

Empirical Finance: Class 1 (Preliminaries)

Hormoz Ramian
Imperial College London
Department of Finance

Introduction

Hormoz Ramian

- Year 4 Finance working on banking, regulation and macroeconomics
- Office Hours: Thursdays 3–4pm
- Email: `h.ramian@imperial.ac.uk`

Today's Outline

- Returns
- Macroeconomic indicators
- Useful approximations
- Exercises

Definition

Simple Return — Suppose time is discrete, with dates $t = 0, 1, 2, \dots$ and the prices of an investment in a financial asset, which pays no intermediate coupon or dividend, between two consecutive periods t and $t + 1$ are P_t and P_{t+1} , respectively, then we can construct a one-period *simple net return* as,

$$r_{t+1} = \frac{\Delta P_{t+1}}{P_t} = \% \Delta P_{t+1}$$

- shows the net investment gain or loss between two periods
- unitless as numerator and denominator are measured with same units e.g. dollars
- simple *gross* return including original investment plus gain or loss

$$R_{t+1} = \frac{P_{t+1}}{P_t}$$

Logarithmic Return — Convenient to transform into logarithmic units

$$\ln R_{t+1} = \Delta \ln P_{t+1}$$

A stock price that moves from 100 today to 110 tomorrow yields a simple return of $\frac{110-100}{100} = 0.10$, similarly, the log return is $\log(110) - \log(100) = 4.7005 - 4.6052 = 0.0953$.

Simple vs log Return

1. Move with each other and are close.
2. Yield different numbers *if* the ratio of two consecutive prices is far from one.
3. Consider expanding a logarithmic function $\ln R_{t+1}$ in the neighbourhood of 1.00:

$$\ln R_{t+1} = \underbrace{\ln(1) + \frac{R_{t+1} - 1}{1}}_{\text{1st order approx}} - \frac{(R_{t+1} - 1)^2}{2} + \dots \approx R_{t+1} - 1 = r_{t+1}$$

2nd order approx

where \approx indicates that we approximated the logarithmic function by ignoring terms beyond higher order to establish that a net simple return is approximately equal to a (net) logarithmic return $r_{t+1} \approx \ln R_{t+1}$ when relative price variations are small.

4. Note that derivative of logarithmic function $f(x) = \ln(x)$ is $f_x = \frac{dx}{x}$. In fact, approximation of the logarithmic function $f(x)$ implies that $\ln x \leq x - 1$ as $x - 1$ is the tangent to $f(x)$ at $x_0 = 1$ but above $f(x)$ elsewhere because the second derivative term is always negative.
5. The economic interpretation is that lower and higher returns are subject to an approximation error of the same sign.

Real and Nominal

Macroeconomic variables (e.g. asset prices) are observed in terms of monetary units:

1. A one-period *nominal return* on an investment shows gains (losses) in terms of its traded monetary unit, but ignores the gains and losses of the monetary unit itself.
2. A one period *real return* measures investment's gain or loss in terms of purchasing power that it delivers in time- t monetary units (time- t dollars).

Variations in the value of each monetary unit is measured by the *percentage change* in the *general level of prices*, or the inflation rate.

- General level of prices refers to an average of one-period price change of a *basket of goods* and services such as Consumer Price Index (CPI).
- We can formulate the real return in the following way:

$$R_{t+1} = \frac{P_{t+1}/Q_{t+1}}{P_t/Q_t} = \underbrace{\frac{1}{(1 + \pi_{t+1})}}_{\text{Inflation Rate}} \times \underbrace{\left(\frac{P_{t+1}}{P_t}\right)}_{\text{Nominal Return}} = \frac{R_{t+1}^{\$}}{1 + \pi_{t+1}}$$

P_{t+1} is the nominal asset price, Q_t is the average price of a basket of goods.

- Real return is a measurement in terms of a real *consumption good*, for instance, an apple at date- t or at date- $t + 1$.

Simple gross nominal return:

$$\begin{aligned}\text{nominal} &= \text{real} \times \text{inflation} \\ 1 + r_{t+1}^{\$} &= (1 + r_{t+1})(1 + \pi_{t+1}) \\ &= 1 + r_{t+1} + \pi_{t+1} + r_{t+1} \times \pi_{t+1} \\ &\Rightarrow \\ r_{t+1}^{\$} &\approx r_{t+1} + \pi_{t+1}\end{aligned}$$

where $r_{t+1} \times \pi_{t+1}$ is small then:

Excess Return

Consider the *nominal* price of a risk-free asset (e.g. Treasuries), B_t , then:

$$\log R_{f,t+1} = \log \frac{B_{t+1}}{B_t} - \log \frac{Q_{t+1}}{Q_t}$$

is the (real, log) risk-free rate, and then:

$$\begin{aligned}\log X_{t+1} &= \log R_{t+1} - \log R_{f,t+1} \\ &= \left(\log \frac{P_{t+1}}{P_t} - \log \frac{Q_{t+1}}{Q_t} \right) - \left(\log \frac{B_{t+1}}{B_t} - \log \frac{Q_{t+1}}{Q_t} \right) \\ &= \log \frac{P_{t+1}}{P_t} - \log \frac{B_{t+1}}{B_t} \\ &= \log R_{t+1}^{\$} - \log R_{f,t+1}^{\$}\end{aligned}$$

is the (real, log) excess return.

