



C7: Quantitative Techniques

Module 4

Copyright

© Commonwealth of Learning 2012

All rights reserved. No part of this course may be reproduced in any form by any means without prior permission in writing from:

Commonwealth of Learning

1055 West Hastings Street

Suite 1200

Vancouver, BC V6E 2E9

CANADA

Email: info@col.org

[Add institute name here]

[Add School/Department name here]

[Add address line 1]

[Add address line 2]

[Add address line 3]

[Add country]

Fax: +[Add country code] [Add area code] [Add telephone #]

Email: [Add e-mail address]

Website: [www.\[Add website address\]](http://www.[Add website address])

Acknowledgements

The Commonwealth of Learning (COL) wishes to thank those below for their contribution to the development of this course:

Course author	Barbara Swart, PhD Professor, Quantitative Management University of South Africa, South Africa
Subject matter experts	Samuel Kwame Amponsah, PhD Kwame Nkrumah University of Science and Technology, Ghana Kim Loy Chee, PhD Wawasan Open University, Malaysia S.A.D. Senanayake Open University of Sri Lanka, Sri Lanka
Educational designers	Symbiont Ltd. Otaki, New Zealand
Course editor	Symbiont Ltd. Otaki, New Zealand

COL would also like to thank the many other people who have contributed to the writing of this course.

Contents

Module 4	1
Modelling in management	1
Introduction	1
Unit 13	2
Optimisation and forecasting	2
Introduction to modelling and decision making	2
The modelling process and model risk	2
Models and assumptions	6
Modelling processes with Markov	8
Activity 4.1	11
Activity 4.2	12
Remember these key points	13
Unit summary	14
Unit 14	15
Simulation	15
Introduction	15
A queuing problem	16
Generation of random numbers	17
Activity 4.3	18
Activity 4.4	19
Remember these key points	19
Unit summary	20
Unit 15	21
Linear and non-linear programming	21
Introduction	21
Introduction to linear programming	21
Activity 4.5	24
Activity 4.6	27
Introduction to non-linear programming	28
Activity 4.7	32
Applications in Excel	32

Activity 4.8	35
Activity 4.9	39
Activity 4.10	41
Remember these key points	42
Unit summary	43
Unit 16	44
<hr/>	
Forecasting: Time series and linear regression	44
Introduction	44
Methods of forecasting	44
Moving averages and exponential smoothing	46
Activity 4.11	50
Activity 4.12	53
Time series models	53
Activity 4.13	60
Activity 4.14	63
Activity 4.15	68
Activity 4.16	69
Remember these key points	70
Unit summary	71
Assessment	73
<hr/>	
References	76
<hr/>	
Further reading	77
<hr/>	

Module 4

Modelling in management

Introduction

Now that you have mastered the mathematical and statistical foundations of modelling using electronic tools, you can move into advanced modelling and decision making.

This module addresses aspects such as:

- the advantages and possible dangers of modelling,
- the use of Markov processes,
- linear and non-linear programming,
- forecasting in decision making and simulation, and
- time series and linear regression.

You will also look at applications in Excel.

In Module 4, you will see examples of the transformation of abstract knowledge, in particular mathematics and probability theory, into practical decision making. This process has been revolutionising economics, business and management and has been accelerated by the huge advances in computers (hardware as well as software), the Internet and telecommunications.

The modelling of processes, from finance and economics to manufacturing and decision making, is now seen as part of the foundations of a successful business.

Algorithms are developed by mathematicians to give companies a millisecond advantage over their competitors. The modern manager and leader must know at least the basic elements of these developments.

Upon completion of this module you will be able to:

- **understand** the philosophy, methods and risks of modelling;
- **apply** modelling to stochastic processes;
- **solve** optimisation problems with linear, integer and non-linear programming techniques;
- **use** simulation and forecasting methods, including regression, to inform decisions; and
- **implement** modelling tools in Excel.



Outcomes



Unit 13

Optimisation and forecasting

Introduction to modelling and decision making

In business and management, you are confronted with problem situations that you have to resolve. Usually you have to simplify the situation and use a model to represent reality. This model should capture the essence of the problem and allow you to analyse its properties and draw conclusions.

In decision making, symbolic models and the power and tools of mathematics are used to represent real-world situations. Models allow parameters and properties to be changed and their effects seen, without actually needing to enforce the changes in the real world. Certain stochastic processes such as price processes are often modelled as Markov chains, which can then be applied to calculate the prices of financial products.

You have already covered most of the modelling tools that are used – graphs and charts, tables of values, variables in algebra, equations, statistics and probability theory.

Upon completion of this unit you will be able to:



Outcomes

- **describe** the steps in the modelling process;
- **explain** the importance and limitations of models;
- **distinguish** between modelling assumptions and reality; and
- **apply** Markov chains to the modelling of stochastic processes.

The modelling process and model risk

A model is an approximation to a real-world situation. It depends on data and information that have been interpreted and sifted by humans for validation and application. For this reason, a model may be affected by political, cultural and individual human biases and subjectivity.

Drastic simplifications are often made in the modelling process, with certain facts or possibilities ignored through necessity or ignorance.

In business management, time and money constraints come into play; solutions may be misunderstood or incorrectly implemented; and implementation may not be monitored properly.

However, progress cannot be made without using models – they are an invaluable tool in decision science.

The modelling process can be summarised by these steps:

1. Observe the situation; identify the problem; collect data; identify objectives.

2. Analyse data; identify variables and relationships; build a model; find solutions.
3. Test the model; see if predictions match observations; change parameters where necessary.
4. Implement the solutions; make decisions; implement and monitor the decisions; update the model.

The process generally has a feedback loop: implementing and updating in Step 4 means new values and variables that have been suggested by the initial solutions are fed back into the model at Step 1 and the process is repeated.

One tool that has become indispensable in modelling and decision making is the computer and its associated software such as spreadsheets and macros. (These were introduced in Module 3.)

Modelling is an essential part of good management and the health of an organisation. It can mean the difference between success and failure. However, you must always be aware of the limitations of modelling – many aspects of an actual situation have to be ignored, many simplifying assumptions are made, and certain variables (such as interest rates, human behaviour and economic conditions) cannot be controlled by the modeller.

Some of the main types of model that have interest for managers are:

- economic models,
- financial and pricing models,
- risk measurement and management models,
- forecasting models, and
- optimisation models.

Model risk and model failure in economic crises

As mentioned, every model has the risk of being implemented incorrectly or inappropriately. This leads to model failure.

Model risk is a form of operational risk, which is the risk of loss for a company from inadequate or failed processes, human error or extreme outside events that the model has not foreseen and allowed for.

Sources and consequences of model risk

Common sources of model risk are wrong assumptions, errors in data, errors in estimating parameters, differences between assumed and actual probability distributions, and incorrect and inappropriate application of models.

We have seen the very serious consequences of model failure in the economic meltdown that spread from America in 2008 to the rest of the world. The many financial crises in the past include:

- the Dutch tulip bubble of 1636;
- the Mississippi bubble caused by John Law in Paris, France in 1729;
- the South Sea bubble caused by John Blunt in London, England in 1720;



- the Latin American debt crisis of 1820;
- the Wall Street crash of 1929, which caused world trade to shrink by a massive 66 per cent;
- the October 1987 crash in the United States;
- the UK Barings Bank collapse caused by Nick Leeson in 1995;
- the Russian debt default of 1998; and
- the Enron debacle of 2001, with Kenneth Lay as chief executive officer.

In many instances, human greed played a big part. However, the current (2008+) crisis is the first that has a substantial component of **model failure**.

Never before have we had such sophisticated models and technology, and never before has the world of finance, business and management been so quantified and mathematised. And yet, these factors have played some part in the collapse – not because mathematical modelling should not be used in the management of our affairs, but because the limits and proper application of such models are not always understood.

This is where you, as a manager with knowledge of a course such as this one, can play an important role.

The following remarks, reproduced from journals, newspapers and the Internet, can be used for discussion.

Note: This section is not intended to devalue the importance of models, but to emphasise the importance of understanding the mathematics and statistics involved in modelling. It also highlights the need to have a healthy respect for both the value and limitations of models in management and decision making.



Discussion

The first financial and economic crisis of the 21st century: To model or not to model?

Analyse and discuss the following statements:

1. Paraphrasing from *Risk Magazine* (Duncan Wood, September 2008):

One school of thought has it that bank risk functions had been taken captive by statisticians, mathematicians and modellers. There has been an unhealthy obsession with risk measurement, while common sense and judgement were allowed to “rust away”.

Models have an important part to play within management, but just handing over risk management to computers and quants [quantitative analysts] is mad. Chastened by the failure of models, banks want to go back to basics, in particular more human involvement.

*Another school of thought has it that models should not be abandoned, but rather improved. Models did all that was expected of them in some cases – they did provide advance warning of the credit bubble burst. In a sense, the system worked in every way but **nobody paid attention**. Instead of getting rid of models, banks need to find a better way of using them.*

2. *The New York Times* (28 October, 2008, David Brooks) comments that we should rethink Step 1 of modelling:

Looking at and perceiving the world is an active process of meaning-making that shapes and biases the rest of the decision-making chain. Human psychology should enter the picture. We should be aware that humans tend to see data that confirm their prejudices more vividly than the data that contradict ideas, and see pattern where there is none. We tend to applaud our supposed skills in circumstances where we’ve benefited from sheer luck. We misperceive the possibility of rare events and don’t include them in our models.

3. Emmanuel Derman and Paul Willmot write on the Wilmott.com blog:

There are no fundamental laws in finance (as there are in physics). Even if there were, there is no way to run repeatable experiments to verify them.

Models are at bottom tools for approximate thinking. Our experience in the financial arena has taught us to be very humble in applying mathematics to markets, and to be wary of ambitious theories which are, in the end, trying to model human behaviour. We do need models and mathematics, but must never forget that models are not the world. Models need to be overlaid with common sense and experience.

We must know what we are assuming when we use a model, and know what has been swept under the carpet. Models can have an enormous effect on society and the economy, beyond our comprehension.



Case study:
Model risk and crises

Choose one of the historical financial crises mentioned in this unit. Research it and write a short report. If possible, mention the positive and/or negative roles that quantitative models played or could have played in the crisis.

Models and assumptions

If you understand the assumptions that underlie a model, you will be able to appreciate its strengths and shortcomings. As examples, we will discuss two well-known economic models and the assumptions they are based upon.

Economic equilibrium and supply and demand

Adam Smith advanced the idea that economic markets should be left alone and would then find stability or equilibrium by a “hidden hand”. It is assumed that humans act rationally, or that any human unreliability and irrationality will be filtered out by the market as a whole.

Equilibrium means balance. At equilibrium, economic forces are balanced; supply of goods equals demand and the reasonable price of a product will be determined. This price is often called the **equilibrium price** and will tend not to change unless demand or supply change. An over-supply should lead to price declines, which return the market to a situation where the quantity demanded equals the quantity supplied. Customers should be happy. Similarly, in an unfettered market, any excess demand (or shortage) would lead to price increases. This gives the producer an incentive to produce more, reducing the demand. As before, the disequilibrium (here, the shortage) disappears.

