Bus 41202: Analysis of Financial Time Series  
Spring 2015, Ruey S. Tsay

Lecture 1: Introduction

Financial time series (FTS) analysis is concerned with theory and practice of asset valuation over time.

Comparison with other T.S. analysis: similarity and difference?

Highly related, but with some added uncertainty, because FTS must deal with the ever-changing business & economic environment and the fact that volatility is not directly observed.

Objective of the course

• to learn ways to get financial information from web directly and to process the information.
• to provide some basic knowledge of financial time series data such as skewness, heavy tails, and measure of dependence between asset returns
• to introduce some statistical tools & econometric models useful for analyzing these series.
• to gain experience in analyzing FTS
• to introduce recent developments in financial econometrics and their applications, e.g., high-frequency finance
• to study methods for assessing market risk, credit risk, and expected loss. The methods discussed include Value at Risk, expected shortfall, and tail dependence.
• to analyze high-dimensional asset returns, including co-movement

Examples of financial time series

1. Daily log returns of Apple stock: 2004 to 2013 (10 years)

2. The VIX index


   Seasonal time series useful in
   • earning forecasts
   • pricing weather related derivatives (e.g. energy)
   • modeling intraday behavior of asset returns

5. US monthly interest rates (3m & 6m Treasury bills)
   Relations between the two series? Term structure of interest rates

6. Exchange rate between US Dollar vs Euro
   Fixed income, hedging, carry trade

7. Size of insurance claims
   Values of fire insurance claims ($\times1000$ Krone) that exceeded 500 from 1972 to 1992.

8. High-frequency financial data:
Figure 1: Daily log returns of Apple stock from 2004 to 2013
Figure 2: Density of daily log returns of Apple stock: 2004 to 2013
Figure 3: Time plot of daily 3-year CDS spreads of JPM: from July 20, 2004 to September 19, 2014.
Figure 4: CBOE Vix index: January 2, 2004 to March 7, 2014.
Figure 5: Quarterly earnings per share of Coca-Cola Company
Figure 6: Daily Exchange Rate: Dollars per Euro
Figure 7: Daily log returns of FX (Dollar vs Euro)
Figure 8: Histogram of daily log returns of FX (Dollar vs Euro)
Figure 9: Monthly US interest rates: 3m & 6m TB
Figure 10: Spread of monthly US interest rates: 3m & 6m TB
Figure 11: Claim sizes of the Norwegian fire insurance from 1972 to 1992, measured in 1000 Krone and exceeded 500.
CAT trade data on January 04, 2010.

<table>
<thead>
<tr>
<th>date</th>
<th>hour</th>
<th>minute</th>
<th>second</th>
<th>price</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>0</td>
<td>57.65</td>
<td>3910</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>0</td>
<td>57.7</td>
<td>400</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>0</td>
<td>57.69</td>
<td>300</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>1</td>
<td>57.65</td>
<td>462</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>1</td>
<td>57.65</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>1</td>
<td>57.65</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>1</td>
<td>57.7</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>1</td>
<td>57.7</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>1</td>
<td>57.72</td>
<td>500</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>1</td>
<td>57.72</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>2</td>
<td>57.73</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>3</td>
<td>57.73</td>
<td>300</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>3</td>
<td>57.72</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>4</td>
<td>57.72</td>
<td>300</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>5</td>
<td>57.57</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>5</td>
<td>57.57</td>
<td>500</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>5</td>
<td>57.56</td>
<td>300</td>
</tr>
<tr>
<td>. . . . .</td>
<td>. . .</td>
<td>. . .</td>
<td>. . .</td>
<td>. . .</td>
<td>. .</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>35</td>
<td>57.77</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>36</td>
<td>57.77</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.54</td>
<td>83600</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.57</td>
<td>100</td>
</tr>
<tr>
<td>. . . . .</td>
<td>. . .</td>
<td>. . .</td>
<td>. . .</td>
<td>. . .</td>
<td>. .</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.55</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.55</td>
<td>2400</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.56</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.55</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.55</td>
<td>100</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.54</td>
<td>170</td>
</tr>
<tr>
<td>20100104</td>
<td>9</td>
<td>30</td>
<td>42</td>
<td>57.54</td>
<td>200</td>
</tr>
</tbody>
</table>

Outline of the course

- Returns & their characteristics: empirical analysis (summary statistics)
- Simple linear time series models & their applications
- Univariate volatility models & their implications
• Nonlinearity in level and volatility
• Neural network & non-parametric methods
• High-frequency financial data and market micro-structure
• Continuous-time models and derivative pricing
• Value at Risk, extreme value theory and expected shortfall (also known as conditional VaR)
• Analysis of multiple asset returns: factor models, dynamic and cross dependence

Asset Returns
Let $P_t$ be the price of an asset at time $t$, and assume no dividend.