Multi-Period Return

A return on an investment made at date- t and held over horizon longer than one period is a *multi-period return*. Suppose the price of an investment today is P_t , then

1. the (gross) simple multi-period return over the next n periods is:

$$\begin{aligned} R_{t,t+n} &= \frac{P_{t+1}}{P_t} \times \frac{P_{t+2}}{P_{t+1}} \times \dots \times \frac{P_{t+n-1}}{P_{t+n-2}} \times \frac{P_{t+n}}{P_{t+n-1}} = \frac{P_{t+n}}{P_t} \\ &= (1 + r_{t+1}) \times (1 + r_{t+2}) \times \dots \times (1 + r_{t+n-1})(1 + r_{t+n}) = \prod_{i=1}^n (1 + r_{t+i}) \end{aligned}$$

where the notation $R_{t,t+n}$ indicates return between n consecutive periods from t to $t + n$ (notation varies e.g. $R_t[n]$).

2. the multi-period log return is:

$$\begin{aligned} \ln R_{t,t+n} &= \ln \left(\frac{P_{t+1}}{P_t} \times \frac{P_{t+2}}{P_{t+1}} \times \dots \times \frac{P_{t+n-1}}{P_{t+n-2}} \times \frac{P_{t+n}}{P_{t+n-1}} \right) = p_{t+n} - p_t \\ &= \ln R_{t+1} + \ln R_{t+2} + \dots + \ln R_{t+n-1} + \ln R_{t+n} = \sum_{i=1}^n \ln R_{t+i} \end{aligned}$$

Suppose that the price of an investment held over n periods varies over time, then each one-period return differs from the following one. A *constant annualized return* is an average (geometric) of consecutive (gross) returns that the investment yields per period:

$$(R_{t,t+n})^{\text{Annualized}} = \left[R_{t,t+n} \right]^{\frac{1}{n}} = \left[\prod_{i=1}^n (1 + R_{t+i}) \right]^{\frac{1}{n}}$$

Similarly, in case of a multi-period log return¹, the following (arithmetic) average describes the per period average (net) log return,

$$(\ln R_{t,t+n})^{\text{Annualized}} = \frac{1}{n} \ln R_{t,t+n} = \frac{1}{n} \sum_{i=1}^n \ln R_{t+i}$$

Macroeconomic data are often observed at various frequencies. For example, GDP and Inflation are reported on monthly and quarterly bases, respectively, therefore adjusting returns of non-synchronous frequencies is a necessary step to obtain like-for-like changes of data over identical times periods.

Discounting

Suppose project A begins with an initial cost of £1000 and produces the following cashflow over five years when it terminates:

Cashflow (Project A)	-£1000	£0	£0	£300	£600	£900
Year	0	1	2	3	4	5

The NPV of this project, given cost of capital r is:

$$\text{NPV}_A(r) = -1000 + \frac{0}{1+r} + \frac{0}{(1+r)^2} + \frac{300}{(1+r)^3} + \frac{600}{(1+r)^4} + \frac{900}{(1+r)^5}$$

Consider the value of this discounted stream at $r_L = 0.08$ and $r_H = 0.18$:

$$\text{NPV}_A(r_L) = +£291.69$$

$$\text{NPV}_A(r_H) = -£114.54$$

Because the NPV depends on cashflow together with cost of capital r , the value changes suggesting that somewhere between r_L and r_H project A's evaluation changes sign.

House Prices

Case Study

Suppose that today's news headline mentions that real estate prices are growing at a fast pace, and that today's price of a certain apartment in London is £180,000. We know from our records that the price of the very same apartment (with identical quality) was £120,000 in 2004. In this scenario prices are measured in pounds but, within respective ongoing economic conditions in 2004 and 2018, and thus the asset value is in nominal prices. The inflation rate in the UK has been modestly increasing during this period and we wish to distinguish between the real gain in the value of the apartment and variations in the value of each pound. For example, due to more population and demand for this apartment, the price has increased, and we wish to know how much purchasing power is delivered by this real estate investment in 2004-terms. We need to deflate the apartment's price by the price deflator index and multiplying by 100 (under the assumption that the deflator has a base value of 100).