Supply and demand problems were considered in Module 1, Units 2 and 3. Look at the problems again. They were so neat and easy to solve!

All this sounds ideal and the theory is often used to suggest one should never interfere with the free market: individuals should be free to pursue maximum satisfaction and profits; centrally planned economies often don't get prices right and suffer from shortages of goods; the model and the mathematics can find equilibrium prices and control the economy.

But most economists caution against attaching such meaning to the equilibrium price. For example, food markets may be in equilibrium at the same time that people are starving because the equilibrium price is so high. In some cases, equilibrium does not bring market clearing but instead leads to unemployment. Equilibrium can also correspond with monopolies where the monopolistic firm creates an artificial shortage in order to increase prices and profits.

Look at the supply and demand problems in Module 1 again and try to be more critical. Market equilibrium (supply and demand) theories can provide insight and help with decision making, but they are models with very specific assumptions. Following models slavishly can lead to injustice and suffering.

The efficient market

An important assumption about financial markets is the **Efficient Market Hypothesis** (EMH). Basically, this says that share prices fully reflect all available information. This idea is also explored later in this unit (“Modelling processes with Markov”).

We can distinguish between three types of market efficiency: weak, semi-strong and strong. Weak efficiency means all historical information is reflected in current prices; semi-strong means all publicly available information is reflected in current prices; and strong means all private and public information is reflected in prices. The basis of the assumption is that information and announcements about prices are disseminated very quickly and incorporated almost immediately into the price. Participants in the market react rationally to information. This means they act sensibly and as thinking beings.

The implications of the hypothesis are that share prices are moved by information that people use in a rational manner. Investors cannot make profits above what should be earned according to the known risks associated with the share. (Insider trading, which is illegal, is not included in this discussion.) They cannot predict the future price on the basis of some special analysis of the available information and use this to earn excessive profits. Even if they could, can they act before all others also start buying and push up prices? What about transaction costs?

Another interpretation of the efficiency of markets is that prices reflect investors' expectations about future cash flows – the markets are rational.

All the above seems to suggest that the financial markets operate in an efficient and fair manner.

Many tests have been done to determine the validity of the efficient market hypothesis, with varying results. There is support for the theory, but many anomalies have been detected.

Some research supports the idea that it is possible to predict future prices to an extent. Market participants (particularly speculators) do not always act rationally. The economist John Maynard Keynes described the stock market as a casino. He emphasised the importance of uncertainty, as opposed to measurable risk.

To summarise, the EMH is an attractive model, but you should be aware of its limitations and the possibility of huge market swings and crashes.



Reflection

Equilibrium and fallibility

Another famous economist, George Soros has some interesting thoughts about the ideas of knowledge, equilibrium and investor rationality.

In his book *The Age of Fallibility* (2006), he writes:

Our view (model) of the world will never correspond to the world as it is because ... what we think automatically becomes part of what we have to think about.

The way we look at (model) the world changes the world, and therefore the model may not be valid any more! Understanding a situation and impacting on that situation by making decisions can interfere with each other.

Soros points out a similar effect in the value of assets and the act of valuing or pricing them. The interference between the two processes often leads to a circular self-reinforcing process that ends up in a bubble. He correctly predicted the real estate and mortgage bubble that has now burst.

Soros feels that equilibrium and rational behaviour are therefore unattainable, saying that “the Age of Reason ought to yield to the Age of Fallibility”.

Modelling processes with Markov

Many management decisions have to be made with the knowledge that the process or quantity under consideration changes or evolves randomly over time. Such processes are called **stochastic processes**. At each time step, the quantity is a stochastic variable – for example, at each time step the variable can be normally distributed.

Developments in finance have suggested treating price processes as stochastic processes of a specific type: the **random walk**. The random walk is a type of Markov process. These are general stochastic processes with the very interesting property that past developments do not affect future developments of the process: all the information needed is contained in the present state of the random variable.

Markov processes can also be used to model inventory levels for products. The theory can become quite complicated, and we restrict ourselves here to a simple discussion.

Remember: Be aware, once again, that the models we build are approximations of reality and based on our perceptions.

Notation

In Module 2, you were introduced to the idea of a random variable X . Such a variable can have different values or be in different states, each with a given probability. Now consider the variable as changing over time. We are at time $t = 0$ where we know X_0 but at each later time step $t = 1, 2, \dots, N$ the variable X can take on new random values. So we look at the sequence or process $\{X_0, X_1, X_2, X_3, \dots, X_N\}$, which we write as

$\{X_t, t = 0, 1, 2, \dots, N\}$. Remember that each X_t for $t > 0$ is a random variable. X_t can be the unknown and unpredictable inventory level at future time t , or the price of a share at time t .

Example 1

The price X of a share changes over time, and the values for future days are not known. The price process over, say, 100 days is a stochastic process $\{X_t, t = 1, 2, \dots, 100\}$. Each daily price X_t is a random variable with a probability distribution showing the different possible values of X_t .

Random walks

Example 2

Suppose you are standing at point $X_0 = 0$. A coin is flipped every second and if it is heads, you take one step forward; if it is tails, you take one step backwards. Your position X_t at any time t is a random variable. As time goes on, you will be moving forwards and backwards randomly. At time $t = 4$, you may be at position $0+4$, $0+2$, $0-2$, $0-4$ or back at 0. Can you see this? The process is a simple random walk.

Example 3

We construct a simple **price process model** where the price can go either up or down one unit at each time step. Suppose today is $t = 0$ where price X_0 is known. If X_0 goes up with one unit its value at time 1 will be $(X_0 + 1)$ and if X_0 goes down with one unit its value at time 1 will be $(X_0 - 1)$. Therefore, at time 1 the random variable price X_1 can take on either value $(X_0 + 1)$ or value $(X_0 - 1)$. At time 2, the random variable price X_2 can take any of three values: $(X_0 + 2)$ or X_0 or $(X_0 - 2)$.

(The model can be made more sophisticated by allowing the price to go up or down by a specified percentage at each time step.)

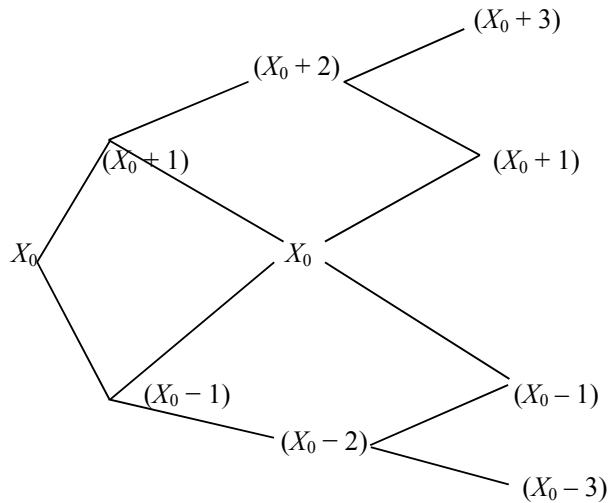
It may be possible to assign probabilities to each state: we assume the probability of an up-move is p and the probability of a down-move is $(1 - p)$. Let us take $p = \frac{1}{2}$.

This process can be displayed in tree form in a similar way to Unit 7 of Module 2. It is called a **binomial tree** (bi- means two) because there are two possible movements, up or down. The tree is recombining in this example (Figure 1) because an up-movement followed by a down-movement has the same effect as a down-movement followed by an up-movement:

$$(X_0 + 1 - 1) = (X_0 - 1 + 1) = X_0$$



The nodes of the tree give the different possible states of price process X_t .



$t=0$	$t=1$	$t=2$	$t=3...$
X_0	X_1	X_2	$X_3...$

Figure 1

This is an example of a random walk. Each X_t has its distribution of values, and in this case you assume each X_t is actually normally distributed. As time unfolds, one of the paths through the tree will be realised and as the time steps get smaller and smaller, the movement approaches a jagged but continuous process called a **Brownian motion**.

Note: Both examples demonstrated these interesting facts:

At each time t the expected value of X_t , namely $E[X_t]$ equals X_0 .
 At each time t the variance of X_t , namely $\text{var}[X_t]$ equals t . The position (state) of the walking person or the price that is realised at time t is only influenced by the position at time $(t - 1)$. What happened at previous times is no longer important. Your prediction of your state at time $t = 5$ will be fully determined by your state at time $t = 4$. What happened earlier is no longer relevant. All previous information is in a sense encapsulated in your position at time $t = 4$. Your next position will be either one step ahead of or one step behind where you are now. In terms of the price process above, to forecast the price at $t = 5$ you only need to know where you are at $t = 4$.

This (in simplified form) is the defining property of a Markov chain. A random walk is a Markov chain or process.

Markov chains

Conditional probabilities were explained in Module 2, Unit 7. You can define Markov chains in terms of conditional probabilities.

For a Markov chain, the probability P of being in state j at time n , conditional on the fact that you know the states at all previous times $0, 1, 2, 3, \dots, (n - 1)$ is just the probability of being in state j at time n ,

conditional on the fact that you know the state at the immediate previous time ($n - 1$).

$$P(X_n = j \mid \text{knowledge of states assumed by } X_{n-1} \dots X_3, X_2, X_1 \text{ and } X_0) = P(X_n = j \mid \text{knowledge of state assumed by } X_{n-1}).$$

Under the Markov model, for a process all you need to know to move one step into the future is the present state. Knowledge of tomorrow comes from knowledge today. The past is not important. If a share price follows a Markov process, our forecast of a future price is not affected by the price a week or month ago, but only by today's price.

Share prices are often modelled in terms of Markov chains. The decision to model price processes with random walks or (geometric) Brownian motion has a long history. The Markov property of share prices is consistent with the assumption of weak market efficiency. This would mean that technical analysis of the history of share prices is of no use. However, there are people who don't agree with this model.

Why do you need to know about this?

Communication processes can be modelled by Markov processes, as can inventories.

Many companies these days invest in sophisticated financial markets. The pricing of exotic financial products such as options (some of which are very complicated) is based on the fact that share prices are modelled with binomial trees or Brownian motion. In fact, the famous Black-Scholes formula relies on this.

Many people blame the options market for part of the financial collapse in 2008. Therefore, it is important that you are aware of the assumptions underlying the pricing model.

Activity 4.1



Activity

Consider random walk

What will you do?

- Consider the random walk of Example 2: suppose you are standing at point $X_0 = 0$. A coin is flipped every second and if it is heads you take one step forward, and if it is tails you take one step backwards. Your position at any time t is a random variable X_t .
 - Construct a binomial tree with four time steps for X_t .
 - Show that the expected value $E[X_t] = 0$ for each t .
 - Prove that $\text{var}[X_t] = t$.
 - Show that at each node on the tree, the value of X_t equals the expected value of the two values X_{t+1} following from that node. Use this fact to reason that if the values of X at the end nodes are given, then you can work backwards to find the value X_0 at time $t = 0$.
- Explain the link between the efficient market hypothesis and the Markov model.



