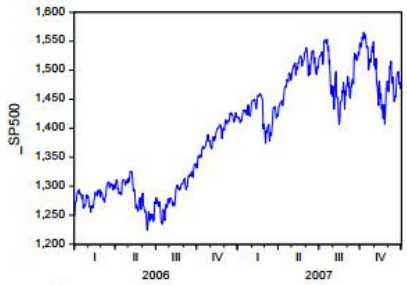
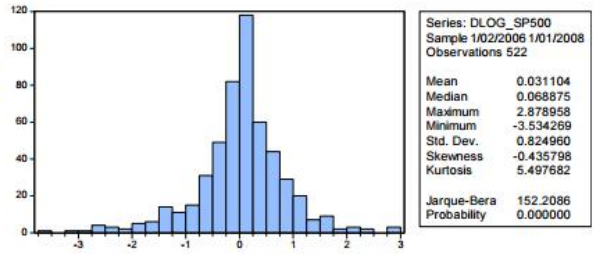
Econ334 Financial Econometrics

Characteristics of Financial Data

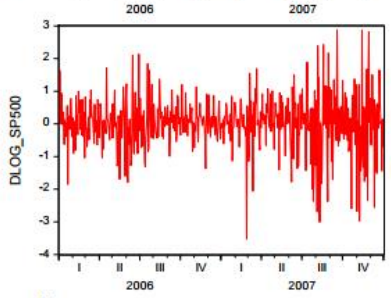
Overview

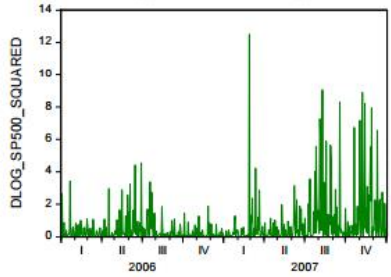




Series: DLOG_SP500
Sample 1/02/2006 1/01/2008
Observations 522

Mean	0.031104
Median	0.068875
Maximum	2.878958
Minimum	-3.534269
Std. Dev.	0.824960
Skewness	-0.435798
Kurtosis	5.497682
Jarque-Bera	152.2086
Probability	0.000000





Characteristics of asset prices

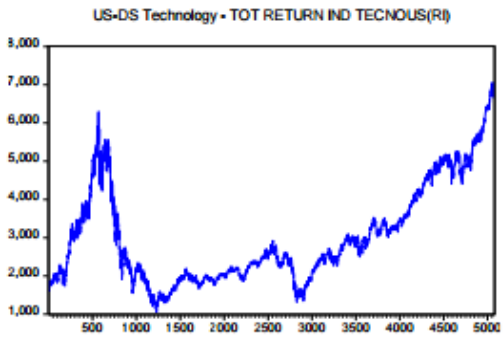
- Large number of *small movements*
- Long-run *trends*
- *Turning Points*
- *No* tendency to return to an *average level*
- *Volatility Clustering* - periods of high volatility and periods of relative quiet
- *Asymmetric Volatility*
- *Thick tails*
- *Skewness*

- **Our Task**
  - Build statistical models for financial time series data that can be used for
    - Forecasting
    - Asset pricing
    - Portfolio management and hedging
    - Testing of economic theories

Types of Financial Data

We will mainly use **Time Series Data**:

- Collected over a period of time
- Have a particular frequency related to the time interval of collection (monthly, weekly, daily, five-minute, etc).
- Example daily price level of US Technology Index (Jan 1998 – July 2017)

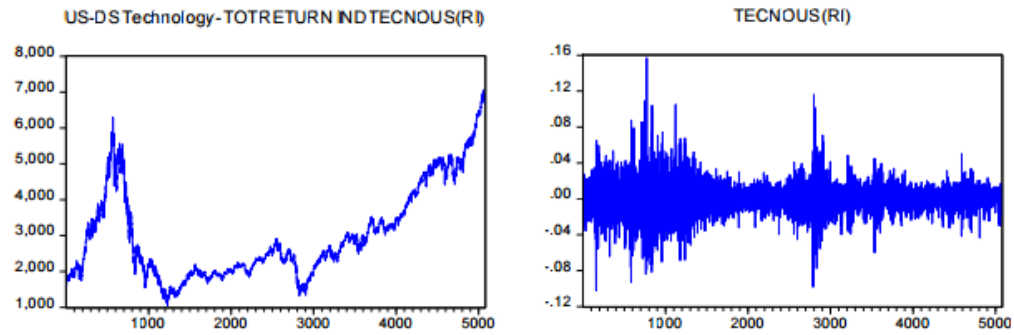


Other types of data includes:

- **Cross Sectional Data:**
  Collected at a single point in time, e.g. a cross-section of stock returns on the ASX at end of day
- **Panel Data:**
  Have both time and cross-section dimensions, e.g. track each company on the ASX over a period of time

Transforming Financial Data

We usually analyse returns, rather than prices. **Prices** depend on scale, for example, $1m buys less in 2017 than what it used to be able to buy in 2000 due to the impact of inflation. **Returns** have better statistical properties than prices. Returns are stationary.