Year	General Price (2004 Levels)	Inflation (%)	Real Prices (2004 Levels)	Nominal Prices (Observed)	Nominal Return (%)	Real Return (%)
2003	99.7	—	119,644	119,376	—	—
2004	100	0.30	120,000	120,000	0.52	0.22
2005	100.8	0.80	120,720	121,686	1.40	0.60
2006	101.21	0.41	122,038	122,534	0.70	0.29
2007	101.81	0.59	123,080	123,810	1.04	0.45
2008	103.03	1.20	124,882	126,378	2.07	0.88
2009	107.47	4.31	130,084	135,690	7.37	3.06
2010	112.62	4.79	139,831	146,532	7.99	3.20
2011	117.35	4.20	150,167	156,474	6.78	2.58
2012	124.75	6.31	161,809	172,013	9.93	3.62
2013	127.62	2.30	174,039	178,043	3.51	1.20
2014	132.85	4.10	181,599	189,041	6.18	2.08
2015	133.91	0.80	189,760	191,274	1.18	0.38
2016	131.37	-1.90	189,522	185,927	-2.80	-0.90
2017	126.9	-3.40	182,758	176,539	-5.05	-1.65
2018	128.55	1.30	177,690	180,000	1.96	0.66

Dividend Yield

Dividend-Price Ratio

Dividend per share, divided by the price per share (real variable). Let D_t denote dividend, then our observed variables are:

1. Stock return without dividend is (simple, nominal): $R_{t+1} = P_{t+1}/P_t$
2. Stock return with dividend is (simple, nominal): $R_{t+1}^d = (P_{t+1} + D_{t+1})/P_t$

Decompose return (dividend included):

$$R_{t+1}^d = \frac{P_{t+1} + D_{t+1}}{P_t} = \underbrace{\frac{D_t}{P_t}}_{\text{Dividend Yield}_t} \cdot \underbrace{\frac{D_{t+1}}{D_t}}_{\text{Dividend Growth}} + \underbrace{\frac{P_{t+1}}{P_t}}_{\text{capital gain}}$$

Divide gross returns:

$$\begin{aligned}\frac{R_{t+1}^d}{R_{t+1}} &= \frac{P_{t+1} + D_{t+1}}{P_t} / \frac{P_{t+1}}{P_t} = 1 + \frac{D_{t+1}}{P_{t+1}} \\ \frac{D_{t+1}}{P_{t+1}} &= \frac{R_{t+1}^d - R_{t+1}}{R_{t+1}} = \frac{R_{t+1}^d}{R_{t+1}} - 1\end{aligned}$$

D_{t+1}/P_{t+1} is the dividend yield (numerator and denominator are nominal)

Dividend Growth

Gordon Dividend Growth Model: Consider a stock is pledged to pay a dividend per share D which grows at a fixed rate g at every subsequent payout date, and that the interest rate is fixed and equal to r ($r > g$) then the price of the stock today is:

$$P_0 = \frac{(1+g)D}{1+r} + \frac{(1+g)^2D}{(1+r)^2} + \frac{(1+g)^3D}{(1+r)^3} + \dots = \sum_{t=1}^{\infty} \frac{(1+g)^t}{(1+r)^t} D$$

re-arrange:

$$P_0 = \underbrace{\frac{1+g}{1+r}}_x \times \underbrace{\left(1 + \frac{1+g}{1+r} + \frac{(1+g)^2}{(1+r)^2} + \dots\right)}_{\text{geometric sum}} \times D = x \times \frac{1}{1-x} \times D$$

note:

$$\frac{x}{1-x} = \frac{\frac{1+g}{1+r}}{1 - \frac{1+g}{1+r}} = \frac{1+g}{r-g}$$

dividend value tomorrow is $D' = (1+g)D$, thus the price today increases when D or g increase or r decrease:

$$P_0 = \frac{1+g}{r-g} D = \frac{D'}{r-g}$$

Stock and Flow

Definition

1. A *stock* is measured at one specific time, and represents a quantity existing at that point in time (say, December 31, 2004), which may have accumulated in the past.
2. A *flow* variable is measured over an interval of time: measured per unit of time (year).

Examples

- the U.S. nominal capital stock is the total value, in dollars, of equipment, buildings, and other real productive assets in the U.S. economy, and has units of dollars.
- the U.S. nominal gross domestic product refers to a total number of dollars spent over a time period, such as a year, is a flow variable, and has units of dollars/year.