One-period simple return: Gross return

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

Simple return:

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

Multiperiod simple return: Gross return

$$1 + R_t(k) = \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}}$$
$$= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}).$$

The $k$-period simple net return is $R_t(k) = \frac{P_t}{P_{t-k}} - 1$.

Example: Table below gives five daily closing prices of Apple stock in December 2011. The 1-day gross return of holding the stock from
12/8 to 12/9 \(1 + R_t = 393.62/390.66 \approx 1.0076\) so that the daily simple return is 0.76%, which is \((393.62 - 390.66)/390.66\).

<table>
<thead>
<tr>
<th>Date</th>
<th>12/02</th>
<th>12/05</th>
<th>12/06</th>
<th>12/07</th>
<th>12/08</th>
<th>12/09</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price($)</td>
<td>389.70</td>
<td>393.01</td>
<td>390.95</td>
<td>389.09</td>
<td>390.66</td>
<td>393.62</td>
</tr>
</tbody>
</table>

**Time interval** is important! Default is one year.

**Annualized (average) return:**

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

An approximation:

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

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

<table>
<thead>
<tr>
<th>Type</th>
<th>#(payment)</th>
<th>Int.</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual</td>
<td>1</td>
<td>0.1</td>
<td>$1.10000</td>
</tr>
<tr>
<td>Semi-Annual</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>(\frac{0.1}{52})</td>
<td>$1.10506</td>
</tr>
<tr>
<td>Daily</td>
<td>365</td>
<td>(\frac{0.1}{365})</td>
<td>$1.10516</td>
</tr>
<tr>
<td>Continuously</td>
<td>(\infty)</td>
<td></td>
<td>$1.10517</td>
</tr>
</tbody>
</table>

\[A = C \exp[r \times n]\]

where \(r\) is the interest rate per annum, \(C\) is the initial capital, \(n\) is the number of years, 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 \( p_t = \ln(P_t) \).

**Multiperiod log return:**

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

**Example** Consider again the Apple stock price.

1. What is the log return from 12/8 to 12/9:
   
   **A:** \[ r_t = \ln(393.62) - \ln(390.66) = 7.5\%. \]

2. What is the log return from day 12/2 to 12/9?
   
   **A:** \[ r_t(4) = \ln(393.62) - \ln(389.7) = 1\%. \]

**Portfolio return:** \( N \) assets

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

**Example:** An investor holds stocks of IBM, Microsoft and Citigroup. Assume that her capital allocation is 30%, 30% and 40%. Use the monthly simple returns in Table 1.2 of the text. What is the mean simple return of her stock portfolio?
**Answer:** $E(R_t) = 0.3 \times 1.35 + 0.3 \times 2.62 + 0.4 \times 1.17 = 1.66$.

**Dividend payment:**

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

**Excess return:** (adjusting for risk)

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

where $r_{0t}$ denotes the log return of a reference asset (e.g. risk-free interest rate).

**Relationship:**

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

If the returns are in **percentage**, then

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

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:**

$$R = (1 + 0.0446)(1 - 0.0734)(1 + 0.1077) - 1 = 1.0721 - 1 = 0.0721 = 7.21\%.$$
Distributional properties of returns

Key: What is the distribution of 
\{r_{it}; i = 1, \cdots, N; t = 1, \cdots, T\}?

Some theoretical properties:

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

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

First moment: mean or expectation of \(X\).
\(\ell\)-th central moment

\[
m_\ell = E[(X - \mu_x)^\ell] = \int_{-\infty}^{\infty} (x - \mu_x)^\ell f(x)dx,
\]

2nd c.m.: Variance of \(X\).

Skewness (symmetry) and kurtosis (fat-tails)

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

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

Why are the mean and variance of returns important?
They are concerned with long-term return and risk, respectively.