Activity 4.2



Activity

Understand the terminology

What will you do?

Use this terminology table to record any terms or words you're uncertain about.

This activity is an opportunity to consolidate your understanding of new terminology and concepts you encountered in Unit 13. Fill in the terms you have learned and then write your own descriptions of them.



Terminology

Term	Description
	:
	:
	:
	:
	:
	:
	:
	:
	:

Remember these key points

- In business and management, you are confronted with situations where decisions need to be made in an informed and responsible way. Decision theory offers various modelling tools.
- First, you have to simplify the situation and use a model to represent reality. This model should capture the essence of the problem and allow us to analyse properties and draw conclusions.
- Remember that every model has the risk of being implemented incorrectly or inappropriately. This leads to model failure.
- Models are based on simplifying assumptions. It is important to be aware of these, so that you can critically evaluate models and theories. Modelling plays an important role in understanding business and management situations.
- Markov chains are a useful way of modelling various processes. The random walk, in particular, allows us to construct pricing formulas for complicated products.
- Understanding models and formulas is the key to their efficient use.



Unit summary

You have successfully completed this unit if you can:



Summary

- **address** the modelling of business and management problems;
- **take** into account model risk and model assumptions;
- **explain** how Markov processes are used to model processes;
- **describe** the steps in the modelling process;
- **discuss** the importance and limitations of models; and
- **distinguish** between modelling assumptions and reality.

Unit 14

Simulation

Introduction

Simulation imitates the behaviour of stochastic situations. There are situations that cannot be studied by repeated observation or by repeatedly running the process in real life – it would just not be possible. Simulation provides a very useful way of getting a picture of the different possible outcomes from such situations.

Simulation models are applied using computers and generate huge numbers of possible paths that a situation can take. In this way, we try to replicate what could possibly happen in an uncertain real-world scenario.

The paths can be generated by random number programs that supply random strings of digits from zero to nine. These types of simulation are called Monte Carlo simulation (referring to the casinos of Monte Carlo). Simulations can also be run under circumstances of risk where the probabilities are known.

The computer simulation runs through the situation repeatedly to produce a sample of different random outcomes. This information could be used to analyse the situation, make rough forecasts and help with the decision-making process. Simulation is dynamic and tracks the developments in a process through time.

Typical situations where simulation can be used are for production lines, queuing and share price movements.

Simulation allows the decision maker to observe possible outcomes and collect data without observing the *actual* process. This information can then be fed into Step 1 of the modelling process.

Upon completion of this unit you will be able to:

- **describe** the steps in running a simulation;
- **generate** random numbers; and
- **understand** how to analyse uncertain problems using simulation.



Outcomes

A queuing problem



Case study: Simulation study for a bottling plant

Bottles are filled with spring water on a production line, examined at an inspection point to see if they are full and, if so, moved along to where a label is fixed to the bottle. If the bottle is not filled properly, it is rejected and sent to another queue. Since we are uncertain whether a bottle will be accepted or rejected, we assign 50 per cent probabilities to both cases. (This is not very realistic, but helps to keep the example simple.)

It takes 10 seconds to fill a bottle. Examination takes almost no time, with rejected bottles removed and the rest moved on to receive the label. It takes two seconds to fix a label.

On the way to the labelling machine the bottles move along a queue where there may be hold-ups, depending on how many bottles were accepted or rejected at the inspection point.

The process is analysed to see whether the production line works efficiently, whether there are bottlenecks and whether decisions for improvements need to be made. Is it possible to forecast how long it will take, on average, for a bottle to pass through the system? We may need this information to enable us to decide whether we should invest in the production system or not.

One way would be to rent the system, set it up with a large number of bottles, observe the production line for a long time and collect the data. This could be very time consuming and expensive.

Another way is to simulate the process, and then the real-life system need not be run.

The simulation process

The uncertainty arises in this example because it is not known at the start which bottles will be rejected and which will continue down the line for labelling. Say 20 bottles in a row all pass the inspection and then there may be hold-ups at labelling. If, on the other hand, 20 bottles in a row fail the inspection, time will be wasted because no labelling will occur for some time.

Simulate the process by assigning random numbers to a sequence of imaginary bottles. These numbers can be generated by a computer in the form of a random string of zeroes and ones. The string can be as long as you like. Then you can decide that zero represents acceptance and one represents rejection. Each zero or one has a 50 per cent chance of being generated. This makes sense for this example because you assumed a 50 per cent probability for acceptance or rejection.

A computer program can then compute the time taken through the process in each case. After that, you can calculate properties such as maximum time spent on the production line, average time per bottle on the production line, amount of time the labelling machine was idle, and so on. All this can be done by sophisticated computer programs in a short time, with little cost and for a very large number of imaginary bottles.

This example is an instance of so-called queuing problems. Simulation is a good tool for analysing these kinds of problem and there are software packages available to run simulations.

One of the steps in the simulation process involves generating random numbers to represent the process you are studying.

Generation of random numbers

A random number is one whose value cannot be predicted with any certainty. Random numbers are generated by random number generators or RNGs. Look under <Tools>, <Data Analysis> in the software package Excel.

One example is the function RAND() in Excel. It produces decimal numbers between zero and one that are from the uniform random distribution. (Uniform random variables were discussed in Module 2, Unit 2 and the Excel function RAND was introduced in Module 3.)

A variation of RAND is RANDBETWEEN(bottom,top). Enter numbers in place of bottom and top to indicate the range where you want the numbers to lie. For example, RANDBETWEEN(0,9) generates random numbers from zero to nine. RANDBETWEEN(0,1) generates random numbers zero or one. You can use this tool to generate numbers for our simulation of bottles in a production line.

In situations where probabilities are available, you can use random number generators as follows. Suppose it is estimated that the probability of a bottle being rejected is 0.25 (or 25%). This means the probability of a bottle being accepted is 0.75 or 75%. To simulate this, you can use RANDBETWEEN(0,99). All numbers from zero to 24 that appear in the generated sequence are then seen as representing rejected bottles. Why? Well, there are 100 numbers from zero to 99. This represents 100%. A percentage of 25% is represented by 25 numbers, which you may choose as zero to 24. All numbers from 25 to 99 are then seen as representing accepted bottles. Therefore, a block of 50 numbers generated from RANDBETWEEN(0,99) can simulate the rejection or acceptance of 50 bottles.

Random numbers can also be generated from the normal distribution.

Note: A single simulation or sample is not enough. A common mistake in using simulation is to take too small a sample and base decisions on the results. You should run the simulation as many times as is feasible and then take the average of the outcomes.

Activity 4.3



Activity

Solve the bottle problem

What will you do?

1. Generate a block of 100 random numbers using function `RANDBETWEEN` in Excel to distinguish between accepted and rejected bottles. Discuss your result.
2. Determine how many bottles were rejected in this simulation.
3. Repeat the process 15 times (a sample of 15 runs) and determine the average number of bottles that will be rejected. Also determine the standard deviation.

Here's our feedback

1. Open a worksheet. Click on cell A1 and enter `=RANDBETWEEN(0,1)`. Press `<Enter>`. Click on A1 again and the Copy icon. Drag to highlight cells A1 to J10. Click Paste. A block of random zeroes and ones will appear.
2. Digit 1 represents rejection. If you sum over all cells, you will know the number of bottles that were rejected because the ones that are accepted contribute zero. Click on A12 and enter `=SUM(A1:J10)`. The answer will be the number of rejected bottles.
3. Click on any cell in the block of numbers and then click in the formula bar fx and press `<Enter>`. A new set of numbers will be generated with a new answer for SUM. Repeat. Determine the average and standard deviation over 15 runs in a separate worksheet. (You can also try to set up a VBA program.)

This activity can be taken further with computer programs that analyse the results and calculate further properties of the process, such as the average time taken for a bottle to go through the process. This falls outside the scope of the course.

Activity 4.4



Activity

Understand the terminology



Terminology

What will you do?

Use this terminology table to record any terms or words you're uncertain about.

This activity is an opportunity to consolidate your understanding of new terminology and concepts you encountered in Unit 14. Fill in the terms you have learned and then write your own descriptions of them.

Term	Description
	:
	:
	:
	:
	:
	:
	:
	:
	:

Remember these key points

- Simulation can be used to describe the behaviour of a system under uncertain conditions. Different scenarios can be played out on a computer in this way, without actually observing the real-life process.
- Each simulation is a sample that is obtained by randomly generating values for the unknown quantity (number of people in a queue, rejected bottles, and so on). Analyses of the observations that are generated then help with modelling and decision making.



Unit summary

You have successfully completed this unit if you can:



Summary

- **use** simulation in the modelling of situations;
- **generate** random numbers; and
- **analyse** uncertain problems using simulation.

Unit 15

Linear and non-linear programming

Introduction

This section looks at ways of finding the optimal values of functions that are subject to constraints. An optimal point can be a point of maximum value (such as for a profit function) or a point of minimum value (such as for an expense function). There will be constraints in these situations – expenses cannot be minimised beyond a certain level and there are limits to the profits that can be obtained.

Upon completion of this unit you will be able to:



Outcomes

- **calculate** the optimal value of quantities under constraints using linear programming methods;
- **understand** the basics of integer programming;
- **determine** maxima or minima with non-linear programming; and
- **apply** the methods using Excel.

Introduction to linear programming

Linear programming (LP) is considered one of the most important tools developed in the 20th century. It can be described as a method for allocating limited resources in an optimal way. It provides a simple and efficient way of analysing situations where the function to be optimised is a linear function. This function will be dependent on a number of variables.

The method is introduced using a simple example of a function of two variables. This section is quite mathematical and if you want to skip the mathematics, you can proceed directly to the next section, “Feasible region”.

Linear functions of two variables

Linear functions of two independent variables x and y are denoted by: $f(x, y) = ax + by + c$, or by $z = ax + by + c$ where a , b and c will be given constants. The variable z ($=f(x, y)$) is now the dependent variable. As you change the values of variables x and y , the value of variable z will change accordingly. If $x = y = 0$, then $z = c$.

Functions of two variables can be displayed in three dimensions in the (x, y, z) -space. Three-dimensional space can be represented by orthogonal co-ordinate axes x , y and z where z is the height above or below the (x, y) -plane.

The graphic representation of $f(x, y) = ax + by + c$ is a flat, angled plane or surface in space, intersecting the z -axis at $z = c$ (because when $x = y =$

0, you see that $z = c$).

So for each given pair (x, y) in the horizontal (x, y) -plane, $f(x, y)$ or z is a point on the inclined plane directly above or below (x, y) . We restrict ourselves to the case where $z \geq 0$, and a and b are non-zero.