Stock	Flow
bank balance (yen)	interest or withdrawals (yen per month)
housing stock (dollars)	housing investment or depreciation (dollars per year)
equity shareholdings (shares)	purchases or sales of shares (shares per month)

- Mapping from stock to flow $Y = AK^\alpha L^{1-\alpha}$

GDP Measurement

Definition: Gross domestic product is, the total monetary value, of all final goods and services produced within the territory of a country, over a particular period of time (quarterly or annually). There are three approaches to measure GDP:

1. **Expenditure Approach:** The sum of all final expenditures within the economy, that is, all expenditure on goods and services which are not used up or transformed in a productive process.
2. **Production Approach:** Concerned with the generation of value added. In other words, the value of all goods and services produced within the economy.
3. **Income Approach:** Sums all income generated by production activity, also known as factor incomes.

notes:

- The methodology to measure GDP is based on *national accounting*.
- The U.S. Bureau of Economic Analysis considers the source data for expenditure components to be more reliable than for either income or production components.

Inflation Measurement

GDP Deflator

- time-varying basket of goods and services
- includes only domestically produced goods
- purchases by government or firms are captured
- Paasche Index
- $\text{GDP Deflator} = \text{Nominal GDP} / \text{Real GDP} \times 100$

Consumer Price Index

- a fixed basket of goods (updated infrequently)
- includes anything *bought* by *consumers* including foreign good
- weights reflect importance of each item (e.g. chicken vs. caviar)
- Laspeyres index

Retail Price Index

- more focused on household sector (used for indexation of pensions, state benefits, etc.)
- includes items such as council tax, housing mortgage payments
- for example, RPI rises when council tax or mortgage interests rise, even though underlying inflationary pressures in the economy stay the same.

Producer Price Index

- tracks the prices of goods bought and sold manufacturers
- including price indices of materials and fuels purchased
- methodology is similar to CPI but basket items are focused on output sector

Notes:

- Measure of percentages price change with respect to a specific base year, for different baskets of goods
- Bank of England and the ECB attempt to achieve a target rate of inflation as measured by the CPI
- CPI and GDP deflators move very closely, however, the CPI tends to overstate inflation.
- For example, the CPI is subject to the substitution bias as it uses a fixed basket, therefore it does not reflect the ability of consumers to substitute cheaper goods for more expensive ones. When relative prices change, the true cost of living rises less rapidly than the CPI

Provider	Data	Source
FRED	Macro Indicators	fred.stlouisfed.org/
CRSP	US Stocks	crsp.com
Commodity Systems Inc.	Futures	csidata.com
Datastream	Stocks, bonds, currencies, etc.	datastream.com/product/has/
IFM	Futures, US Stocks	theifm.org
Olsen & Associates	Currencies, etc.	lsen.ch
Trades and Quotes	US Stocks	yse.com/marketinfo
US Federal Reserve	Currencies, etc.	federalreserve.gov/releases/

Compound Interest

Annually — Consider an initial investment at date-0 of amount x_0 that offers a constant net interest rate $r > 0$ per year

- interest rate is fixed, known and risk-free
- initial investment one year later is worth $x_1 = (1 + r)x_0$ with net payoff is rx_0
- repeated one year after to delivers

$$x_2 = (1 + r)x_1 = (1 + r)^2x_0 \quad (1)$$

which implies that the net payoff rx_0 is reinvested in together with the initial investment amount.

- In general, if we repeat re-investment year after year, then at date- n (years), the investment grows to

$$x_n = (1 + r)^n x_0 \quad (2)$$

Re-investing the intermediate payoffs at the end of each year is known as compounding interest and in this example is compounded annually.

Biannually — When the payoff from this investment is collectable biannually (with intermediate re-investments), then after the first six months obtain $x_0(1 + 0.5r)$, and thus obtain:

$$x_1 = (1 + 0.5r)(1 + 0.5r)x_0 = (1 + 0.25r^2 + r)x_0 > (1 + r)x_0$$

the overall payoff increases after one year.

***k*-times**

This additional gain due to intermediate re-investments increases, as the frequency of collections increases e.g. to quarterly, monthly, etc. and in general, if the time period is of size $\frac{1}{k}$ and compounding take place k -times per year, at a linear rate $\frac{r}{k}$ per period, then the value of investment after one year is:

$$x_1(k) = \left(1 + \frac{r}{k}\right)^k x_0$$

which by repeating one year ahead yields $\left(1 + \frac{r}{k}\right)^{2k} x_0$ and after n -year $\left(1 + \frac{r}{k}\right)^{nk} x_0$.