Simple Returns

One period simple net return:

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

E.g. buy 1 share in AMP at end June 2014 for $5.968 per share and sell it at end June 2015 for $6.470 per share.

- 1 year net return: 6.470/5.968 − 1 = 0. 0841
- This is equal to 8.41% in decimal notation

Compounding of Simple Returns

The k period (holding) return Net return $R_t[k]$ is

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

- May be computed as the product of the single period returns over k periods prior to time t minus 1
- Multi-period returns are called *compound* returns
- e.g let R_t be monthly returns then …

- The k period net return

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

- If $R_t = R$ for all t (single period returns are the same)

$$P_t = P_{t-k}\left(1+R_t[k]\right) = P_{t-k}(1+R)^k$$

Continuous Returns

Consider what happens as the number of compounding intervals n per year grows and R is *the annual interest rate*.

- Annual compounding

$$P_t = P_0(1+R)^t$$

Monthly

$$P_t = P_0\left(1+\frac{R}{12}\right)^{12t}$$

- Weekly

$$P_t = P_0\left(1+\frac{R}{52}\right)^{52t}$$

Continuous

$$P_t = P_0 \lim_{n\to\infty}\left(1+\frac{R}{n}\right)^{nt} = P_0 e^{Rt}$$

This is useful for multi-period returns where we can add continuous returns

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

Log Returns

Log returns are convenient and have some statistical advantages over simple returns. We will use letter r instead of R when referring to log returns. One period log (continuously compounded) return is:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln P_t - \ln P_{t-1} = \ln(1+R_t)$$

E.g. buy 1 share in AMP at end June 2014 for \$5.968 per share and sell it at end June 2015 for \$6.470 per share

- r = ln(6470)– ln(5968) = 0.0807 or 8.07%
- Notice that 6.470 = 5.968exp(0.0807)
- Remember that the simple return R was 8.41%

A Summary of the Relationships:

If the returns are in decimals, *e.g.* 0.1:

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

If the returns are in percentages, *e.g.* 10%:

$$r_t = 100\ln\left(1+\frac{R_t}{100}\right) \qquad R_t = 100\left(e^{r_t/100} - 1\right)$$

Temporal aggregation (compounding over time):

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

Example:

If the monthly log return of an asset is 4.46%, then the corresponding simple monthly return is

$$100[\exp(4.46/100) - 1] = 4.56\%$$

If the monthly log returns of the asset within a quarter are 4.46%, -7.34% and 10.77% then the quarterly log return of the asset is

$$(4.46 - 7.34 + 10.77) = 7.89\%$$

Random Variables

The return on a stock over the next period is uncertain today – it is a random variable. Without the uncertainty, the problems in finance would reduce to basic microeconomics.

Basic Probability and Random Variables

Experiments:
A random experiment is an action or observation whose outcome is uncertain in advance of its occurrence.

E.g: toss a coin, spin a roulette wheel, observe a share price

Sample Space S:
Is the set of all possible outcomes of an experiment

- E.g. 1 experiment: toss a coin twice $S = \{HH, HT, TH, TT\}$
- E.g. 2 experiment: observe a share price $S = [0, \infty)$

Rule for Independence:

P(HT) = P(H and T) = P(H) x P(T)
→Because coin flips are independent

Event:
A subset of the sample space
E.g. Simple Event A – two heads $A = H$, Composite Event B – two heads or two tails $B = \{HH \cup TT\}$ where $\cup$ is the union operator, it means "or"

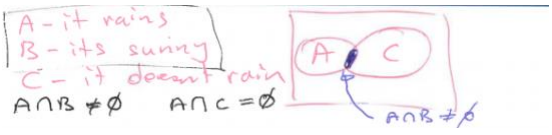
Rule for Mutually Exclusive Events:

P(B)= P(HH U TT) = P(H & H) or P(T & T)
= P(H & H) + P(T & T)
Outcomes only happen once→ Mutually Exclusive

Broadly speaking probability expresses a degree of belief a person has about an event/statement by a number between zero and one. Probabilities should follow the following three rules (axioms):

1. $0 \leq P(A) \leq 1$ for any event A
2. $P(S) = 1$ where S is the sample space
3. If $\{A_i\}$, $i = 1,2, \ldots$ are mutually exclusive events , i.e., $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$