Here's an example:

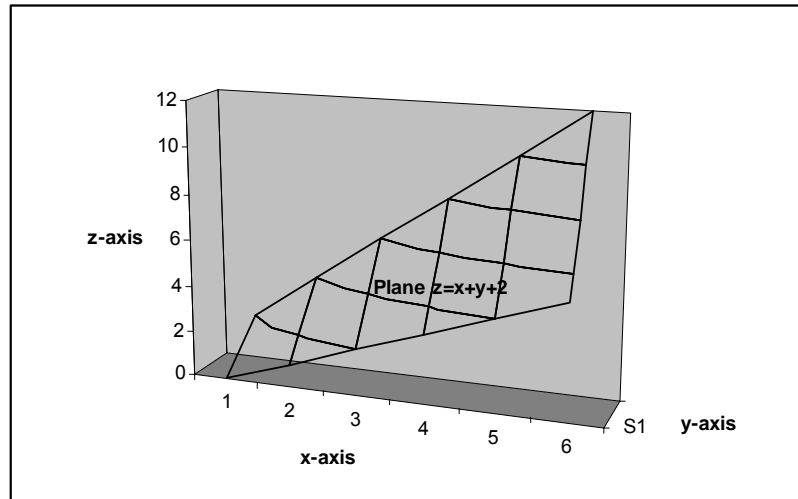


Figure 2

We sketched part of the plane (surface) $z = x + y + 2$ above. The origin of the axes is point $(0, 0, 0)$ and lies in the bottom deep corner (look “through” the page). The plane itself cuts the z -axis where $x = y = 0$ (that is, at $z = 2$). It slopes upwards (in front of your face) much like a mountain slope.

Now $f(1, 0) = 1 + 0 + 2 = 3$. That is, $z = 3$ when $x = 1$ and $y = 0$. The point $(1, 0, 3)$ lies on the inclined surface at height 3 directly above point $(1, 0)$ in the flat (x, y) -plane. The point $(0, 1, 3)$ lies on the surface, also at height 3 directly above point $(0, 1)$ in the (x, y) -plane. In plain words, if you beam up points in the (x, y) -plane, you land on the inclined plane or surface $z = x + y + 2$ above. If you beam up point $(1, 0)$ in the (x, y) -plane you land on the point $(1, 0, 3)$ in the surface above.

Note: Both points, $(0, 1, 3)$ and $(1, 0, 3)$, lie at height $z = 3$. There are infinitely many points in the (x, y) -plane that can be beamed up to the surface to lie at height $z = 3$. Conversely, if we beam down onto the (x, y) -plane from all the points on the surface at height 3, we will get a line of points below.

In general:

For each given height $z = H$, there is a straight line of points in the (x, y) -plane satisfying the relation $H = ax + by + c$. You can see this by solving for y in the equation and writing this as:

$$y = \frac{-a}{b}x + \frac{H - c}{b}$$

There are only two variables now because z has been given a specific fixed value H and a , b and c are also constants. This expression for y is clearly a straight line in the (x, y) -plane. Refer back to Module 1. For every value H , you get a line with a different y -intercept.

These lines can be thought of as the contour lines of the inclined surface above, corresponding to different heights H . The contours are all parallel with the same slope $m = \frac{-a}{b}$ and different y -intercepts (which are determined by varying H at points $\frac{H - c}{b}$).

A **contour line** is a line in the (x, y) -plane (such as on a map), so that all points on this line represent points on the z -plane (mountain slope or surface) above that are at height H . These contour lines are also called **level lines**. Take a few moments to let this sink in.

Example 4

Draw the contour lines corresponding to $z = x + y + 2$ for heights $z = 2$, $z = 4$ and $z = 6$, respectively. The equations are: $y = -x$, $y = -x + 2$ and $y = -x + 4$, respectively. Make sure you can derive these equations.

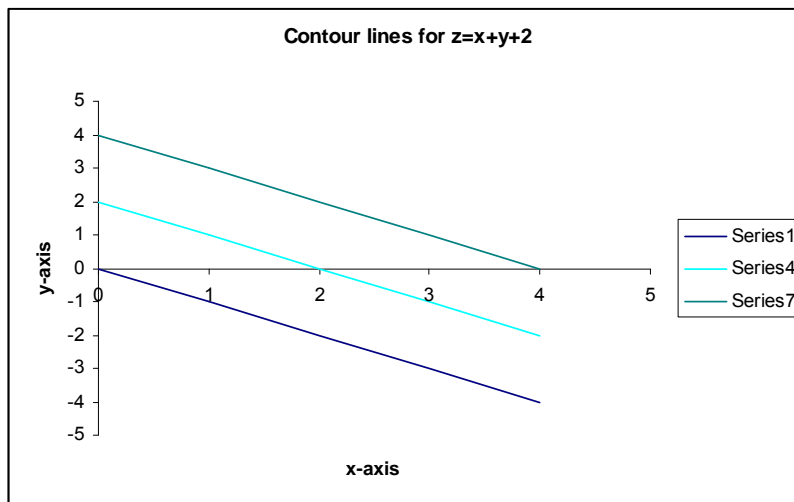


Figure 3

The series 1 line (bottom) is the contour line corresponding to $z = 2$, the series 4 line (middle) corresponds to $z = 4$ and the series 7 line to $z = 6$.

Feasible region

In linear programming problems, we will consider a function $F(x, y)$ that has to be optimised with respect to x and y . The variables x and y will be subject to restrictions or constraints defined by linear inequalities such as $x \geq 0$, $y \geq 0$, $ax + by \geq 0$, and so on. Each restriction defines a region in the (x, y) -plane (see Module 1, Unit 2, “Linear inequalities”) and the intersection of all such regions forms the feasible region for the problem. The optimal values for x and y must lie in the feasible region.



Example 5

The feasible region will have a number of sharp corners. Constraints $x \geq 0$, $y \geq 0$, $x \leq 4$, and $y \leq 2$ define a feasible region that is a rectangle in the (x, y) -plane. The rectangle has corners at $(0, 0)$, $(0, 2)$, $(4, 0)$; $(4, 2)$.

Object function

The function $F(x, y)$ that has to be optimised is called the **object function**. If it is linear, it will be a sloping plane in three dimensions. Optimising F means finding the point (x, y) where $F(x, y)$ reaches its optimum value (maximum or minimum height). Therefore we will consider the level curves or contour lines of F corresponding to all possible values or heights of F . The point (x, y) we are looking for must also lie in the feasible region determined by the constraints.

So we look for the contour line for the optimal height of F that also touches the feasible region. It is a mathematically proven fact that the optimal line will touch the feasible region at a corner point (extreme point). This corner point (x^*, y^*) then gives the solution: (x^*, y^*) satisfies the constraints and $F(x^*, y^*)$ is the optimal value of F .

Don't worry if you're feeling a little lost after this long explanation, as the linear programming method will be illustrated with an example. First, note the general graphic method.



Note it!

Basic linear programming

- Step 1: Identify the two unknown quantities that you want to find and name them x and y .
- Step 2: Identify the (linear) quantity or function of x and y that should be optimised. This function is called the object function $F(x, y)$.
- Step 3: Identify all constraints and regions in the (x, y) -plane. These will define a feasible region.
- Step 4: Graph the feasible region. Graph level curves (contour lines) of the object function F that run across the feasible region.
- Step 5: Determine the point (x^*, y^*) on a contour line that touches the feasible region so that F is optimised there. This yields the optimal quantities for x and y and the corresponding optimum value of F .

Activity 4.5



Activity

Find the maximum profit

What will you do?

Find the maximum of profit function $F(x, y) = 2x + 4y + 5$. The units are in thousands of dollars. Variables x and y represent quantities of two items A and B that are sold. The constraints are $x \geq 0$, $y \geq 0$ and $x + y \leq 1$. (This means the quantities of items sold must be non-negative and you cannot sell more than 1,000 items altogether.)

Here's our feedback

Follow Steps 1–5 in basic linear programming.

Steps 1–2 are complete. To complete Step 3, solve the inequality $x + y \leq 1$. Write it in the form $y \leq 1 - x$; draw line $y = 1 - x$ and note that the inequality gives the region below the line because point $(0, 0)$ satisfies the inequality.

In Step 4, the feasible region is the intersection of the regions $x \geq 0$, $y \geq 0$ and the region below $y = 1 - x$. It is the triangular region bounded by the positive x -axis, positive y -axis and line section $y = -x + 1$. It has corners at $(0, 1)$, $(0, 0)$ and $(1, 0)$.

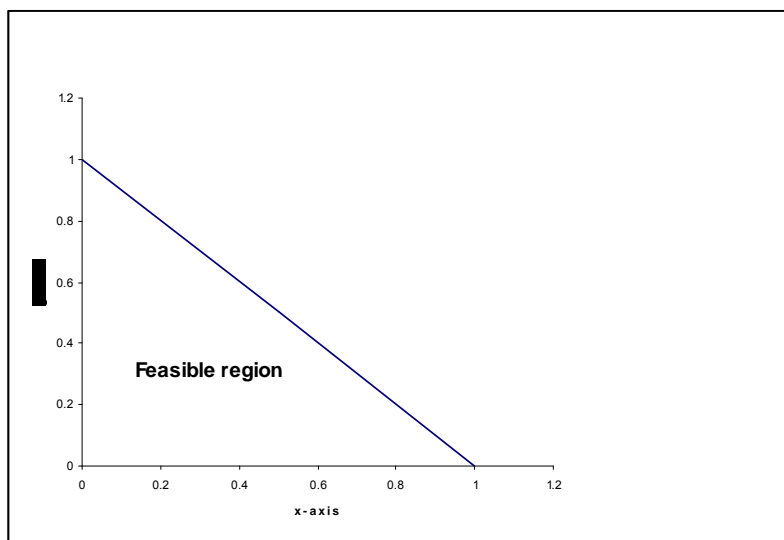


Figure 4

Function F is a plane sloping upwards above the feasible region. The contour lines (level curves) for F are given by solving for y from the equation $2x + 4y + 5 = H$, where H can be any given value of object

$$\frac{H - 5}{4}$$

function F . You obtain $y = -0.5x + \frac{H - 5}{4}$.

These lines all have slope -0.5 and move upwards as H increases.