- For example, re-investment on monthly basis yields $\left(1 + \frac{r}{12}\right)^{12n} x_0$
- Over an arbitrary interval of time $(0, t]$, compounding k -times implies re-investment every $\frac{t}{k}$ units of time at linear rate $\frac{rt}{k}$ which yields the following amount at time t :

$$x_t(k) = \left(1 + \frac{rt}{k}\right)^k x_0$$

When it is possible to repeat compounding over an infinitesimally small time interval $k \rightarrow \infty$, the re-investment is called *continuous compounding* which delivers² as $k \rightarrow \infty$,

$$\lim_{k \rightarrow \infty} x_t(k) = \lim_{k \rightarrow \infty} x_0 \left(1 + \frac{rt}{k}\right)^k = \exp(rt)x_0$$

The economic interpretation is that the time- t value of a continuously compounded initial investment of x_0 during $t \in (0, t]$ is worth $\exp(rt)x_0$.

Frequency	Payments	Interest Rate per Period	Value after One Year	EIR
Annual	1	0.1	\$1.10000	10.00%
Biannual	2	0.05	\$1.10250	10.25%
Quarterly	4	0.025	\$1.10381	10.38%
Monthly	12	0.0083	\$1.10471	10.47%
Weekly	52	0.1/52	\$1.10506	10.50%
Daily	365	0.1/365	\$1.10516	10.51%
Continuously	∞	—	\$1.10517	10.52%

Doubling the Investment

Given an annual fixed and risk-free interest rate $r = 4\%$ and an initial investment capital x_0 , it takes $t \approx 17$ years to generate 100% net gain on initial investment. This is because $\exp(rt)x_0 = 2x_0$ yields $t = 0.6931/r \approx 17$.

Taylor Approximation

Definition: Let $f(z)$ be any real-valued continuous (nonlinear) function then we can approximate $f(z)$ around a (i) *given point* z_0 , up to a certain (ii) *order* n if the function is n -times differentiable:

$$f(z) \approx \underbrace{f(z_0) + f'(z_0)(z - z_0)}_{\text{1st order } (n = 1)} + \underbrace{\frac{1}{2}f''(z_0)(z - z_0)^2 + \dots}_{\text{2nd order } (n = 2)} \quad (3)$$

where $f'(\cdot)$ and $f''(\cdot)$, the first and second order derivatives of $f(\cdot)$, respectively, exist. The first order provides a linear approximation and the second order a quadratic approximation of $f(z)$, and so on. The approximation expresses the original nonlinear function in terms of base polynomials and the approximation error decreases as further higher order base polynomials are considered.

Example (Log):

Let $f(r) = \log(1 + r)$ then $f'(r) = \frac{1}{r}$ and $f''(r) = -\frac{1}{r^2}$ then approximating f around point $r_0 = 0$ up to the first order yields,

$$\begin{aligned} f(1 + r) &\approx f(1 + r_0) + f'(1 + r_0)(r - r_0) \\ &\approx \log(1 + 0) + \frac{1}{1 + 0}(r - 0) \\ &\approx r \end{aligned}$$

this implies that in a small neighbourhood of 1.00 then $\log(1 + r) \approx r$ if $r \approx 0$, for instance when (returns or interest rate) $r = 0.01$ then:

$$\log\left(1 + \frac{1}{100}\right) \approx \frac{1}{100}$$

Example (Fractional):

Let $f(r) = \frac{1}{1+r}$ then $f'(r) = -\frac{1}{(1+r)^2}$ then approximating f around point $r_0 = 0$ up to the first order yields,

$$\begin{aligned} f(r) &\approx \frac{1}{1+r_0} - \frac{1}{(1+r_0)^2}(r-r_0) \\ &\approx 1-r \end{aligned}$$

which is applicable when evaluating the present value of the face value of a £100 bond that is receivable one year from now, at interest rate $r = 0.01$ gives:

$$\begin{aligned} \text{PV} &= \frac{\pounds 100}{1+r} \\ &\approx \pounds 100 \times (1-r) \\ &\approx \pounds 100 \times 0.99 \end{aligned}$$

that is the price of the bond today.

Example (Power):

Let $f(r) = (1 + r)^n$ then $f'(r) = n(1 + r)^{n-1}$ then approximating f around point $r_0 = 0$ up to the first order yields,

$$\begin{aligned} f(r) &\approx (1 + r_0)^n + n(1 + r_0)^{n-1}(r - r_0) \\ &\approx 1 + nr \end{aligned}$$

which is applicable to the case in which a fixed-income security pays a constant interest rate r over n number of years, then the forward value e.g. of a £100 over 25 years given the interest rate of 1% is $\pounds 100 \times (1 + 25 \times 1\%) = \pounds 125$.

