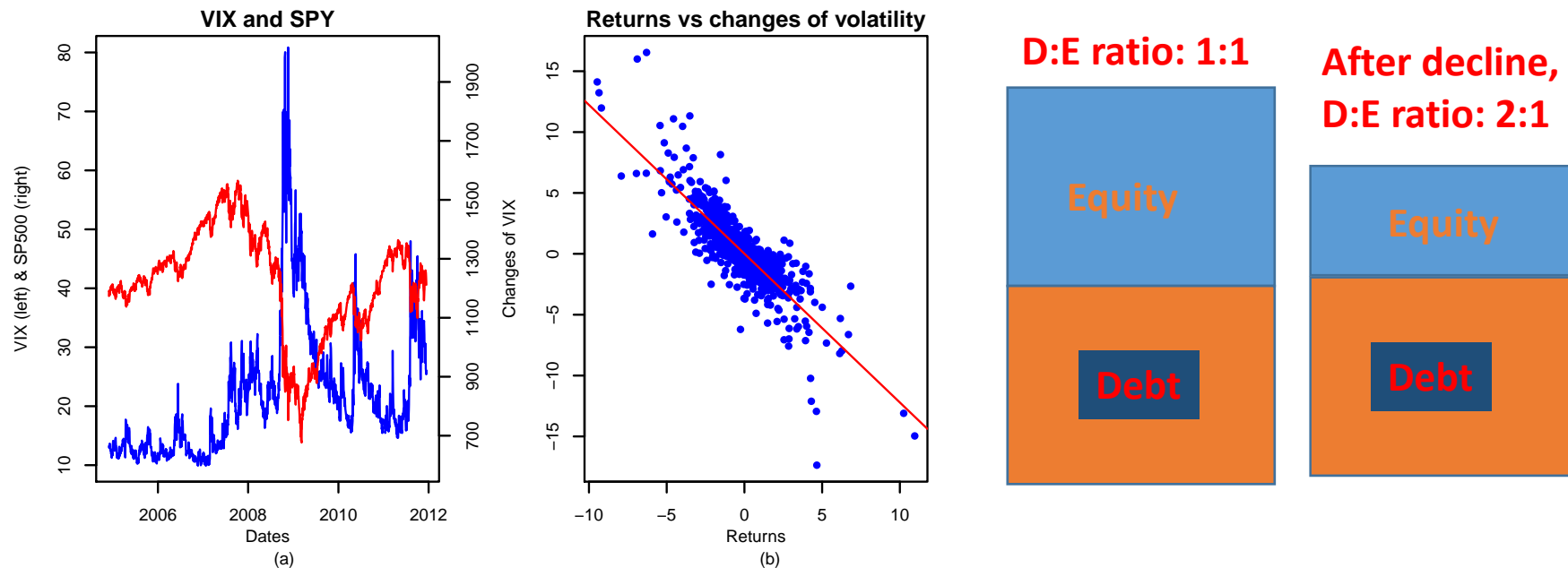
Figure 1.10: VIX, S&P500 index, and Leverage Effect. Data: Nov. 29, 2004 – Dec. 14, 2011.

1.3 Stationarity

Time series techniques are frequently employed to model the returns of financial assets. An important concept is the stationarity, which is about **structural invariability** across time so that historical relationship can be aggregated.