Why is return 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:
  \[ \hat{K}(x) = \frac{1}{(T-1)\hat{\sigma}_x^4} \sum_{t=1}^{T} (x_t - \hat{\mu}_x)^4. \]

Under normality assumption,
\[ \hat{S}(x) \sim N(0, \frac{6}{T}), \quad \hat{K}(x) - 3 \sim N(0, \frac{24}{T}). \]

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 \(H_o\) of a symmetric distribution if \(|S^*| > Z_{\alpha/2}\) or p-value is less than \(\alpha\).

2. Test for tail thickness:
   \[ K^* = \frac{\hat{K}(x) - 3}{\sqrt{24/T}} \sim N(0, 1) \]
if normality holds.

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

3. A joint test (Jarque-Bera test):

$$JB = (K^*)^2 + (S^*)^2 \sim \chi^2_2$$

if normality holds, where $\chi^2_2$ denotes a chi-squared distribution with 2 degrees of freedom.

**Decision rule:** Reject $H_o$ of normality if $JB > \chi^2_2(\alpha)$ or p-value is less than $\alpha$.

**Empirical properties of returns**

Data sources:

- Course web:
  
  http://faculty.chicagobooth.edu/ruey.tsay/teaching/bs41202/sp2015/

- CRSP: Center for Research in Security Prices (Wharton WRDS)
  
  http://wrds.wharton.upenn.edu/

- Various web sites, e.g. Federal Reserve Bank at St. Louis
  
  http://research.stlouisfed.org/fred2/

- Data sets of the textbook:
  
  http://faculty.chicagobooth.edu/ruey.tsay/teaching/fts3/

Empirical dist of asset returns tends to be skewed to the left with heavy tails and has a higher peak than normal dist. See Table 1.2 of the text.

**Demonstration of Data Analysis**
Figure 12: Comparison of empirical IBM return densities (solid) with Normal densities (dashed)
**R demonstration:** Use monthly IBM stock returns from 1967 to 2008.

**** Task: (a) Set the working directory
(b) Load the library ‘‘fBasics’’.
(c) Compute summary (or descriptive) statistics
(d) Perform test for mean return being zero.
(e) Perform normality test using the Jaque-Bera method.
(f) Perform skewness and kurtosis tests.

```r
> setwd("C:/Users/rst/teaching/bs41202/sp2015")  # set working directory
> library(fBasics)  # Load the library ‘‘fBasics’’.

> da=read.table("m-ibm6708.txt",header=T)  # Load data with header on top
> da[1,]  # Show the first row of the data
   date   ibm   sprtn
1 19670331 0.048837 0.03941

> ibm=da[,2]  # Get "ibm" simple returns
> rt=log(ibm+1)  # Transform into log returns
> plot(rt,type='l')  # Plot log returns with caption.
> title(main='Time plot of monthly log returns of IBM stock from 1967,3 to 2008.12')

> mean(rt)  # Compute sample mean
[1] 0.006208082

> var(rt)  # Compute sample variance
[1] 0.005258775

> skewness(rt)  # Compute sample skewness
[1] -0.1353432

> kurtosis(rt)  # Compute sample excess kurtosis.
[1] 1.693092

> basicStats(rt)  # Compute all summary (or descriptive) statistics

rt
 nobs 502.000000
NAs 0.000000
Minimum -0.303683
Maximum 0.302915
1. Quartile -0.037641
3. Quartile 0.048443
Mean 0.006208
Median 0.005260
Sum 3.116457
SE Mean 0.003237
LCL Mean -0.000151
UCL Mean 0.012567
Variance 0.005259
```
Stdev 0.072517
Skewness -0.135343
Kurtosis 1.693092

> t.test(rt)  # Test the mean log return being zero.

    One Sample t-test
    data: rt
    t = 1.9181, df = 501, p-value = 0.05567
    alternative hypothesis: true mean is not equal to 0
    95 percent confidence interval:
     -0.0001509194 0.0125670842
    sample estimates:
      mean of x
     0.006208082

> normalTest(rt,method='jb')  # Test for the normality assumption

    Title:
    Jarque - Bera Normality Test

    Test Results:
      STATISTIC:
        X-squared: 62.8363
      P VALUE:
        Asymptotic p Value: 2.265e-14

> s3=skewness(rt)  # Perform skewness test
> T=length(rt)
> tst=s3/sqrt(6/T)
> tst
[1] -1.237977
> pv=2*pnorm(tst)  # Compute two-sided p-value of the test statistic.
> pv
[1] 0.2157246

> k4=kurtosis(rt)  # Perform kurtosis test.
> tst=k4/sqrt(24/T)
> tst
[1] 7.743311
>
>q()  # quit R.