Example (Exponential):

Consider the following,

$$\log \left(\frac{P_{t+1}}{P_t} \right) = r$$

exponentiating both sides yields $P_{t+1}/P_t = \exp(r)$ which can be re-written as $P_{t+1}/P_t - 1 = \exp(r) - 1$, then let $h(r) = \exp(r)$ and approximating around point $r_0 = 0$ up to the first order by applying (3) and knowing that $h(r) = h'(r) = h''(r)$ yields:

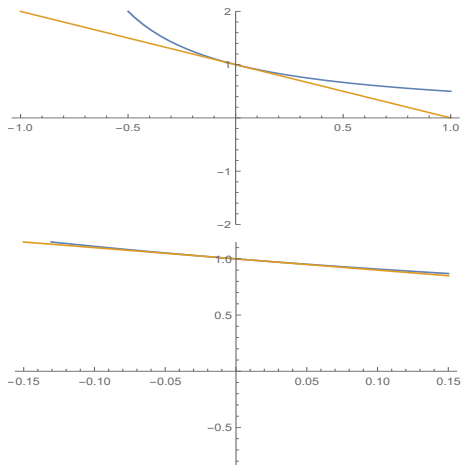
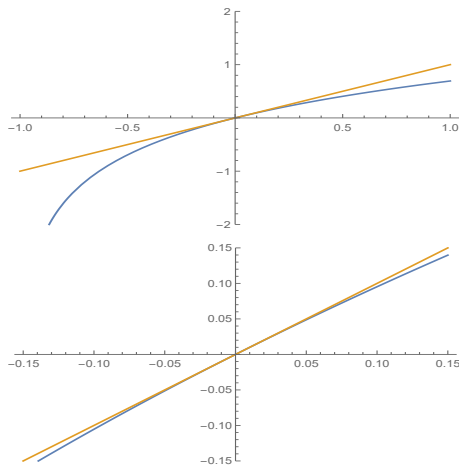
$$\begin{aligned} \exp(r) &\approx \exp(r_0) + \exp(r_0)(r - r_0) \\ &\approx 1 + 1 \times (r - 0) \\ &\approx 1 + r \end{aligned}$$

then the expression $(P_{t+1} - P_t)/P_t = \exp(r) - 1$ can be substituted for in the following way:

$$\begin{aligned} (P_{t+1} - P_t)/P_t &= \exp(r) - 1 \\ &\approx (1 + r) - 1 \\ &\approx r \end{aligned}$$

Use Taylor expansion (order one, around zero):

1. $\log(1+x) \approx x$ when $x \approx 0$
2. $\frac{1}{1+x} \approx 1-x$ when $x \approx 0$
3. $\frac{1+y}{1+x} \approx 1+y-x$ when $x, y \approx 0$ where $(1+y)(1-x) = 1+y-x-xy$ but $x \times y$ negligible



Basics

Revision from Financial Econometrics (Fall Term), e.g. handout 5. Particularly, knowledge of the following is assumed in this course:

- Unit-root and stationarity
- Vector Autoregressive
- Co-integration

Estimation with R implementation is preferred.

Stationarity and Unit Root Test

(Review from Econometrics Course)

Time-series econometrics requires stricter conditions to ensure that estimation and inference methods remain valid when working with time-series data:

- Time-invariant distribution (time invariant moments e.g. mean, var, etc.)
Prices, consumption, labour income, etc. are non-stationary
- Unit root test examines if the underlying distribution of a *single* times-series remains invariant across time

$$H_0 : I(1) \quad \text{Non-stationary} \quad (4)$$

$$H_A : I(0) \quad \text{Stationary} \quad (5)$$

A stationary distribution (observed macro time-series) is:

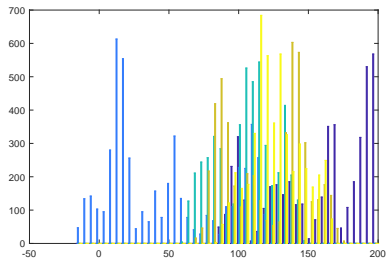
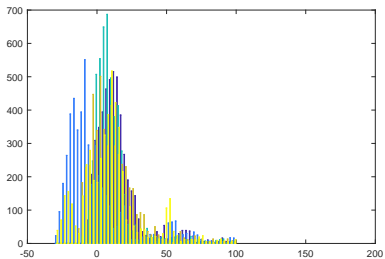
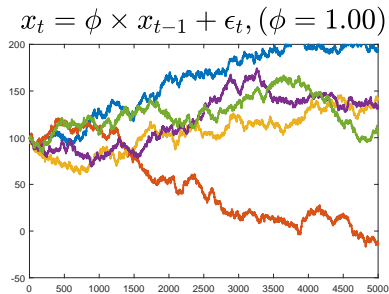
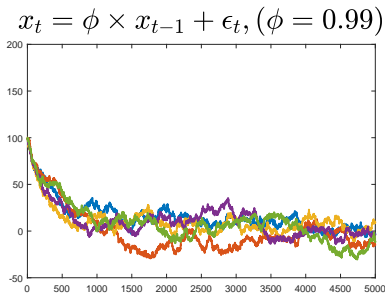
1. non-stationary when a shock has a permanent effect
2. stationary when a shock only has a transitory effect

Two possible conclusions (null hypothesis):

- Reject: Evidence to distinguish from $I(1)$ (test aims at rejecting the null)
- Fail to reject: cannot distinguish from $I(1)$ (need to check $I(2)$ vs. $I(1)$, ...)