Figure 1.11: Autocorrelations of the daily, weekly and monthly log returns, the squared daily, weekly and monthly log returns, and the absolute daily, weekly and monthly log returns of the **S&P500 index in Jan. 1985 − Feb. 2011**.
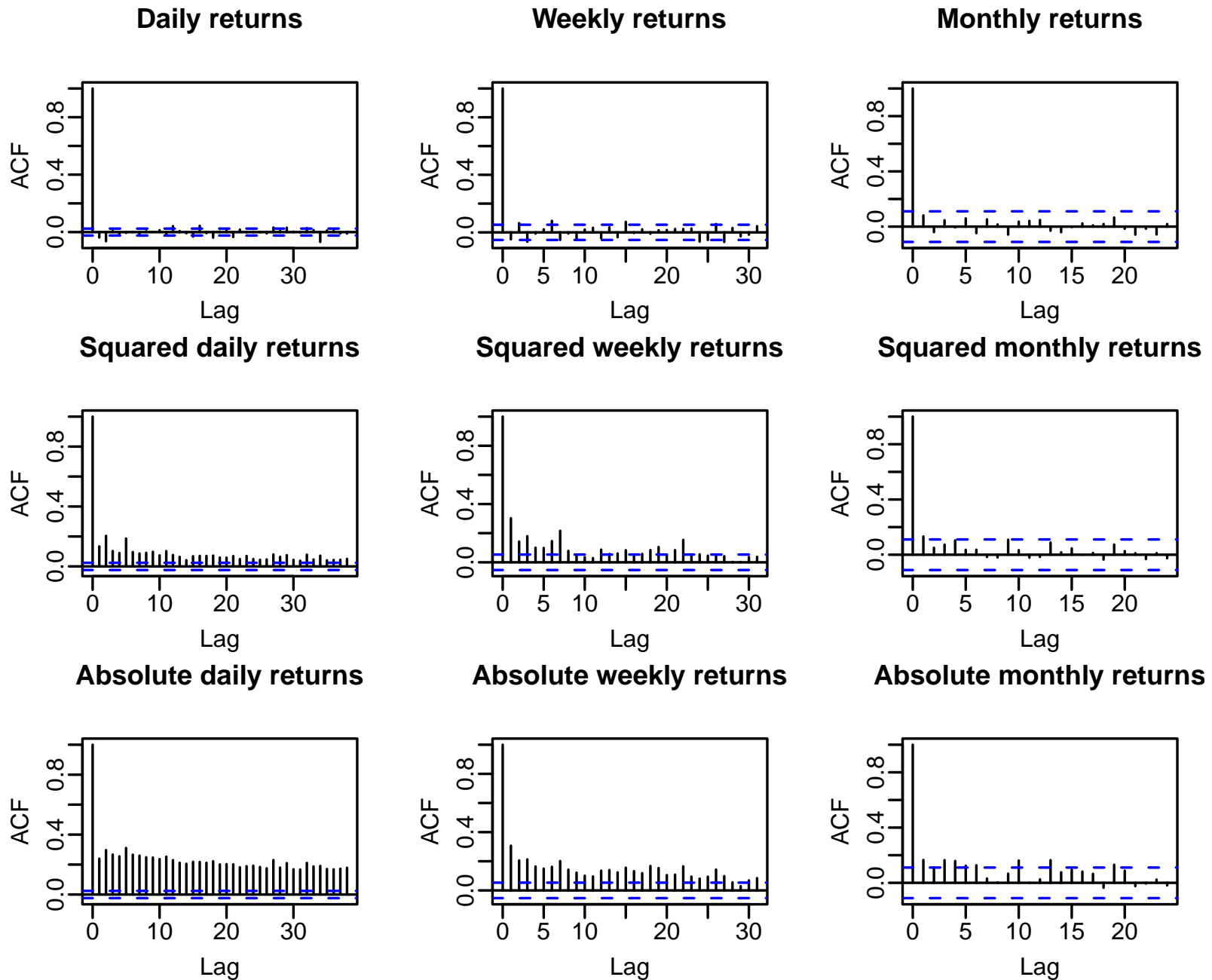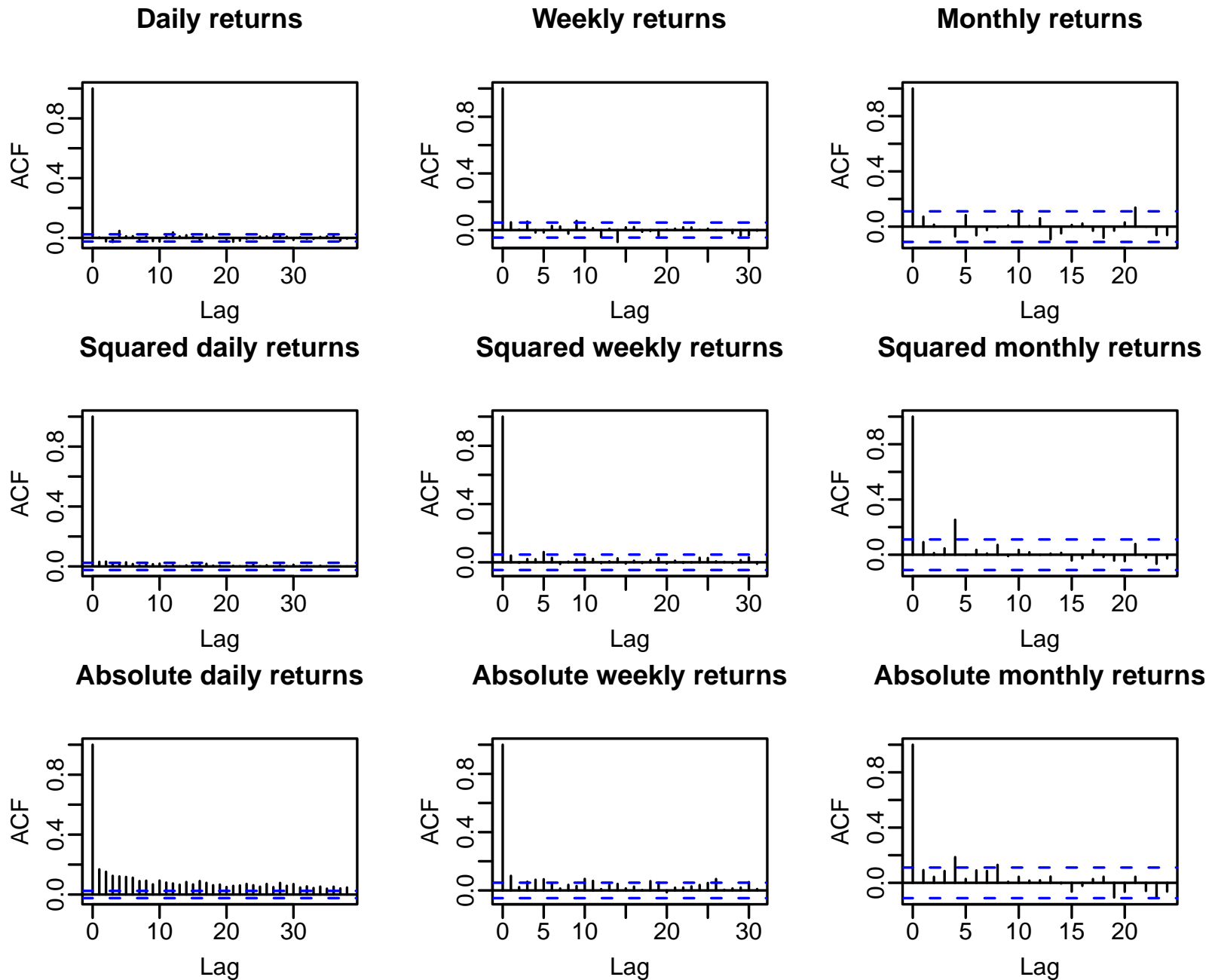
Figure 1.12: Autocorrelations of the daily, weekly and monthly log returns, the squared daily, weekly and monthly log returns, and the absolute daily, weekly and monthly log returns of the **Apple stock in Jan 1985 – Feb 2011**.