Normal and lognormal dists
Y is lognormal if X = ln(Y) is normal.
If \( X \sim N(\mu, \sigma^2) \), then \( Y = \exp(X) \) is lognormal with mean and variance

\[
E(Y) = \exp(\mu + \frac{\sigma^2}{2}), \quad V(Y) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1].
\]

Conversely, if \( Y \) is lognormal with mean \( \mu_y \) and variance \( \sigma_y^2 \), then \( X = \ln(Y) \) is normal with mean and variance

\[
E(X) = \ln \left[ \frac{\mu_y}{\sqrt{1 + \frac{\sigma_y^2}{\mu_y^2}}} \right], \quad V(X) = \ln \left[ 1 + \frac{\sigma_y^2}{\mu_y^2} \right].
\]

**Application**: If the log return of an asset is normally distributed with mean 0.0119 and standard deviation 0.0663, then what is the mean and standard deviation of its simple return?

**Answer**: Solve this problem in two steps.

**Step 1**: Based on the prior results, the mean and variance of \( Y_t = \exp(r_t) \) are

\[
E(Y) = \exp \left[ 0.0119 + \frac{0.0663^2}{2} \right] = 1.014
\]

\[
V(Y) = \exp(2 \times 0.0119 + 0.0663^2)[\exp(0.0663^2) - 1] = 0.0045
\]

**Step 2**: Simple return is \( R_t = \exp(r_t) - 1 = Y_t - 1 \). Therefore,

\[
E(R) = E(Y) - 1 = 0.014
\]

\[
V(R) = V(Y) = 0.0045, \quad \text{standard dev} = \sqrt{V(R)} = 0.067
\]

**Remark**: See the monthly IBM stock returns in Table 1.2.

**Processes considered**

- return series (e.g., ch. 1, 2, 5)
• volatility processes (e.g., ch. 3, 4, 10, 12)
• continuous-time processes (ch. 6)
• extreme events (ch. 7)
• multivariate series (ch. 8, 9, 10)

Likelihood function (for self study)
Finally, it pays to study the likelihood function of returns \( \{r_1, \cdots, r_T\} \) discussed in Chapter 1.

Basic concept:
Joint dist = Conditional dist \( \times \) Marginal dist, i.e.
\[
f(x, y) = f(x|y)f(y)
\]
For two consecutive returns \( r_1 \) and \( r_2 \), we have
\[
f(r_2, r_1) = f(r_2|r_1)f(r_1).
\]
For three returns \( r_1, r_2 \) and \( r_3 \), by repeated application,
\[
f(r_3, r_2, r_1) = f(r_3|r_2, r_1)f(r_2, r_1)
= f(r_3|r_2, r_1)f(r_2|r_1)f(r_1).
\]
In general, we have
\[
f(r_T, r_{T-1}, \cdots, r_2, r_1)
= f(r_T|r_{T-1}, \cdots, r_1)f(r_{T-1}, \cdots, r_1)
= f(r_T|r_{T-1}, \cdots, r_1)f(r_{T-1}|r_{T-2}, \cdots, r_1)f(r_{T-2}, \cdots, r_1)
= \cdots
= \left[ \prod_{t=2}^{T} f(r_t|r_{t-1}, \cdots, r_1) \right] f(r_1),
\]
where \( \prod_{t=2}^{T} \) denotes product.
If $r_t | r_{t-1}, \ldots, r_1$ is normal with mean $\mu_t$ and variance $\sigma^2_t$, then likelihood function becomes

$$f(r_T, r_{T-1}, \ldots, r_1) = \prod_{t=2}^T \frac{1}{\sqrt{2\pi\sigma_t}} \exp \left[ -\frac{(r_t - \mu_t)^2}{2\sigma^2_t} \right] f(r_1).$$

For simplicity, if $f(r_1)$ is ignored, then the likelihood function becomes

$$f(r_T, r_{T-1}, \ldots, r_1) = \prod_{t=2}^T \frac{1}{\sqrt{2\pi\sigma_t}} \exp \left[ -\frac{(r_t - \mu_t)^2}{2\sigma^2_t} \right].$$

This is the *conditional* likelihood function of the returns under normality.