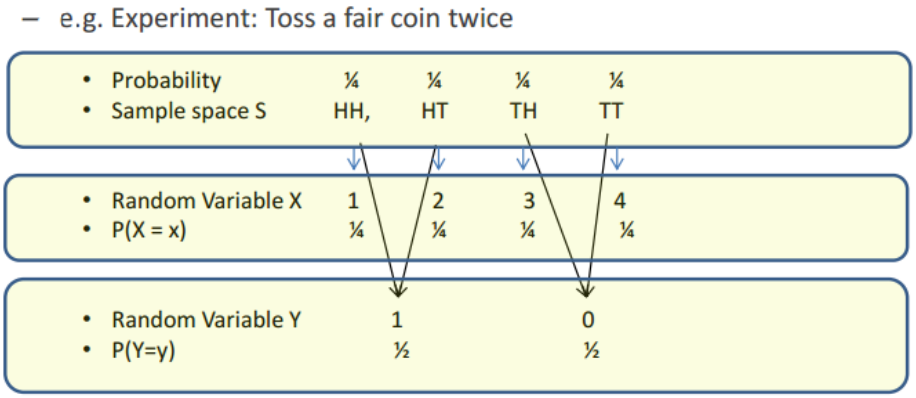
- Axiom 1: What is the probability of picking any one card from a deck of cards (e.g. Queen of Hearts)?
  - Answer: 1/52>0
- Axiom 2: What is the probability of selecting any card?
  - Answer: $P(S) = 1$
- Axiom 3: What is the probability of selecting one card that is either a queen of hearts, or a king of diamonds, or a ten of spades?
  - Answer: $P(Q_h \cup K_d \cup 10_s) = P(Q_h) + P(K_d) + P(10_s)$
    $= \frac{1}{52} + \frac{1}{52} + \frac{1}{52} = \frac{3}{52} = 0.06$



A random variable is a real-valued function defined over a sample

Discrete Random Variables

A discrete random variable is a variable that takes a countable number of real numbers with certain probabilities.



- Explain in words the definitions of X and Y.
  - P(HH or HT) = P(HH) + P(HT) →Mutually exclusive = P(H)P(H) + P(H)P(T)
  - = ½ ½ + ½ ½ = ½
- Note: in almost all problems we can forget about the original sample space and pay attention only to what values a random variable takes with what probabilities.

Continuous Random Variables

A random variable X is continuous if its CDF is a continuous function.

Note: – for a continuous random variable we always have $P\ X = x = 0$

we can calculate $P(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a)$
$f(x) = dF(x)/dx = F'(x)$ is the probability distribution (density) function (PDF)

Probability Distributions

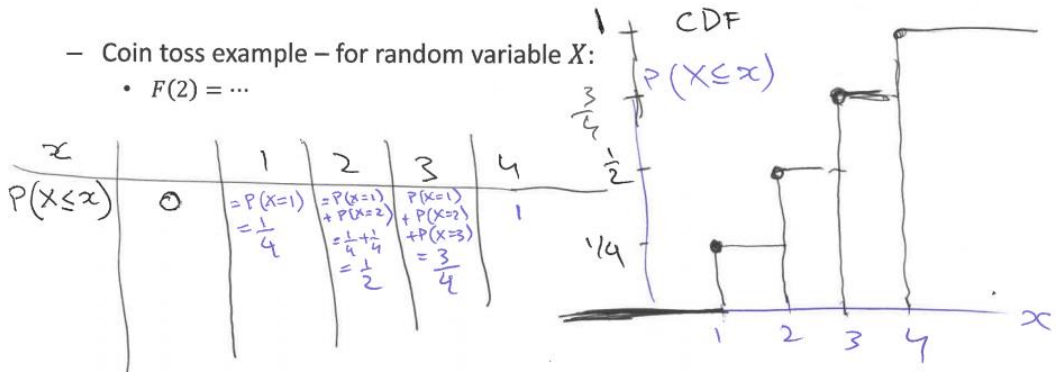
A discrete probability distribution function (PDF) gives all possible realizations of X and the probability that each value occurs.

A discrete probability distribution is a table, graph or rule linking a probability ($X = x_i$) with all possible realizations $X = x_i$.

Cumulative Distribution Function

For any random variable X, the cumulative distribution function (CDF) gives the probability that X is less than or equal to a specific value.

$(x) = P(X \leq x)$
where X is a random variable, and x is any real



The Normal Distribution

The normal distribution of a random variable X with mean μ and variance σ^2 is denoted by X~(μ, σ^2) Its probability density function (PDF) is given by:

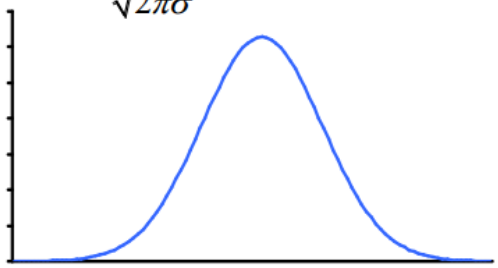
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$$

A normal random variable X can be standardized by subtracting the mean and dividing by the standard deviation:

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

A r.v. X that is Normally distributed is an example of a continuous random variable

$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<\infty$$



The Normal distribution is:

- Bell shaped
- Symmetric (Skewness = 0)
- Kurtosis = 3

Sometimes returns are assumed to be normally distributed. Assuming normality has the following advantages:

- Only the mean and variance are needed to describe the distribution (if we know the mean and variance we can compute all probabilities)
- Estimation and testing become easier

The disadvantages are:

- The true distribution of simple net returns cannot go below -1 (-100%)
- This happens when the value of your asset goes to zero

Products of normal variables are not normal so if R_t is normal then $R(k)$, which is the k period return is not. For example, if daily returns are normal then weekly or monthly returns cannot be normal. Financial data do not look normal in many instances. There are tests for normality which suggest that the data is not normal.