<u>Definition</u>: $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ is (weak) stationary if

(i) $\mathrm{E}X_t = \mu$,

(ii) $\mathrm{Cov}(X_t, X_{t+k}) = \gamma(k)$, independent of t — called ACF.

Thus, $\gamma(0) = \mathrm{var}(X_t)$, independent of t.

The covariance and correlation between X_t and X_{t+k} **do not change** over time. It is sufficient for **linear** predictions.

<u>Definition</u>: $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ is strict (strong) stationary if the distribution of $(X_1, \ldots, X_k)$ and $(X_{t+1}, \ldots, X_{t+k})$ are the same for any k and t.

Strong stationarity $\Longrightarrow$ weak stationarity (if $EX_t^2 < \infty$) & needed for **nonlinear** prediction.

Question: What is the covariance matrix of $(X_1, \cdots, X_T)$?

1.4   Autocorrelation Function

Autocorrelation Function is defined as

$$\rho(k) = \mathrm{Corr}(X_{t+k}, X_t) = \frac{\mathrm{Cov}(X_{t+k}, X_t)}{\sqrt{\mathrm{var}(X_{t+k})}\sqrt{\mathrm{var}(X_t)}}$$

$$= \frac{\gamma(k)}{\gamma(0)} \quad \text{for stationary time series}$$

♠measures linear relations of X_t and X_{t+k};

♠$-1 \leq \rho(k) \leq 1$ and is even $\rho(k) = \rho(-k)$;

♠is a semi-positive definite function: for any a_i

$$\sum_{i=1}^{k}\sum_{j=1}^{k}\rho(i-j)a_ia_j \geq 0.$$

Estimate of ACVF:

$$\widehat{\gamma}(k) = \frac{1}{T}\sum_{t=1}^{T-k}(X_t - \bar{X})(X_{t+k} - \bar{X}), \qquad \bar{X} = \sum_{t=1}^{T}X_t/T.$$

Sample ACF: $\widehat{\rho}(k) = \widehat{\gamma}(k)/\widehat{\gamma}(0).$

White noise: Call $\{\varepsilon_t\} \sim \text{WN}(0, \sigma^2)$, if

$$\gamma(0) = \sigma^2, \quad \gamma(j) = 0, \qquad j > 0.$$

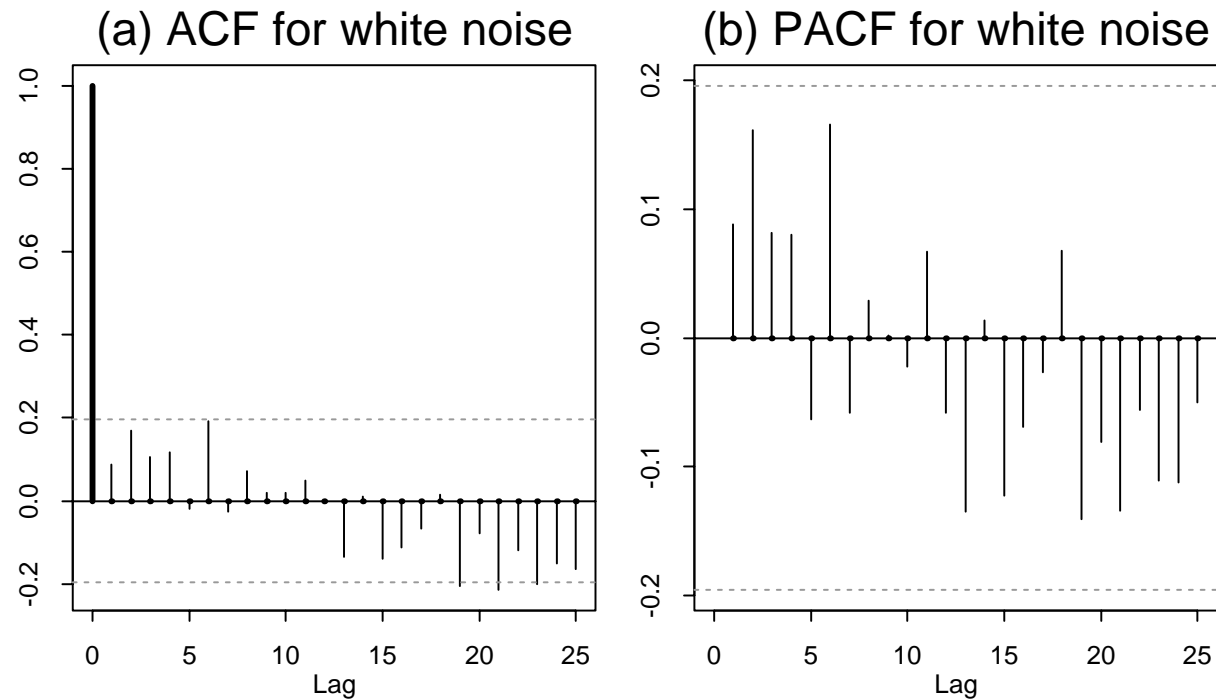But the sample ACF will not be exactly zero (Fig 1.13). What is the **confidence limit**?

Figure 1.13: ACF and PACV for a simulated white noise series with T=100.

Asymptotic normality: For independent white noise process,

$$\sqrt{T}\widehat{\rho}(j) \xrightarrow{D} \mathcal{N}\left(0, 1\right), \; j > 0$$

Thus, about 95% of acf falls in $\pm 1.96/\sqrt{T}$.