Stationary vs. Non-Stationary Distribution (Time-Series)

(Review from Econometrics Course)



Dickey-Fuller Distribution (Test-Statistic)

(Review from Econometrics Course)

ADF test is one approach to examine stationarity using the following test-statistic:

$$\text{ADF-test-stat} = \frac{\hat{\phi} - 1}{\sigma_{\hat{\phi}}} \sim \text{DF Distribution (one tail)} \quad (6)$$

Large-Sample Critical Values of the Augmented Dickey-Fuller Statistic for Testing for a Unit Root:

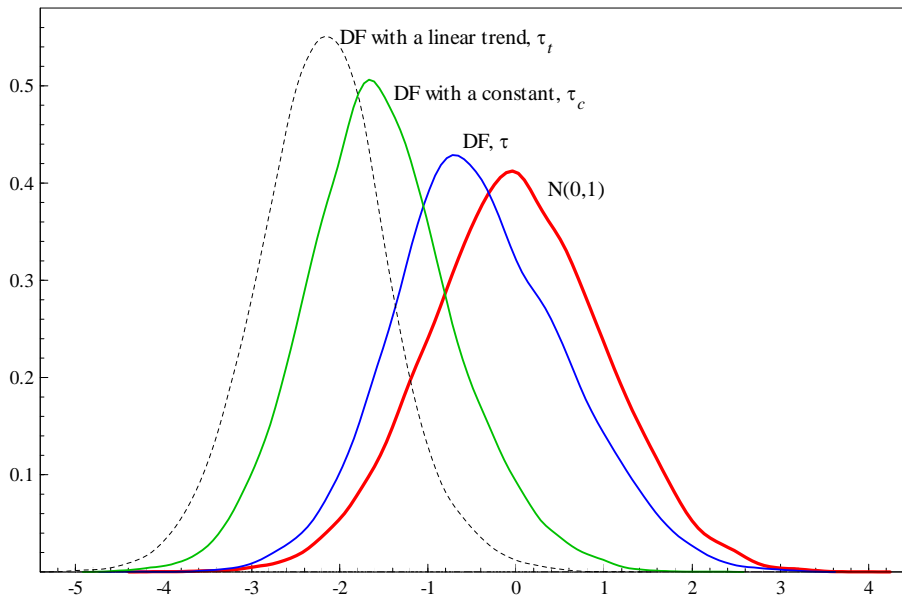
	10%	5%	1%
Intercept only (τ_c)	-2.57	-2.86	-3.43
Intercept & time trend (τ_t)	-3.12	-3.41	-3.96

Question

Why stationarity could be an important consideration to a policy maker?

Dickey-Fuller Distribution (Test-Statistic)

(Review from Econometrics Course)



Co-integration

An ideal setup to study economic variables is to use them on level (without differencing, etc.). However, when variables are non-stationary on level, they amount to spurious regressions, unless one can establish a co-integrating relationship:

$$C_t = \phi_1 + \phi_2 A_t + \phi_3 Y_t + \epsilon_t \quad (7)$$

- Individual variables, C_t , A_t , Y_t are non-stationary (fail to reject ADF unit root)
- Their *combination* $\epsilon_t = C_t - \phi_1 - \phi_2 A_t - \phi_3 Y_t$ can be tested:

$$H_0 : I(1) \quad \text{Non-stationary} \quad (8)$$

$$H_A : I(0) \quad \text{Stationary} \quad (9)$$

- Significance level $\alpha : 5\%$ and test against Dickey-Fuller distribution
- Rejection of null hypothesis implies ruling out non-stationarity: linear combination of series ϵ_t is stationary, thus C_t , A_t , Y_t follow a long-term equilibrium relationship