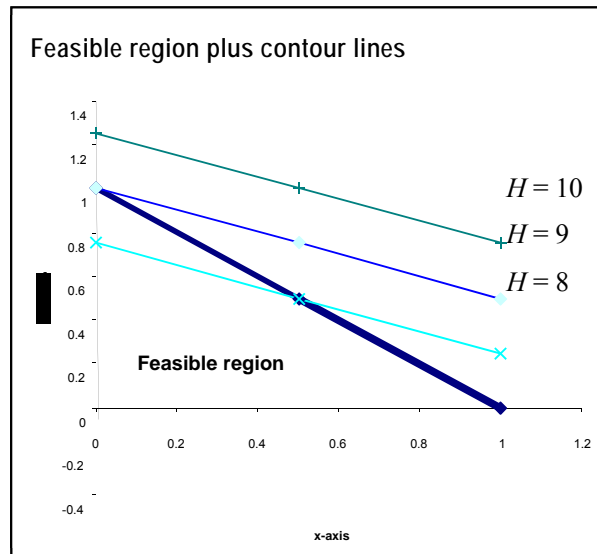


Figure 5

The top three parallel lines correspond to $H = 10$, 9 and 8, respectively. Clearly, the maximum value of F for values of (x, y) **in the feasible region** is obtained where the middle line touches the feasible region at the corner point $(0, 1)$ on the y -axis.

For Step 5, the optimum values are for $x = 0$, $y = 1$ and the maximum of object function F under the restrictions is then

$$F(0, 1) = 2(0) + 4(1) + 5 = 9.$$

Conclusion: The maximum profit of \$9,000 is obtained by selling 1,000 of item B and 0 of item A.

Note: This was a simple example, with a small number of constraints, leading to a simple feasible region. The level lines were easy to draw. It may become difficult to draw level lines to discover graphically where the object function reaches a maximum or minimum. This leads us to a simpler method.



Note it!

Corner-point method

It is a mathematically proven fact that the optimal level line will touch the feasible region at a **corner point**. Therefore, the solution to a linear programming problem is always at a corner point.

Follow Steps 1–4, but do not draw level curves for object function F . Instead, find all corner points of the feasible region.

Step 5. Evaluate F at all the corner points. The point (x^*, y^*) that optimises F yields the optimal quantities for x and y .

This process may also become difficult because more constraints could lead to a region with many corners that are difficult to identify. In this case, you will need a computer and special software.

Example 6

In Activity 4.5, the corner points are $(0, 0)$, $(0, 1)$ and $(1, 0)$. We know the solution must be one of these points. The values of F at the corner points are $F(0, 0) = 5$, $F(0, 1) = 9$ and $F(1, 0) = 7$. We conclude that F reaches its maximum value of 9 at point $(0, 1)$.

Activity 4.6



Activity
Find the maximum

What will you do?

1. Find the maximum of the function F in Activity 4.5 using the corner-point method.
2. Find the minimum of the function F in Activity 4.5 using the graphic and corner-point methods.
3. Find the maximum of function $G(x, y) = 4x + 2y + 8$ subject to the constraints $0 \leq x \leq 3$, $y \geq 0$ and $y - x \leq 1$.
4. Find the minimum of function $F(x, y) = x - 2y + 6$ subject to the constraints $0 \leq x \leq 3$ and $0 \leq y \leq 3$.

The formal name for the preferred general algorithm for solving LP problems is the **simplex method**, developed in 1947. It is very efficient, and usually done on a computer with software packages.

Simplex method

The simplex method works as we just described it. It relies on these facts:

1. Solutions to LP problems are corner points.
2. If a corner-point solution has no adjacent corner point that is better in terms of optimising F , then that corner point is the solution.

First, identify the corner points of the feasible region. Start at one, say (a_1, b_1) , and evaluate F there. Now evaluate F at the two points adjacent to point (a_1, b_1) . If the values are not higher at adjacent (neighbouring) points, you have found the solution at (a_1, b_1) .

If the values are higher at adjacent points, then move to the point with the highest value for F , say (a_2, b_2) . Calculate the value of F at its neighbour. If it is not higher than the value at (a_2, b_2) , then F has its maximum at (a_2, b_2) . If not, move to the neighbour. And so on.

The simplex method is slightly more complicated than discussed here, but essentially this is how it works.

Small and medium-sized problems can be solved in Excel using the Solver tool. (This is discussed in Unit 15.) Large problems need a modelling language or mathematical programming language such as AMPL, MPL and LINGO.

Integer linear programming

Integer linear programming (ILP) is used when there is an additional constraint – that the solutions for x and y *must be integers* (*whole*



numbers). This is typical in problems where x and y refer to the numbers of items that have to be manufactured or sold, as you cannot make or sell fractions of items. The integer feasible region now consists of only those points (x, y) in the ordinary LP feasible region where both x and y are integers. They need not be corner points at all.

Example 7

Maximise $F(x, y) = 2x + 3y$ subject to the constraints:

$$x + 3y \leq 8.25$$

$$2.5x + y \leq 8.75$$

$$x, y \geq 0$$

x, y must be integers

As a first step, ignore the integrality constraint and solve the “relaxed” LP problem:

Maximise $F(x, y) = 2x + 3y$ subject to: $x + 3y \leq 8.25$

$$2.5x + y \leq 8.75$$

$$x, y \geq 0$$

If we are lucky, the optimal values for x and y may turn out to be integers and we will have solved the ILP. Indeed, in the last example this is exactly what happens: $x = 122$ and $y = 78$.

You will not always be so lucky and the solutions to the relaxed LP problem will not be integers. Could you round off the numbers to the nearest integers? No, because rounding off could move you outside the integer feasible region.

There is an algorithm (known as the branch-and-bound algorithm) that allows you to solve any ILP by solving a series of related LP problems, but it takes us beyond the scope of this module. However, we will show how Excel can solve ILPs very quickly.

Let's summarise

Linear programming allows us to solve problems involving minimising or maximising an object function $F(x, y)$, subject to constraints on x and y . The object function can be cost, profit, and so on and can be represented by a plane in 3-D with a family of parallel contour lines down in the (x, y) -plane.

The constraints define the feasible region in the (x, y) -plane. The aim is to find the corner point of the feasible region where the contour line representing the optimal object function touches it.

Introduction to non-linear programming

Up to now, the objective function F and the constraints have all been linear expressions. Many scenarios require us to use non-linear functions and/or constraints, requiring the use of non-linear programming (NLP).

In this case, the graph of object function F is not a plane but may be a surface more like a mountain, with the level lines curved and the feasible region not necessarily having corners. Even if the region has corners,

optimal solutions need not occur there. This situation is far more complicated than for linear programming.

Example 8

Minimise $F(x, y) = xy$

subject to $x, y \geq 0$ and $x^2 + y^2 \leq 4$.

Those of you with more advanced knowledge of mathematics will see that in this case, the feasible region is a quarter circular disc of radius 2, and the level curves for F are hyperbolas of the form $y = \frac{H}{x}$, H constant.

Figure 6 shows two level curves for $H = 1$ and $H = 2$, respectively, and the boundary of the circular feasible region.

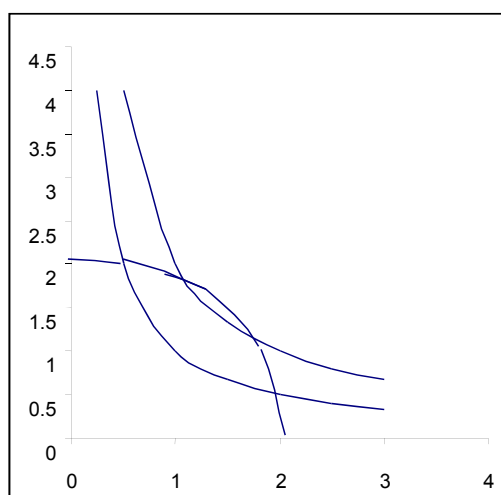


Figure 6

Excel can be used to solve this kind of problem. (See “Applications in Excel” later in this unit.)

Local and global optimal solutions

One important difference between linear and non-linear functions is that linear functions have a unique maximum or minimum in the feasible region, whereas non-linear functions can have many local optima. One of these will be the overall or global optimum, and this is the solution we must find.

Here is a simple example, illustrated in two dimensions.

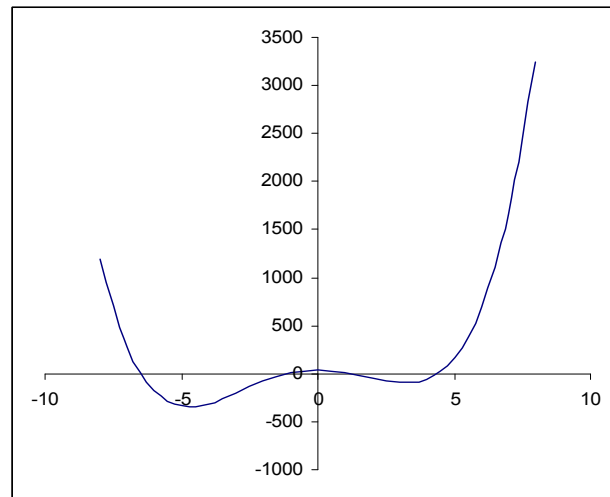


Figure 7

In Figure 7, there are two local minima at $x = -5$ and at $x = 3$. The global minimum is at $x = -5$. On the interval $[-8, 8]$ there are three local maxima at $x = -8$, $x = 0$ and $x = 8$. The global maximum is at $x = 8$.

Markowitz portfolio selection (optional)

A famous non-linear optimisation problem is that of portfolio selection for investment in risky shares.

A collection of risky shares is called a **portfolio**. The idea is that the investor wants to maximise the expected rate of return on the portfolio, subject to a specified risk he/she is willing to accept, or to minimise risk subject to a specified rate of return he/she requires.

Remember that risk is measured as the standard deviation of rates of return and is called the **volatility** of the shares. The problem is that an increase in return generally leads to an increase in volatility and risk.

Risk can be decreased to some extent by diversifying the portfolio to combine shares that are negatively correlated. (See Module 2, Unit 6, “Skewness and the spread of data” and Activity 2.5.) This approach to investment was pioneered by H. Markowitz and W. Sharpe.

Example 9

Your portfolio P contains two shares, X and Y. Assume that shares X and Y have expected rates of return and volatilities (standard deviations of rates of return):

$$E(R_X) = 5\% \quad E(R_Y) = 10\% \quad \sigma(R_X) = 25\% \quad \sigma(R_Y) = 45\%$$

The co-variance of rates of return is denoted by $\sigma_{X,Y}$. Assume that $\sigma_{X,Y} = -0.024$, which means the shares are negatively correlated.

Portfolio P consists of proportions x_X and x_Y of shares X and Y, respectively. The proportions or fractions x_X and x_Y are unknown variables and the decision problem is to determine what their values should be. They should add up to 1. If you want a quarter of the portfolio to consist of shares X, then the remaining three-quarters will be shares Y:

$$x_X = 0.25 \text{ and } x_Y = 0.75 \text{ so that } x_X + x_Y = 1.$$

(Technically, it is possible for one of these to be negative. Say x_X is negative, $x_X = -2$. This means you are **short-selling** two of share X: you borrow the shares from a broker for a certain period of time, sell them and invest the money in buying share Y. At the end of the period, you must buy and return the two shares of X.)