Other dists, e.g. Student-$t$, can be used to handle heavy tails.

**Model specification**

- $\mu_t$: discussed in Chapter 2
- $\sigma^2_t$: Chapters 3 and 4.

**Quantifying dependence**: Consider two variables $X$ and $Y$.

- Correlation coefficient:
  $$\rho = \frac{\text{Cov}(X, Y)}{\text{std}(X)\text{std}(Y)}.$$

- Kendall’s tau: Let $(\tilde{X}, \tilde{Y})$ be a random copy of $(X, Y)$.
  $$\rho_\tau = P[(X - \tilde{X})(Y - \tilde{Y}) > 0] - P[(X - \tilde{X})(Y - \tilde{Y}) < 0] = E[\text{sign}((X - \tilde{X})(Y - \tilde{Y}))].$$

This measure quantifies the probability of *concordant* over *discordant*. Here concordant means $(X - \tilde{X})(Y - \tilde{Y}) > 0$. For spherical distributions, e.g., normal, $\rho_\tau = \frac{2}{\pi} \sin^{-1}(\rho)$. 

27
• Spearman’s rho: rank correlation. Let $F_x(x)$ and $F_y(y)$ be the cumulative distribution function of $X$ and $Y$.

$$\rho_s = \rho(F_x(X), F_y(Y)).$$

That is, the correlation coefficient of probability-transformed variables. It is just the correlation coefficient of the ranks of the data.

Why do we consider different measures of dependence?

• Correlation coefficient encounters problems when the distributions are not normal (spherical, in general). This is particularly relevant in risk management.

• Correlation coefficient focuses no linear dependence and is not robust to outliers.

• The actual range of the correlation coefficient can be much smaller than $[-1, 1]$.

R Demonstration

```r
> da=read.table("m-ibm6708.txt",header=T)
> head(da)
      date  ibm    sprtn
1 19670331 0.048837 0.039410
2 19670430  0.15317  0.11661
3 19670531  0.06712  0.05841
4 19670630  0.03720  0.02800
5 19670731 -0.013589 -0.011715
6 19670831 -0.013589 -0.011715
> dim(da)
[1] 502  3
> ibm=da$ibm
> sp=da$sprtn
> plot(sp,ibm)  # not show in this lecture note.
> cor(ibm,sp)
[1] 0.5949041
> cor(sp,ibm,method="kendall")
[1] 0.4223351
> cor(sp,ibm,method="spearman")
```
> x=rank(ibm)
> y=rank(sp)
> cor(x,y)
> x[1:10]
> [1] 375 455 130 408 295 191 453 432 328 305
> y[1:10]
> [1] 398 414 44 310 427 157 365 90 219 348

****
> z=rnorm(1000)
> x=exp(z)
> y=exp(20*z)
> cor(x,y)
[1] 0.3187030
> cor(x,y,method='kendall')
[1] 1
> cor(x,y,method='spearman')
[1] 1

Takeaway
1. Understand the summary statistics of asset returns
2. Understand various definitions of returns & their relationships
3. Learn basic characteristics of FTS
4. Learn the basic R functions. (See Rcommands-lec1.txt on the course web.)

R commands used to produce plots in Lecture 1.

> x=read.table("d-aapl0413.txt",header=T)  
> dim(x)  
[1] 2517 3 
> x[1,]  
Permno date    rtn
1   14593 20040102 -0.004212
> y=ts(x[3],frequency=252,start=c(2004,1))  
> plot(y,type='l',xlab='year',ylab='rtn')
> title(main='Daily returns of Apple stock: 2004 to 2013')
> par(mfcol=c(2,1))  # To put two plots on a single page
> y=y*100  # percentage returns
> hist(y,nclass=50)
> title(main='Percentage returns')
> d1=density(y)
> plot(d1$x,d1$y,xlab='returns',ylab='den',type='l')

> x=read.table("m-tb3ms.txt",header=T)  # Load 3m-TB rates
> dim(x)
[1] 914 4

> y=read.table("m-tb6ms.txt",header=T)  # Load 6m-TB rates
> dim(y)
[1] 615 4
> 914-615
[1] 299
> x[300,]  # Check date of the 3m-TB
          year mon day value
300 1958 12 1 2.77
> y[1,]  # Check date of the 1st observation of 6m-TB
          year mon day value
 1 1958 12 1 3.01