1.5   Predicability of asset returns

Predicability of asset returns is fundamental and stated in several forms, depending on the mathematical requirement. One of the forms is the **I.I.D. random walk hypothesis** for the asset returns $\{X_t\}$:

$$\{X_t\} \sim \text{IID}(0, \sigma^2).$$

Another is that the **returns are uncorrelated**. Under i.i.d.,

$$\sqrt{T}\widehat{\rho}(j) \xrightarrow{D} \mathcal{N}(0, 1), \quad \text{for} \quad j \neq 0.$$

Weak random walk hypothesis:

$$H_0 : \rho(j) = 0, \quad \forall j \neq 0.$$

<u>Portmanteau statistic</u> (Box and Pierce 1970): $Q^*(m) = T \sum_{j=1}^{m} \widehat{\rho}^2(j).$

<u>Ljung and Box (1978):</u> $Q(m) = T(T+2) \sum_{j=1}^{m} \dfrac{\widehat{\rho}^2(j)}{T-j}.$

★ Under the i.i.d. white noise hypothesis, both test statistics follow asymptotically χ^2_m-distribution.

★ Ljung-Box test statistic is usually used.

<u>Example 2</u>: Apply the techniques to monthly SP500 returns in Fig. 1.3. The ACF of the data was given in Fig. 1.11.

<u>Results</u>: P-values should be read keeping $T = 313$ in mind.

m		1	6	12	24
returns	Q_m	2.101	5.149	8.958	14.080
	p-value	0.147	0.525	0.707	0.945
squared returns	Q_m	5.517	12.292	16.964	23.474
	p-value	0.019	0.056	0.151	0.492
absolute returns	Q_m	8.687	39.283	49.721	76.446
	p-value	0.003	0.000	0.000	0.000

■ The return process is white noise.

■ Squared-returns is not white noise, but weakly correlated. The correlation at daily frequency is much larger monthly.

■ Absolute return has long memory. See also Fig 1.11.

1.6   Efficient Market Hypothesis

★ Strong Form: security prices reflect all available information, public or private

★ Semi-Strong Form: security prices reflect all public information

★ Weak form: security prices reflect all past information

The log-price $\log S_t = \log S_{t-1} + r_t$ is a random walk:

■ White noise: too weak and hence too vague. e.g. Are returns predicable? **By nonlinear rule?**

■ Independent Identically Dist (i.i.d): too strong (Nothing can be predicted, contradicting with stylized feature)

Motivation: The returns of assets are uncorrelated, but not independent. This is evidenced by the following corregrams.
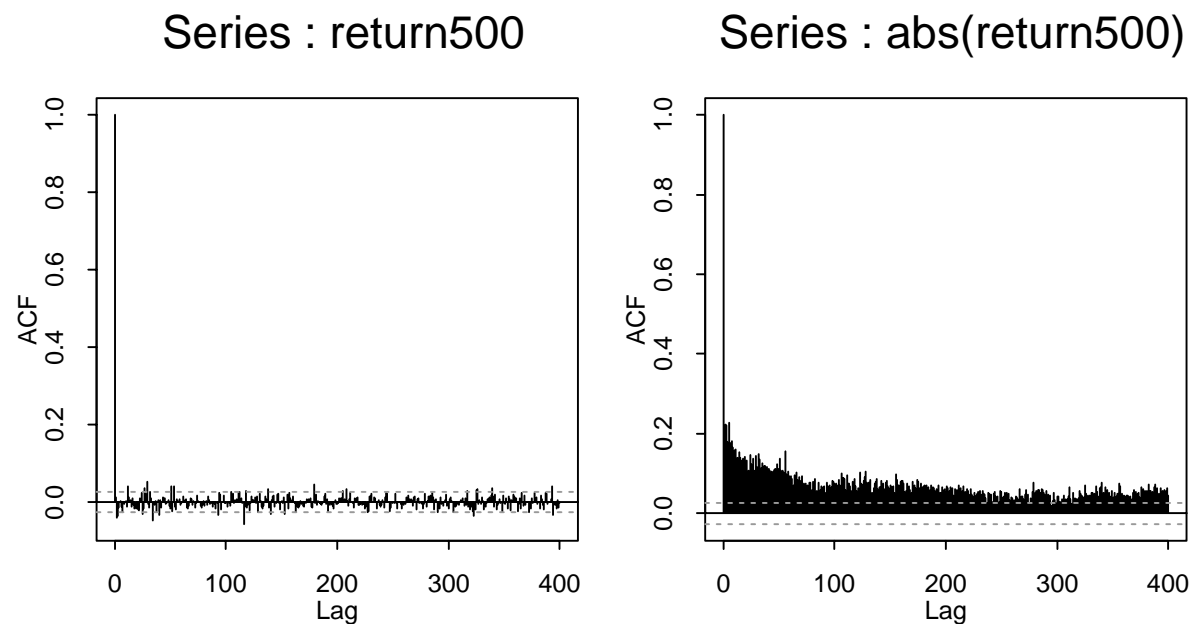