The expected rate of return on portfolio P is the proportional sum of returns on X and Y:

$E(R_P) = x_X E(R_X) + x_Y E(R_Y) = x_X (0.05) + x_Y (0.10)$ and the volatility or risk of the portfolio is:

$$\begin{aligned}\sigma(R_P) &= \sqrt{x_X^2 \sigma^2_X + x_Y^2 \sigma^2_Y + 2x_X x_Y \sigma_{X,Y}} \\ &= \sqrt{x_X^2 (0.25)^2 + x_Y^2 (0.45)^2 + 2x_X x_Y (-0.024)}\end{aligned}$$

Suppose you will accept a level of risk of not more than 30 per cent volatility on your portfolio P. There are different values for x_X and x_Y that will achieve this, leading to different rates of return on P. How should you choose?

The idea is that you will try to maximise your expected rate of return for that level of risk. This is a non-linear programming problem:

Maximise $F(x_X, x_Y) = E(R_P) = x_X (0.05) + x_Y (0.10)$

subject to constraints:

$$\begin{aligned}\sqrt{x_X^2 (0.25)^2 + x_Y^2 (0.45)^2 + 2x_X x_Y (-0.024)} &\leq 0.30 \text{ and} \\ x_X + x_Y &= 1\end{aligned}$$

If we do not allow short-selling, we will add the constraints:

$$x_X \geq 0 \text{ and } x_Y \geq 0$$

The object function F (or $E(R_P)$) is linear, but the constraint on risk is non-linear. This is a non-linear programming problem.

You can, of course, have more than two shares in a portfolio, say shares X_1, X_2, \dots, X_N with proportions x_1, x_2, \dots, x_N where N is any number.

You can also decide to minimise the risk, subject to a constraint that the expected return must be at least 10 per cent. The problem will then look like this:

$$\text{Minimise } F(x_1, x_2, \dots, x_N) = \sigma(R_P) = \sqrt{\sum_{i \neq j=1}^N x_i^2 \sigma^2_{X_i} + 2x_i x_j \sigma_{X_i, X_j}}$$

subject to constraints:

$$\begin{aligned}\sum_{i=1}^N x_i E(R_{X_i}) &\geq 0.10 \\ \sum_{i=1}^N x_i &= 1\end{aligned}$$

In this case, the object function F (or $\sigma(R_P)$) is non-linear, but the constraint on return is linear. This is also a non-linear programming



problem.

We will show how to solve portfolio selection problems next.

Activity 4.7



Activity
Maximise and
minimise

What will you do?

Classify the problems as LP, ILP or NLP. Draw all feasible regions and a level line for the object function where possible.

1. Maximise $F(x, y) = x + y$ subject to the constraints $0 \leq x \leq 3$, $y \geq 0$ and $y^2 + x^2 \leq 1$.
2. Minimise $G(x, y) = 2x - 3y$ subject to the constraints $x, y \geq 0$ and $y + x \leq 1$.
3. Maximise $F(x, y) = x y + 1$ subject to the constraints $0 \leq x \leq 3$ and $0 \leq y \leq 1$.
4. Maximise $K(x, y) = x + y$ subject to the constraints $0 \leq x \leq 3$, $y \geq 0$, $y + x \leq 3$ and x and y both integers.

Applications in Excel

The linear, integer and non-linear programming problems discussed earlier in this unit may look daunting, but if you can read mathematical symbols and type you can solve these problems. Alternatively, as a manager you will recognise the type of problem you are dealing with and will be able to instruct someone else on what to do to solve it using Excel. At the very least, you will not be intimidated when others throw around jargon such as LP and NLP algorithms.

The Solver tool for linear programming

Solver is a tool in Excel that solves LP problems by implementing the simplex method. To ensure that the Solver Add-in has been activated, go to Excel, click <Tools>, <Add-ins> and check that Solver Add-In has been ticked.

Example 10

In Activity 4.5, you considered this problem:

Find the maximum of function $F(x, y) = 2x + 4y + 5$. The constraints are $x \geq 0$, $y \geq 0$ and $x + y \leq 1$.

You applied the techniques of basic linear programming (drawing the feasible region and contour lines, or using the corner-point method) to solve the problem. Although visually appealing, the process was rather time consuming and would not be easy for more complicated object functions or constraints.

The Solver tool can be used for LP problems such as this one.

Step 1: Create a workbook and enter:

Cell A1: "Object function F" Cell B1: $= (2*B2 + 4*B3 + 5)$

Cell A2: “x-variable” Cell B2: 0
 Cell A3: “y-variable” Cell B3: 0
 Cell A4: “x+y” Cell B4: =(B2+B3)

The values of 0 in B2 and B3 are merely “starter” values for x and y and you could enter any reasonable values. (Discuss what would be “reasonable values”.) You need these cells because they contain the variables for which you are optimising F . They also appear in the constraints. You need cell B4 because “ $x+y$ ” appears in a constraint.

Click on B1. Now click <Tools> and then Solver.

A box with Solver Parameters will appear. \$B1\$ should be in “Set Target Cell”. Remember that the \$ symbol on either side of a cell name means that the cell is an absolute reference and will not change.

In the next line, click “Max” because you want to maximise F . In “By changing cells”, enter B2,B3.

Click in the “Subject to Constraints” box and then on <Add>. The “Add Constraint” box pops up. In “Cell Reference”, enter B2, in the next box choose inequality \geq and in the final box enter 0. Click <OK>. You have entered constraint $x \geq 0$. Click <Add>.

The “Add Constraint” box pops up. In “Cell Reference”, enter B3, in the next box choose inequality \geq and in the final box enter 0. Click <OK>. You have entered constraint $y \geq 0$. Click <Add>.

The “Add Constraint” box pops up. In “Cell Reference”, enter B4, in the next box choose inequality \leq and in the final box enter 1. Click <OK>. You have entered constraint $x+y \leq 1$.

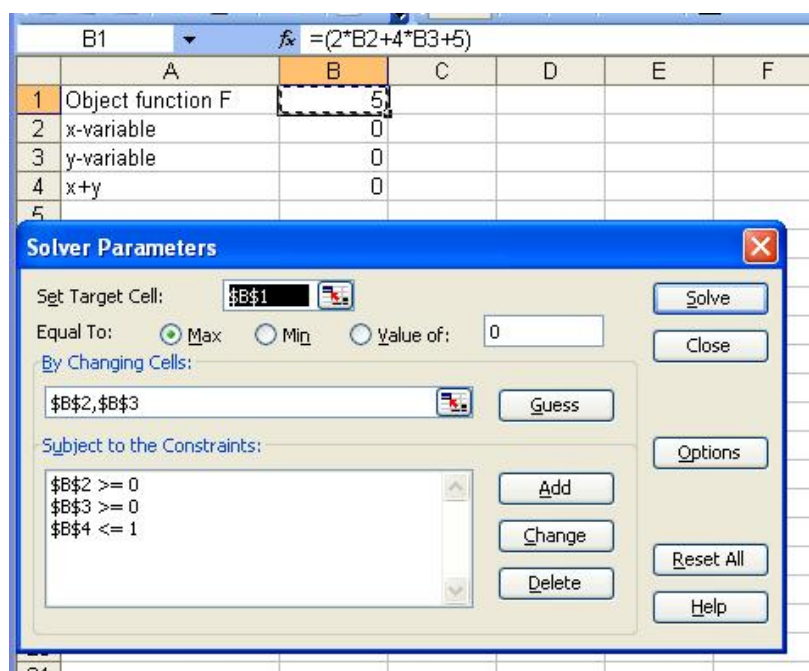


Figure 8

Click the <Solve> button. The “Solver Results” box announces that Solver has found a solution. At the same time, you will see that the



optimal values of F , x and y have appeared in the spreadsheet cells B1:B3.

$$F = 9, x = 0, y = 1.$$

These are exactly the values you found in Unit 15.

Example 11

A Malaysian company wants to produce an information guide about nutrition. The guide will be sold to non-governmental organisations (NGOs), schools, clinics, hospitals, community organisations, and so on. The guide will be published as a booklet and in the form of a video disc. The booklet will sell for 10 ringgits (MYR) and the video for 20 ringgits.

The company wants to maximise the income from sales. It estimates that combined sales will be no more than 12,000 units. It also seems that sales of the booklet will be at least 6,000 and sales of the video discs between 1,000 and 4,000 units. The printer has also advised the company that it only has 2,000 hours available for printing the booklet and that it takes 15 minutes on average to print a booklet.

As consultant to the company, you have to advise how many video discs and booklets it should produce in order to maximise its income.

Here's the solution:

Model the problem. Denote the number of booklets to be sold as B and the number of CDs as D . The object function is the income function $I(B, D) = 10B + 20D$ (each book sold brings in MYR 10 and each disc brings in MYR 20. So B booklets sold bring in $10 \times B$ ringgits).

The printing time constraint is that each booklet takes 0.25 hours, and for B booklets the time taken will be $0.25 B$ hours. Therefore, in terms of hours, you must have $0.25 B \leq 2,000$. Multiplying throughout by 4 gives $B \leq 8,000$.

The model is:

Maximise $(10B + 20D)$

Subject to constraints: $B + D \leq 12,000$

$$B \geq 6,000$$

$$1,000 \leq D \leq 4,000$$

$$B \leq 8,000$$

Use Excel to solve for B and D and find the maximum value of income I .

Create a workbook as shown in Example 10, using variables B and D instead of x and y , and activate Solver. Enter the constraints.

Note: If the problem did not have the constraints $B \geq 6,000$ and $D \geq 1,000$, you should also add the constraints $B \geq 0$ and $D \geq 0$ to ensure you do not get negative numbers as answers.

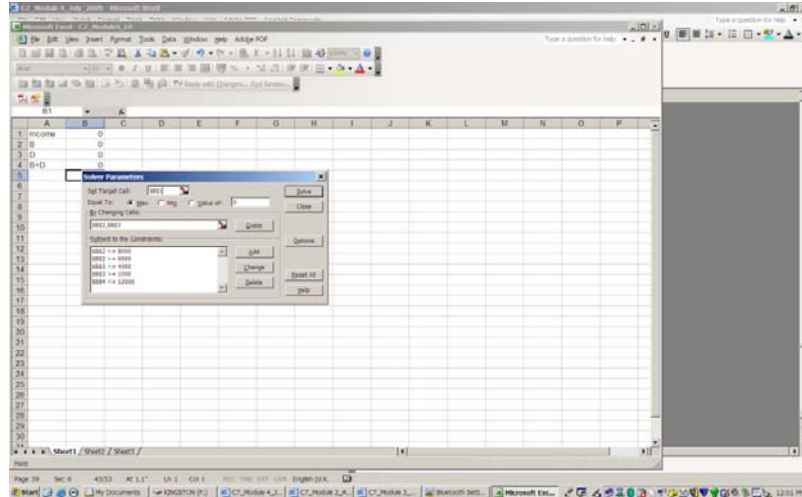


Figure 9

Click on <Solve>. This should appear in your worksheet:

Income	160000
B	8000
D	4000
B+D	12000

You can now present these recommendations to the company:

Arrange production of 8,000 booklets and 4,000 videos. This will give the maximum income of MYR 160,000.