> int=cbind(x[300:914,4],y[,4])  # Line up the two TB rates
> tdx=(c(1:615)+11)/12+1959
> par(mfcol=c(1,1))
> max(int)
[1] 16.3
> plot(tdx,int[,1],xlab='year',ylab='rate',type='l',ylim=c(0,16.5))
> lines(tdx,int[,2],lty=2)  # Plot the 6m-TB rate on the same frame.

> plot(tdx,int[,2]-int[,1],xlab='year',ylab='spread',type='l')
> abline(h=c(0))  # Draw a horizontal line to 'zero'.

> x=read.table("q-ko-earns8309.txt",header=T)  # Load KO data
> dim(x)
[1] 107 3
> x[1,]
          pends anntime value
 1 19830331 19830426 0.0375
> tdx=c(1:107)/12+1983
> plot(tdx,x[,3],xlab='year',ylab='earnings',type='l')
> title(main='EPS of Coca Cola: 1983-2009')
> points(tdx,x[,3])

>
> y=read.table("d-exuseu.txt",header=T)  # Load USEU exchange rates
> dim(y)
[1] 3567   4
> y[1,]
  year mon  day value
1 1999 1    4   1.1812
> tdx=c(1:3567)/252+1999
> plot(tdx,y[,4],xlab='year',ylab='eu',type='l')
> title(main='Dollars per Euro')

> r=diff(log(y[,4]))  # Compute log returns
> plot(tdx[2:3567],r,xlab='year',ylab='rtn',type='l')
> title(main='ln-rtn: US-EU')

> hist(r,nclass=50)
> title(main='useu: ln-rtn')
Linear Time Series (TS) Models

Financial TS: collection of a financial measurement over time
Example: log return $r_t$

Data: $\{r_1, r_2, \ldots, r_T\}$ (T data points)
Purpose: What is the information contained in $\{r_t\}$?

Basic concepts

- Stationarity:
  - Strict: distributions are time-invariant
  - Weak: first 2 moments are time-invariant

What does weak stationarity mean in practice?

Past: time plot of $\{r_t\}$ varies around a fixed level within a finite range!

Future: the first 2 moments of future $r_t$ are the same as those of the data so that meaningful inferences can be made.

- Mean (or expectation) of returns:
  \[ \mu = E(r_t) \]

- Variance (variability) of returns:
  \[ \text{Var}(r_t) = E[(r_t - \mu)^2] \]

- Sample mean and sample variance are used to estimate the mean and variance of returns.
  \[ \bar{r} = \frac{1}{T} \sum_{t=1}^{T} r_t \quad \& \quad \text{Var}(r_t) = \frac{1}{T - 1} \sum_{t=1}^{T} (r_t - \bar{r})^2 \]
• Test $H_o: \mu = 0$ vs $H_a: \mu \neq 0$. Compute

$$t = \frac{\bar{r}}{\text{std}(\bar{r})} = \frac{\bar{r}}{\sqrt{\text{Var}(r_t)/T}}$$

Compare $t$ ratio with $N(0, 1)$ dist. 

**Decision rule:** Reject $H_o$ of zero mean if $|t| > Z_{\alpha/2}$ or p-value is less than $\alpha$.

• Lag-$k$ autocovariance:

$$\gamma_k = \text{Cov}(r_t, r_{t-k}) = E[(r_t - \mu)(r_{t-k} - \mu)].$$

• Serial (or auto-) correlations:

$$\rho_\ell = \frac{\text{cov}(r_t, r_{t-\ell})}{\text{var}(r_t)}$$

Note: $\rho_0 = 1$ and $\rho_k = \rho_{-k}$ for $k \neq 0$. Why?

Existence of serial correlations implies that the return is predictable, indicating market inefficiency.

• Sample autocorrelation function (ACF)

$$\hat{\rho}_\ell = \frac{\sum_{t=1}^{T-\ell}(r_t - \bar{r})(r_{t+\ell} - \bar{r})}{\sum_{t=1}^{T}(r_t - \bar{r})^2},$$

where $\bar{r}$ is the sample mean & $T$ is the sample size.