Figure 1.14: ACFs of the log-returns and absolute log-returns of the S&P 500 index.

Martingale: Let E_t be the conditional expectation given the information up to time t. A sequence $\{Y_t\}$ is called **martingale** if it is observable at time t and

$$E_t Y_{t+1} = Y_t, \qquad \forall t.$$

Hence, $\varepsilon_t = Y_t - Y_{t-1}$ is called a martingale difference, satisfying $E_t \varepsilon_{t+1} = 0$. What is $E\varepsilon_t$?

Double Expectation: $E(Y) = E\{E(Y|\mathbf{X})\}$ for any $\mathbf{X}$.

Thus, $E\varepsilon_t = E\{E_{t-1}(\varepsilon_t)\} = 0$.

Relation with prediction:

- Forecast future X_{T+m}, T — forecast origin, m — forecast horizon.

- The prediction error $= \mathrm{E}(X_{T+m} - f(X_T, X_{T-1}, \cdots))^2$.
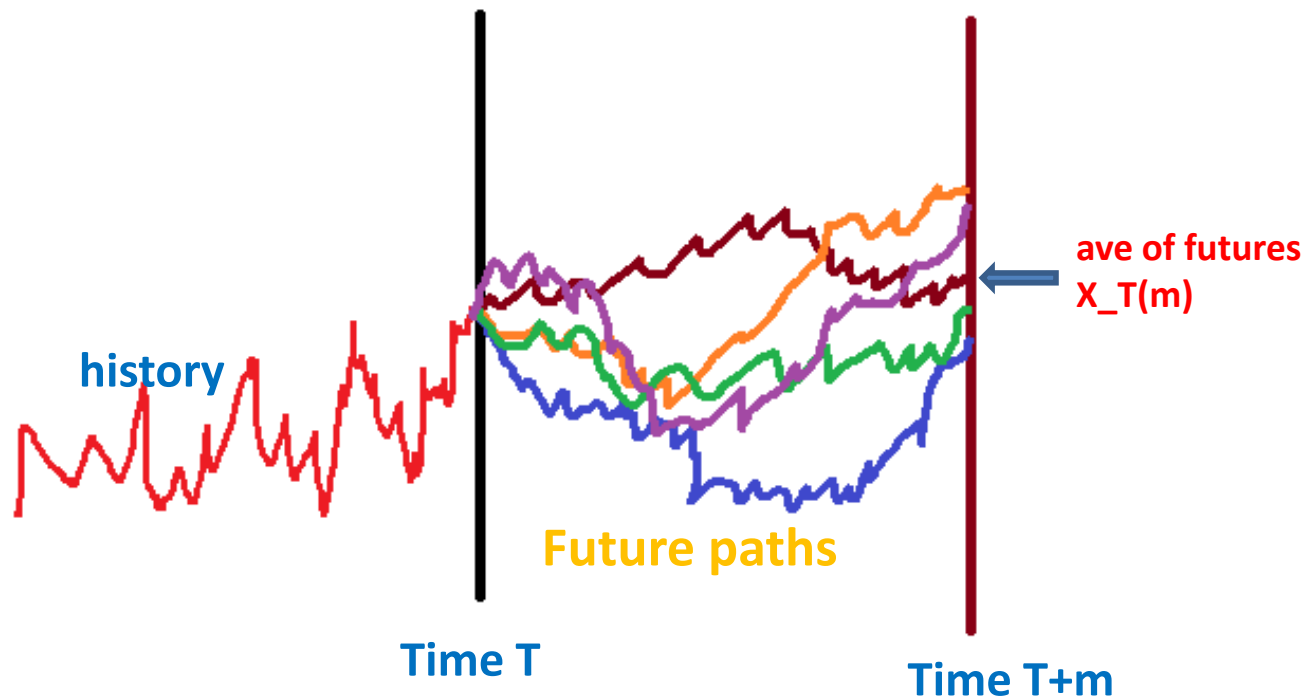
Figure 1.15: Illustration of the best prediction of X_{T+k} given $X_1, \cdots, X_T$, which is the average of possible values.