Activity 4.8



Activity Minimise the cost

What will you do?

The cost C associated with manufacturing cables depends on the cost of copper wire and plastic tubing. Let x denote the number of metres of copper wire and y the number of metres of plastic tubing. There are fixed costs of INR 200,000 before manufacturing even starts. The cost of copper wire is INR 200 per metre and the cost of plastic tubing is INR 5.45 per metre. Material is only delivered if you order a combined length of copper and plastic that is more than 500 m. On the other hand, restrictions on storage space mean the combined length of copper and plastic can be 1,000 m at most. As well, operationally you should always have at least 200 m of copper wire and 100 m of plastic tubing at the factory.

Use Solver to find the numbers of metres of copper wire and of plastic tubing that will minimise the cost.

The Solver tool for integer linear programming

In Example 11, we were fortunate that the optimal values for B and D happened to be integers, although we did not formulate the problem with



the constraints that they should be.

Example 12

Suppose the company advises you that the costs and constraints should be changed:

The income from a booklet is MYR 10.50 and from a video is MYR 20.45. The constraints change from $B \leq 8,000$ to $B \leq 7,900.5$ and $D \leq 4,000$ changes to $D \leq 2,000.76$. Also change constraint $B + D \leq 12,000$ to $B + D \leq 9,000$. The solutions are then:

Income	114407.6
B	6999.24
D	2000.76
B+D	9000

Fractional values for B and D are now a problem, as you cannot produce fractions of books or videos. Rounding off answers is also not always optimal. Your model is not robust enough to handle these changes and therefore should be changed to an ILP problem.

Excel handles this kind of problem very effectively:

1. Click <Add> in the “Add Constraint” box.
2. In “Cell Reference”, enter B2, in the next box choose “int” and click <OK>.
3. You have entered the constraint “ B is an integer”. Click <Add> again, in “Cell Reference” enter B3, in the next box choose “int” and click <OK>.
4. You have entered the constraint “ D is an integer”.
5. Now click <Solve>. This is the solution that appears:

Income	114400
B	7000
D	2000
B+D	9000

Arrange production of 7,000 booklets and 2,000 videos. This will give the maximum income of MYR 114,400.



Discussion

Sub-optimality tolerance

Before you accept this as the answer, be aware that Solver uses a suboptimal tolerance factor of 5 per cent. This means that the Solver algorithm stops searching for an integer solution when it is within 5 per cent of the optimal integer solution. To make sure that Solver finds the actual best solution, we change the tolerance factor to zero in the following way.

In the Solver Parameter box, after you have entered the constraints, click <Options>. Click in the Tolerance box, delete 5 and enter 0. Click <OK>.

Now click <Solve>. In this example we get the same solution, but in general you may get a different answer that is the better solution.

The Solver tool for non-linear programming

Solver uses the generalised reduced gradient method to solve NLP problems. As mentioned previously, NLP problems can have more than one local optimal value and, unfortunately, Solver does not tell you if the solution it found is local or global. It is a good idea to run the algorithm a number of times from different starting values to see whether you get different solutions. In general, if your starter values lie close to a certain local optimal value, the Solver algorithm will stop at that optimal value and not move on to find the actual global solution.

Note: When you work with non-linear problems, it is best to not use starting values of zero (0) for all the relevant variables, as Solver does not always like starting from 0. However, zero starting values are fine for linear problems.

We will now apply Solver to the NLP problem of portfolio selection discussed in Unit 15, “Introduction to NLP”. Fortunately, it is known that the portfolio problem has a unique global solution, so we only run the problem once from one set of starter values.

Finding optimal portfolios of risky shares

Consider a portfolio P of three shares X1, X2 and X3. It is known that their average annual rates of return over the last five years were:

$$E(R_{X1}) = 7.6\%, E(R_{X2}) = 14.4\% \text{ and } E(R_{X3}) = 16.9\%$$

Their annual volatilities are estimated as:

$$\sigma_{X1} = 25.1\%, \sigma_{X2} = 45.25\% \text{ and } \sigma_{X3} = 49.2\%$$

We also have their co-variances:

$$\sigma_{X1,X2} = -0.025, \sigma_{X1,X3} = 0.044 \text{ and } \sigma_{X2,X3} = -0.0542$$

The problem lies in determining the proportions x_1 , x_2 and x_3 of the investor’s money that should be put in shares X1, X2 and X3, respectively. The investor wants to minimise the risk but requires a rate of return of at least 12 per cent on portfolio P. No short-selling is allowed.

We formulate the problem as follows (refer to Unit 15, “Introduction to NLP”):

$$\text{Minimise } F(x_1, x_2, x_3) = \sigma(R_P) = \sqrt{\sum_{i \neq j=1}^3 x_i^2 \sigma^2_{X1} + 2x_i x_j \sigma_{X_i, X_j}}$$

Subject to the constraints:

$$E(R_P) = \sum_{i=1}^3 x_i E(R_{X_i}) \geq 0.12$$

$$\sum_{i=1}^3 x_i = 1; \quad x_1, x_2 \text{ and } x_3 \geq 0$$



Modelling the NLP problem

Rewrite the problem with the data included.

Minimise $\sigma(R_p) =$

$$\sqrt{x_1^2(0.063) + x_2^2(0.2048) + x_3^2(0.242) + 2x_1x_2(-0.025) + 2x_1x_3(0.044) + 2x_2x_3(-0.0054)}$$

Subject to:

$$x_1(0.076) + x_2(0.144) + x_3(0.169) \geq 0.12$$

$$x_1 + x_2 + x_3 = 1; x_1, x_2 \text{ and } x_3 \geq 0.$$

Open an Excel workbook with cell names in A1 to A9 as shown, and enter the data in cells B1 to B9.

Enter the formulas for $E(R_p)$ and $\sigma(R_p)$ in B11 and B12.

Use starter values for x_1 , x_2 and x_3 as indicated in cells B13, B14 and B15.

Note: We did not choose 0 for all three starter values because Solver will not recognise those values.

Enter the formula for $\sum_{i=1}^3 x_i$ into cell B16.

E(RX1)	0.076
E(RX2)	0.144
E(RX3)	0.169
vol(X1)	0.251
vol(X2)	0.4525
vol(X3)	0.492
cov(X1,X2)	-0.025
cov(X1,X3)	0.044
cov(X2,X3)	-0.054
E(RP)	0.076
vol(RP)	0.251
x1	1
x2	0
x3	0
Sum(xi)	1

Figure 10

Click <Tools> and <Solver>. Click on “Min” for cell B12 and enter B13, B14, B15 for “by changing cells”. Enter the constraints. Click on <Solve>.

Cells B11 to B16 should look like this:

E(RP)	0.12
vol(RP)	0.201919
x1	0.434607
x2	0.343263
x3	0.22213
Sum(xi)	1

Figure 11

Analysing the portfolio solution

Your investor can now make an informed decision based on the Excel solution to the problem. The values indicate that the investor should invest about 44% of the money in share X1, 34 per cent in share X2 and 22 per cent in share X3.

This will give a portfolio P with an expected rate of return of 12 per cent and risk in the form of volatility of 20.19 per cent.

This risk is quite low if you consider that each share on its own has volatility in excess of 25 per cent. How is this possible? It is the negative co-variances that reduce risk. This is why portfolios should be diversified and contain a number of shares, some of which should be negatively correlated.

In simplistic terms, shares from different sectors (such as financials, resources, retail) have a good chance of being negatively correlated.

Activity 4.9



Activity Solve an ILP problem

What will you do?

1. Use Excel to solve this ILP problem.

Maximise $(2x + 3y - 5z)$ subject to:

$$x + y \leq 12, x + y + z \leq 20, x \leq 10, x, y \text{ and } z \text{ positive integers.}$$

2. In the portfolio problem above, why can you not use these starter values: a: $x_1 = 1, x_2 = 1$ and $x_3 = 0$?

Change the co-variances to positive values:

$$\sigma_{x_1, x_2} = 0.025, \sigma_{x_1, x_3} = 0.044 \text{ and } \sigma_{x_2, x_3} = 0.0542$$

and repeat the optimisation problem.

Discuss your findings.



Case study:
Employee work schedules

A factory operates seven days a week, 24 hours a day, with workers doing shifts. On certain days, more workers are needed to finish off and load products for shipping. The manager estimates the minimum numbers of workers needed on different days are:

Mon	Tue	Wed	Thu	Fri	Sat	Sun
18	12	20	12	22	20	8

On Mondays at least 18 workers are needed; on Tuesdays at least 12 workers, and so on. The shifts are 24 hours long but designed to let workers rest in between. Certain shifts are better paid because they extend late at night or are on Saturdays. There are four kinds of shift, described as:

Shift type	Days on duty	Weekly wage
1	Monday, Wednesday, Saturday	USD 560
2	Tuesday, Friday, Sunday	USD 570
3	Monday, Thursday, Saturday	USD 550
4	Wednesday, Friday	USD 300

Problem: The manager wants to minimise wage expenses.

How many workers should the manager assign to each shift so that the minimum number of workers will be available each day as required?

Hint: Assign variables as follows:

x_1 = number of workers assigned to Shift 1, x_2 = number of workers assigned to Shift 2, x_3 = number of workers assigned to Shift 3 and x_4 = number of workers assigned to Shift 4.

Formulate the object function as the total weekly wage expense. Then formulate the seven constraints and use Solver in Excel. Remember, this is an ILP problem. To help you, here is one constraint: since you must have at least 18 workers on Mondays, you must have $x_1 + x_3 \geq 18$.

Activity 4.10



Activity

Understand the terminology

What will you do?

Use this terminology table to record any terms or words you're uncertain about.

This activity is an opportunity to consolidate your understanding of new terminology and concepts you encountered in Unit 15. Fill in the terms you have learned and then write your own descriptions of them.



Terminology

Term	Description
------	-------------

:

:

:

:

:

:

:

:

:

:

:



Remember these key points

- Linear programming (LP) allows you to solve problems involving minimising or maximising a linear object function $F(x, y)$, subject to linear constraints on variables x and y .
- The object function can be cost, profit, and so on, and can be represented by a plane in 3-D with a family of parallel contour lines down in the (x, y) -plane.
- The constraints define the feasible region in the (x, y) -plane.
- The aim is to find that corner point (x^*, y^*) of the feasible region where the straight line representing the optimal value of the object function touches it.
- The most used method is the simplex method, which is based on moving from corner point to the next corner to find the optimal value of F .
- In integer linear programming (ILP), one of the constraints is that the solutions for variables x and y must be integers, which complicates the situation. In non-linear programming (NLP), the object function and/or the constraints are non-linear. These two types of problem are best solved using a computer. When more than two variables are involved, it is also necessary to use a computer.
- Excel has an excellent tool, Solver, to handle all three instances and for any number of variables. In particular, it can easily handle the selection of shares for an optimal portfolio.