• Test zero serial correlations (market efficiency)

  – Individual test: for example,

$$H_o: \rho_1 = 0 \text{ vs } H_a: \rho_1 \neq 0$$

$$t = \frac{\hat{\rho}_1}{\sqrt{1/T}} = \sqrt{T\hat{\rho}_1}$$
Asym. $N(0, 1)$.

**Decision rule:** Reject $H_o$ if $|t| > Z_{\alpha/2}$ or p-value less than $\alpha$.

- Joint test (Ljung-Box statistics):
  
  $H_o : \rho_1 = \cdots = \rho_m = 0$ vs $H_a : \rho_i \neq 0$

  
  $$Q(m) = T(T + 2) \sum_{\ell=1}^{m} \frac{\hat{\rho}_{\ell}^2}{T - \ell}$$

  Asym. chi-squared dist with $m$ degrees of freedom.

  **Decision rule:** Reject $H_o$ if $Q(m) > \chi^2_m(\alpha)$ or p-value is less than $\alpha$.

- Sources of serial correlations in financial TS
  - Nonsynchronous trading (ch. 5)
  - Bid-ask bounce (ch. 5)
  - Risk premium, etc. (ch. 3)

  Thus, significant sample ACF does not necessarily imply market inefficiency.

**Example:** Monthly returns of IBM stock from 1926 to 1997.

- $R_t$: $Q(5) = 5.4(0.37)$ and $Q(10) = 14.1(0.17)$

- $r_t$: $Q(5) = 5.8(0.33)$ and $Q(10) = 13.7(0.19)$

**Remark:** What is p-value? How to use it?

Implication: Monthly IBM stock returns do not have significant serial correlations.

**Example:** Monthly returns of CRSP value-weighted index from 1926 to 1997.
• \( R_t: Q(5) = 27.8 \) and \( Q(10) = 36.0 \)

• \( r_t: Q(5) = 26.9 \) and \( Q(10) = 32.7 \)

All highly significant. Implication: there exist significant serial correlations in the value-weighted index returns. (Nonsynchronous trading might explain the existence of the serial correlations, among other reasons.) Similar result is also found in equal-weighted index returns.

**R demonstration:** IBM monthly simple returns from 1967 to 2008.

```r
> da=read.table("m-ibm6708.txt",header=T)
> ibm=da[,2]
> library(fSeries)
Loading required package: robustbase
.....
The new version of 'fSeries' has been renamed to 'timeSeries'

> acf(ibm,lag=15) \<== Obtain plot of ACFs
> m1=acf(ibm,lag=15) \<== Obtain plot of ACFs and more.
> names(m1)
[1] "acf" "type" "n.used" "lag" "series" "snames"
> m1$acf
     [,1]  
[1,] 1.0000000000 
[2,] -0.0007050032 
[3,] -0.0028369916 
[4,] 0.0353196823 
[5,] -0.0740449728 
......
[15,] 0.0013865609 
[16,] -0.0360026726

> m2=pacf(ibm,lag=10) \<== Obtain partial ACFs.
> names(m2)
[1] "acf" "type" "n.used" "lag" "series" "snames"
> m2$acf
     [,1]  
[1,] -0.0007050032 
[2,] -0.0028374901 
[3,] 0.0353195823 
......
```

35
> Box.test(ibm,lag=10)  \textless{}\textgreater{}  \texttt{Perform Box-Pierce test for serial correlations}

\textbf{Box-Pierce test}

data: ibm
X-squared = 6.8338, df = 10, p-value = 0.741

> Box.test(ibm,lag=10,type='Ljung')  \textless{}\textgreater{}  \texttt{Perform Box-Ljung test for serial correlations.}

\textbf{Box-Ljung test}

data: ibm
X-squared = 6.9444, df = 10, p-value = 0.7307

\textbf{Back-shift (lag) operator}

A useful notation in TS analysis.

\begin{itemize}
  \item Definition: $Br_t = r_{t-1}$ or $Lr_t = r_{t-1}$
  \item $B^2r_t = B(Br_t) = Br_{t-1} = r_{t-2}$.
\end{itemize}

$B$ (or $L$) means time shift! $Br_t$ is the value of the series at time $t - 1$.