Best prediction: find f to minimize PE:

$$X_T(m) = \arg\inf_f \mathrm{E}(X_{T+m} - f)^2 = \mathrm{E}(X_{T+m}|X_1, \cdots, X_T).$$

●more precise, hard to estimate from data.

Best linear prediction: Find $\beta_0, \cdots, \beta_p$ to min prediction error

$$\mathrm{E}(X_{T+m} - \beta_0 - \beta_1 X_T - \cdots - \beta_p X_{T+1-p})^2,$$

using past p-lagged variables. Let $b_{m,0}, \cdots, b_{m,p}$ be the minimizer.

Then, the best m-step linear predictor is

$$X_T^L(m) = b_{m,0} + b_{m,1} X_T + \cdots + b_{m,p} X_{T+1-p},$$

This can be found by the least-squares method, depends only on ACF.

Martingale: The best prediction is itself (**can not** be predicted).

Example 3: A model to capture stylized features is $r_{t+1} = \sigma_t \varepsilon_{t+1}$, where $\{\varepsilon_t\}$ i.i.d, ε_t independent of history.

- Allow time-varying volatility

- Inpredicatbility of return: $E_t r_{t+1} = \sigma_t E_t \varepsilon_{t+1} = 0$.

- $\{r_t^2\}$ can be predictable: $E_t r_{t+1}^2 = \sigma_t^2 E \varepsilon_t^2$.

Properties of martingale difference:

(i) $E\varepsilon_t = E\{E_{t-1}\varepsilon_t\} = 0$ (double expectation formula);

(ii) $E_t \varepsilon_{t+m} = E_t\{E_{t+m-1}\varepsilon_{t+m}\} = 0, \quad \forall \, m \geq 1$

   (unpredictable conditional mean);

(iii) $E\varepsilon_t \varepsilon_{t+m} = E\{E_t(\varepsilon_t \varepsilon_{t+m})\} = 0, \forall m > 0$ (uncorrelated).

Relationship: $\{\text{i.i.d.}\} \Longrightarrow \{\text{martingale difference}\} \Longrightarrow \{\text{uncorrelated}\}$.