VARX: Vector Autoregression with Exogenous Variable

Let's examine how cay_t could explain one-step ahead variations of:

$$(\Delta \log C_t, r_t - r_{f,t}, \Delta \log Y_t) \quad (10)$$

One statistical model is to use the following (reduced form) VAR with an additional exogenous variable:

$$\begin{bmatrix} \Delta \log C_t \\ r_t - r_{f,t} \\ \Delta \log Y_t \end{bmatrix} = \mathbf{A}_0 + \underbrace{\mathbf{A}_1 \begin{bmatrix} \Delta \log C_{t-1} \\ r_{t-1} - r_{f,t-1} \\ \Delta \log Y_{t-1} \end{bmatrix} + \dots + \mathbf{A}_3 \begin{bmatrix} \Delta \log C_{t-3} \\ r_{t-3} - r_{f,t-3} \\ \Delta \log Y_{t-3} \end{bmatrix}}_{\text{Number of Lags}} + \beta cay_{t-1} + \varepsilon_t$$

- \mathbf{A}_0 is 3×1 vector of constants, \mathbf{A}_i is 3×3 matrix of autoregressive coefficients, β is 3×1 vector of exogenous process coefficients
- VAR's endogenous vector $(\Delta \log C_t, r_t - r_{f,t}, \Delta \log Y_t)$
- cay_t is an exogenous variable (excluded from the VAR vector)
- Determine optimal lag structure (SC, HQC, etc.)
- Lags capture autocorrelations thus ε_t satisfies classical assumptions
- VAR is different from using three single-equation regressions (with identical variables) e.g. error terms are allowed to correlate.

VARX: Some Background

(From Econometrics Course: not part of this course)

Optimal Lag Structure

Kilian (2005) suggests the following practitioner's guide to lag selection:

- Schwarz (SC): Quarterly VAR models with sample size > 120
- Hannan-Quinn (HQ): Quarterly VAR models with sample size < 120

Estimation

VAR is a linear model thus least squares method is readily applicable. Let $X_t = (\Delta \log C_t, r_t - r_{f,t}, \Delta \log Y_t)$, then:

$$X_t = \mathbf{A}_0 + \mathbf{A}_1 X_{t-1} + \dots + \mathbf{A}_K X_{t-K} + \beta cay_{t-1} + \varepsilon_t$$

Denote $\Theta = [\mathbf{A}_0 \ \mathbf{A}_1 \ \dots \ \mathbf{A}_K \ \beta]$, $Z = [\mathbf{1} \ X_{t-1} \ \dots \ X_{t-K} \ cay_{t-1}]$, and re-arrange:

$$X_t = \begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_1 & \dots & \mathbf{A}_K & \beta \end{bmatrix} \begin{bmatrix} \mathbf{1} & X_{t-1} & \dots & X_{t-K} & cay_{t-1} \end{bmatrix} + \varepsilon_t = \Theta Z + \varepsilon_t$$

Using *Vec* operator and Kronecker product (\otimes):

$$Vec(X_t) = Vec(\Theta Z) + Vec(\varepsilon_t) \quad (11)$$

Least squares estimation for Θ is $\hat{\Theta} = [(ZZ')^{-1}Z \otimes \mathbf{I}]X$.

Q — Assume that both Mexico's and Argentina's average annual per capita GDP growth rates are 3 percent per year, and both countries began with an initial per capita GDP of \$1,000. However, Argentina has been growing since 1940 and Mexico only since 1965. In 2015, Mexico's per capita GDP would have been about ..., while Argentina's would have been about

1. \$19.42; \$12.94
2. \$48,544; \$72,816
3. \$9,179; \$4,384
4. \$4,384; \$9,179
5. \$51,500; \$77,250

Q — If the nominal GDP rises by 6 percent and the price level rises by 3 percent, then the real GDP ... by ... percent.

1. falls; 3
2. falls; 9
3. rises; 9
4. rises; 3

Q — You have saved £0.78 in your Big Apple Bank and are waiting to welcome New Year's Eve 2000 while working a night shift at Applied Cryogenics. Unluckily, you trip over some wire and fall into a cryogenic cell, which locks until New Year's Eve 3000. Assuming the bank kept paying an annual interest of 2.25%, how much money would you have in your account once you arrive at New Year's Eve 3000?

Using constant interest rate 0.0225 and 1000 years as the horizon, we have:

$$\begin{aligned}\text{Terminal Investment Value} &= \text{£}0.78 \times \underbrace{(1 + 0.0225)^{1000}}_{\text{Compound interest}} \\ &= \text{£}0.78 \times 4,605,923,064 \approx 3,592,619,990\end{aligned}$$

Q — Between November 2015 and August 2016, Microsoft paid an annualised dividend of \$1.44 per share. If dividends kept growing at a rate of 12% and the rate of return of Microsoft's equity were assumed to be 14.54%, what would be the price of a single share?

Microsoft's share price today, given the annualised dividend D , equity expected return r and dividend growth g , is:

$$\text{price} = \frac{D}{r - g} = \frac{\$1.44}{0.1454 - 0.12} = \$56.6929 \approx \$57$$