Suppose that the daily log returns are

<table>
<thead>
<tr>
<th>Day</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_t$</td>
<td>0.017</td>
<td>-0.005</td>
<td>-0.014</td>
<td>0.021</td>
</tr>
</tbody>
</table>

Answer the following questions:

\begin{itemize}
  \item $r_2 =$
  \item $Br_3 =$
  \item $B^2r_5 =$
\end{itemize}

\textbf{Question:} What is $B2$?
What are the important statistics in practice?
Conditional quantities, not unconditional

**A proper perspective:** at a time point $t$

- Available data: $\{r_1, r_2, \ldots, r_{t-1}\} \equiv F_{t-1}$
- The return is decomposed into two parts as

$$r_t = \text{predictable part} + \text{not predictable part}$$

$$= \text{function of elements of } F_{t-1} + a_t$$

In other words, given information $F_{t-1}$

$$r_t = \mu_t + a_t$$

$$= E(r_t | F_{t-1}) + \sigma_t \epsilon_t$$

- $\mu_t$: conditional mean of $r_t$
- $a_t$: shock or innovation at time $t$
- $\epsilon_t$: an iid sequence with mean zero and variance 1
- $\sigma_t$: conditional standard deviation (commonly called volatility in finance)

Traditional TS modeling is concerned with $\mu_t$:

Model for $\mu_t$: **mean equation**

Volatility modeling concerns $\sigma_t$.

Model for $\sigma_t^2$: **volatility equation**

**Univariate TS analysis serves two purposes**

- a model for $\mu_t$
- understanding models for $\sigma_t^2$: properties, forecasting, etc.
**Linear time series**: $r_t$ is linear if

- the predictable part is a linear function of $F_{t-1}$
- $\{a_t\}$ are independent and have the same dist. (iid)

Mathematically, it means $r_t$ can be written as

$$r_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i},$$

where $\mu$ is a constant, $\psi_0 = 1$ and $\{a_t\}$ is an iid sequence with mean zero and well-defined distribution.

In the economic literature, $a_t$ is the *shock* (or *innovation*) at time $t$ and $\{\psi_i\}$ are the *impulse* responses of $r_t$.

**White noise**: iid sequence (with finite variance), which is the building block of linear TS models.
White noise is not predictable, but has zero mean and finite variance.

**Univariate linear time series models**

1. autoregressive (AR) models
2. moving-average (MA) models
3. mixed ARMA models
4. seasonal models
5. regression models with time series errors
6. fractionally differenced models (long-memory)

**Example** Quarterly growth rate of U.S. real gross national product (GNP), seasonally adjusted, from the second quarter of 1947 to the first quarter of 1991.
An AR(3) model for the data is

\[ r_t = 0.005 + 0.35r_{t-1} + 0.18r_{t-2} - 0.14r_{t-3} + a_t, \quad \hat{\sigma}_a = 0.01, \]

where \( \{a_t\} \) denotes a white noise with variance \( \sigma_a^2 \). Given \( r_n, r_{n-1} \& r_{n-2} \), we can predict \( r_{n+1} \) as

\[ \hat{r}_{n+1} = 0.005 + 0.35r_n + 0.18r_{n-1} - 0.14r_{n-2}. \]

Other implications of the model?

In this course, we use *statistical methods* to find models that fit the data well for making inference, e.g. prediction. On the other hand, there exists economic theory that leads to time-series models for economic variables. For instance, consider the *real business-cycle theory* in macroeconomics. Under some simplifying assumptions, one can show that \( \ln(Y_t) \), where \( Y_t \) is the output (GDP), follows an AR(2) model. See *Advanced Macroeconomics* by David Romer (2006, 3rd, pp. 190).

**Example:** Monthly simple return of Center for Research in Security Prices (CRSP) equal-weighted index

\[ R_t = 0.013 + a_t + 0.178a_{t-1} - 0.13a_{t-3} + 0.135a_{t-9}, \quad \hat{\sigma}_a = 0.073 \]

Checking: \( Q(10) = 11.4(0.122) \) for the residual series \( a_t \).

Implications of the model?

Statistical significance vs economic significance.

In this course, we shall discuss some reasons for the observed serial dependence in index returns. See, for example, Chapter 5 on nonsynchronous trading.

**Important properties of a model**
• Stationarity condition
• Basic properties: mean, variance, serial dependence
• Empirical model building: specification, estimation, & checking
• Forecasting