Martingale hypothesis: The asset price (log-price) is a martingale sequence, a random walk with martingale difference. Its return is a martingale difference.
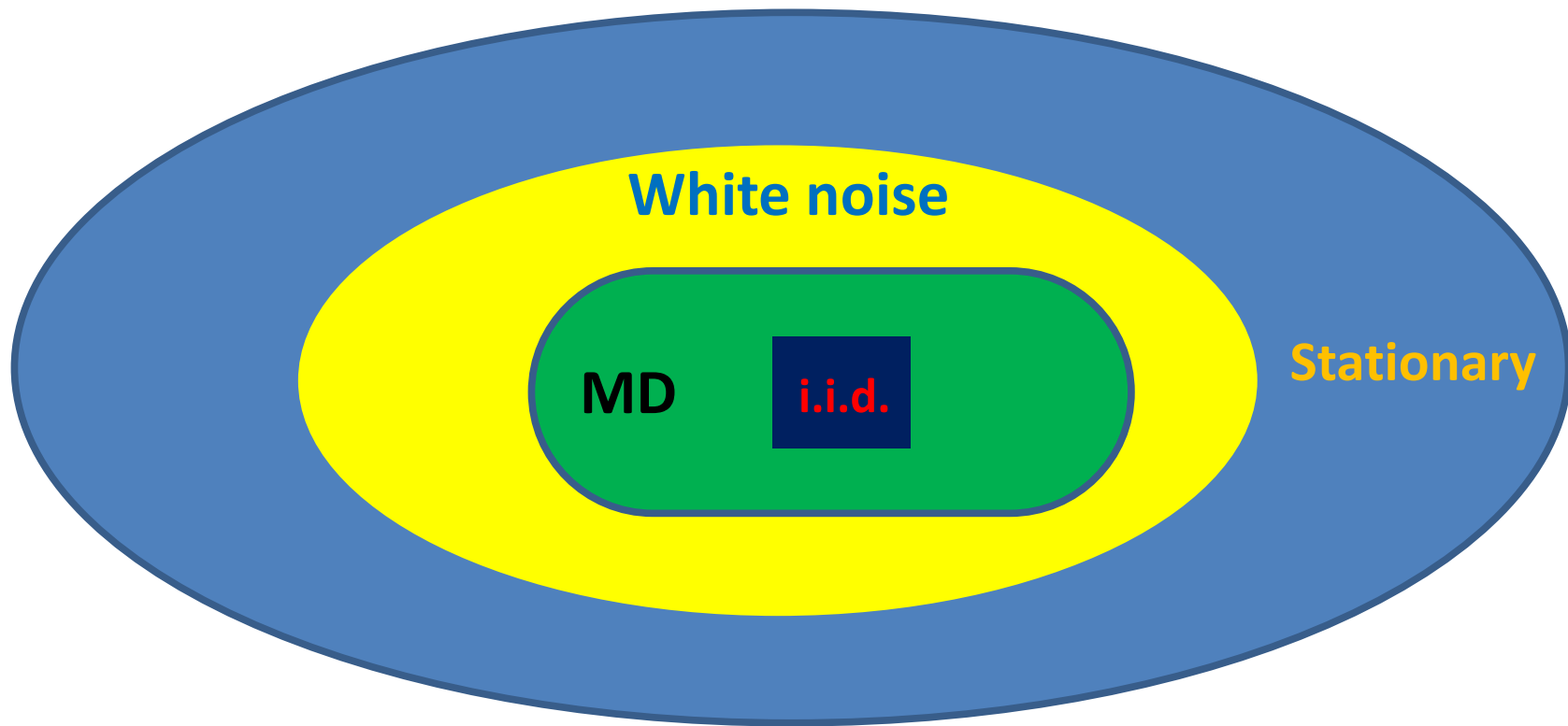
Figure 1.16: Relationship among different processes: Stationary processes are the largest set, followed by white noise, Martingale Difference (MD), and i.i.d. processes. There are many processes between stationary processes and white noise processes, to be detailed in Chapters 2 and 3.

<u>Remarks</u>: The in-predicability of asset returns r_t is the same as the random walk of (logarithm) asset prices. The meaning of in-predicability takes three mathematical forms (weakest to strongest):

- the series $\{r_t\}$ is uncorrelated (the future return can not be predicted by **linear** rules);

- the series $\{r_t\}$ is a martingale difference (the future return can not be predicted even by **nonlinear** rules);

- the series $\{r_t\}$ is an i.i.d. sequence (**nothing** can be predicted, including volatility).

The margingale hypothes is more **reasonable** on the price dynamics (returns are not predicable, but volatility can possibly be). However, the Ljung and Box test is only testing i.i.d. white noise.

Implications of martingale hypothesis: is a form of **efficient market hypothesis** with no arbitrage.

1.7   Summary

<u>**Returns**</u>: ●Simple- and log-returns (additive, approx the same); ●bond yields; ●excess returns.

<u>**Stylized features of financial returns**</u>: ●random walks; ●volatility clustering; ●heavy tails;

●asymmetry; ●aggregational Gaussianity; ●leverage effect

<u>**Tools**</u>: QQ-plot (chekcing dist), JB-test (normality), ACF (checking WN), Ljung-Box test (WN)

<u>**Use of data**</u>: What make historical data usable is the assumption:

★ Stationary for linear prediction;

★ Strong stationary for nonlinear prediction;

★ Best nonlinear prediction is conditional expectation given available information.

<u>**Martingale**</u>:

★ Best prediction is itself (fair game); suitable for log-prices (martingale hypothesis);

★ MD suitable log-returns; (stronger than white noise; weaker than independent);

★ Stylized formula for MD: $r_t = \sigma_{t-1}\varepsilon_t$ (predicable times independent).