Unit summary

You have successfully completed this unit if you can:



Summary

- **identify** linear (LP), integer linear (ILP) and non-linear programming (NLP) problems;
- **formulate** various object functions and constraints;
- **solve** simple LP problems using graphic and corner-point methods;
- **set up** portfolio selection problems based on the Markowitz theory;
- **apply** the Solver tool in Excel to solve LP, ILP and NLP problems; and
- **analyse** solutions in order to make informed decisions.



Unit 16

Forecasting: Time series and linear regression

Introduction

Managing a company includes trying to build up information about the future so you can make decisions that will position your company favourably. You have seen throughout the previous modules that many events are risky and uncertain because you do not know what the future will bring. At best, you can try to assign probabilities to outcomes or try to predict the future values of important business variables. Such variables include future sales figures, future interest rates, future costs and future trends.

This unit discusses the general ideas behind forecasting and the methods for predicting future values. You will see how to identify and quantify patterns that can help with planning and strategising.

Upon completion of this unit you will be able to:



Outcomes

- **describe** different methods for forecasting in management;
- **calculate** trends with moving averages;
- **apply** forecasting with time series;
- **apply** basic linear regression in forecasting; and
- **explain** the limitations of forecasting methods.

Methods of forecasting

There are two approaches to forecasting: **qualitative** and **quantitative**. Both have a role to play.

Qualitative methods

These methods are sometimes called judgemental or subjective – that is, they depend on people’s opinions to a large extent. Qualitative approaches include:

- personal assessment,
- group consensus,
- market research,
- historical assessment, and
- the Delphi method.

Personal assessment

This method can be effective because people who have a lot of experience in a certain area of work often have a reliable “gut feeling” about future developments. Workers on machines have a good sense for when the machine may be close to breaking down; sales staff have a good

idea about how long stocks will last; and experienced stock brokers may have a feeling for future financial trends.

However, these are people's personal feelings and assessments and they are usually affected by bias, prejudice or lack of knowledge. For this reason, you should be careful when using personal assessments (although they are useful in checking predictions made by other methods).

Group consensus

This is an extension of the personal assessment method. You ask for a range of opinions about a future event or a panel of people get together to discuss information and assumptions. Through consensus, they may be able to give a forecast for a business variable such as future sales figures.

However, this method depends on group dynamics and whether people are able to discuss ideas openly and reach consensus. The process may also be time consuming.

Market research

Suppose your business is thinking of launching a new product. You may contract a market researcher to test the public reaction to the new concept. This information can be gathered through questionnaires or interviews.

The important fact here is that only a small section or sample of the population of potential customers can be questioned. Therefore, the results have to be analysed statistically and interpreted carefully. This is where the tools you acquired in Module 2 become very important. (See Unit 8 on "Sampling and hypothesis testing".)

Historical assessment

If a company has good historical data about past sales figures, it may be analysed to forecast future figures. There are often patterns of demand that can be found in past data. For example, new products tend to follow a cycle of getting known in the market (slow sales), then growth (good sales as people become excited by the new product), reaching a plateau (the product is established) and then possibly declining (as people get bored with the product). Historical data on share prices is often used to calculate the expected future returns and volatilities of shares.

Again, care must be taken when basing future forecasts on past data. The turkey had 364 days' worth of historical data showing he always received excellent treatment and feed, and he forecast the same for the future. Little did he know.

The Delphi method

This method involves sending a questionnaire to a panel of experts who usually don't know each other. Their answers are combined in a report, which is sent back to the panel members. They can then revise their original answers to the questionnaire and send back the result. The process may be repeated until consensus is reached. The method uses the expertise of a group of people and can be very useful. Using electronic communications such as email makes this method cheap and fast.

(The name of the method comes from the story of the Oracle of Delphi in Greek mythology. Dating back to 1400 BC, the Oracle of Delphi was the



most important shrine in Greece. People came from all over to ask the Pythia, Apollo's priestess, questions about the future. Because the answers were usually cryptic, the correct interpretation of an "oracle" was rather vague. The *Oxford English Dictionary* defines an oracle as a person or thing regarded as an infallible guide to future events.)

Quantitative methods

There are two types of quantitative methods of forecasting: projective or trend forecasting methods and causal methods.

Projective forecasting uses historical data to forecast future values and trends. Past trends are projected into the future. We will discuss moving averages and time series analysis as a tool for forecasting.

Causal methods look at dependencies between quantities and the effect of outside influences that cause variables to behave in predictable ways. We will introduce linear regression as the tool for forecasting based on dependencies.

Moving averages and exponential smoothing

Rather than embarking on a full analysis of a time series, we only want to forecast a future value or trend based on the average of historical values. When we average, we tend to smooth out irregularities in the past data. We establish the trend in the data by using either a simple average or a moving average method.

The **simple average method** takes all data values collected over N time periods, y_t , $t = 1, 2, \dots, N$ and calculates the average. This is taken as F_{N+1} , the forecast for the next time period where $t = N+1$.

$$\text{Simple average method: } F_{N+1} = \frac{\sum_{t=1}^N y_t}{N}$$

(Remember: \sum means "sum".) The method works well as long as there are no recent dramatic changes in observations that indicate a change in trend. In such circumstances, it would be better to give old data less weight or to ignore old data completely.

Example 13

A newly opened shop has weekly sales figures for six weeks. It wants to project sales figures for the next week or two.

Week t	Number of items sold y_t
1	301
2	312
3	310
4	313
5	320
6	308

By the simple average method, the forecast value for week seven is:

$$F_7 = \frac{\sum_{t=1}^6 y_t}{6}$$

$$= 310.67 = 311 \text{ (rounded off)}$$

Testing the simple averages model yields this forecast for week six:

$$F_6 = \frac{\sum_{t=1}^5 y_t}{5}$$

$$= 311.2 = 311 \text{ (rounded off)}$$

This compares well with the actual value of $y_6 = 308$.

We should calculate the forecast values over all time periods and compare them with the given historical data, to determine how accurately our method can then forecast future values. The trend in this example seems to be constant.

The **moving averages method** ignores data older than a specified period of time. In the previous example, you may decide to use only data from the previous two weeks to forecast the next week. Weeks two and three are averaged to forecast week four; weeks four and five are used to forecast week six, and so on. In this way, the averaging process moves forward and the two-week moving averages forecast for week seven is:

$$F_7 = \frac{\sum_{t=5}^6 y_t}{2}$$

$$= \frac{320 + 308}{2} = 314.$$

Testing this model for week six:

$$F_6 = \frac{\sum_{t=4}^5 y_t}{2}$$

$$= \frac{313 + 320}{2} = 316.5 = 316 \text{ (rounded down)}.$$

This value seems less accurate than the simple averages method, but moving averages are generally better at smoothing out irregularities and indicating the trend in the data.

Once again, you should calculate all forecast values F_1 to F_6 so you can analyse the accuracy of the moving averages method for your data. This is explained in the section on forecast errors.

The moving averages method can be extended in many ways. Suppose you have monthly data for 60 months, but decide that only the last six months are relevant. You then take the average over the last six months to



find the six-month moving averages forecast value F_{61} :

$$F_{61} = \frac{y_{60} + y_{59} + \dots + y_{55}}{6}$$

For the next forecast F_{62} you *move the average forward* by using observations $y_{61}, y_{60}, y_{59}, y_{58}, y_{57}, y_{56}$.

The choice of how many data values to use is yours. In general, the moving averages method for forecast value F_{N+1} using n periods of data is:

$$F_{N+1} = \frac{\sum_{t=N+1-n}^N y_t}{n}$$

A disadvantage of this method is that all past observations are treated as equally important.

The **weighted moving averages method** assigns weights to data, according to how important the data values are. The weights are positive numbers between zero and one and must add up to one. The forecast value F_{N+1} using n periods of data with weights $w_1, w_2 \dots w_n$ then becomes:

$$F_{N+1} = \frac{w_1 y_N + w_2 y_{N-1} + \dots + w_n y_{N+1-n}}{n}$$

with $0 \leq w_i \leq 1$ and $\sum w_i = 1$ for $i = 1, 2, \dots, n$

How do you choose the weights? You can decide that recent data should carry greater weight than older data. For example, for the six-month moving averages forecast value F_{61} above, you can decide that data values $y_{60}, y_{59}, \dots, y_{55}$ should carry the weights $w_1 = 0.4, w_2 = 0.2, w_3 = 0.15, w_4 = 0.1, w_5 = 0.1$ and $w_6 = 0.05$.

$$\text{Then } F_{61} = \frac{0.4y_{60} + 0.2y_{59} + 0.15y_{58} + 0.1y_{57} + 0.1y_{56} + 0.05y_{55}}{6}$$

You can also decide to choose the weights so that the forecast error is minimised.

Forecast errors

We need to measure the errors in forecasting to decide on the accuracy of our forecast and to compare forecasting methods. The error is the difference between the actual and forecast value.

$$e_t = y_t - F_t \text{ and } t = 1, 2, \dots, N$$

where e_t is the error at data point or time t , y_t is the actual observation at time t and F_t is the forecast value at time t .

Clearly, if $y_t > F_t$ for some t , then we have a positive error: $e_t > 0$

If $y_t < F_t$ for some t , then we have a negative error: $e_t < 0$

However, we are mostly interested in the **magnitude** of the error. For example, if $y_1 = 2.5$ and $F_1 = 2.7$, then $e_1 = 2.5 - 2.7 = -0.2$. The magnitude of e_1 is $+0.2$.

Remember: The magnitude of any negative number is its positive equivalent and the magnitude of any positive number is itself. The mathematical terminology is “absolute value” or “modulus” and the symbol is $| \quad |$

So $|-3| = 3$ and $|3| = 3$

The accuracy of the specific forecast will depend on the overall accuracy of the method — that is, the size of the overall error for the N data values. We can use mean error, mean absolute deviation or mean squared error.

Measures of error

$$\text{Mean error ME} = \frac{\sum_{t=1}^N e_t}{N}$$

$$\text{Mean absolute deviation MAD} = \frac{\sum_{t=1}^N |e_t|}{N}$$

$$\text{Mean squared error MSE} = \frac{\sum_{t=1}^N (e_t)^2}{N}$$

You should calculate these values for different moving average methods, the 2-week average method and the 4-week average method, to decide which method is best for your particular problem.

The Solver tool for the weighted moving averages method

We mentioned that the weights for the weighted moving averages method can be chosen so that the forecast error is minimised. We choose to minimise the MSE. The formula for the MSE will then contain the weights $w_1, w_2 \dots w_n$. How is the MSE minimised with respect to the weights? We apply the Solver tool in Excel.

The method is exactly the same as the one applied in Section 15.3 when we chose the weights of shares in a portfolio to minimise the risk. The object function is now the MSE and the constraints are simply $0 \leq w_i \leq 1$ and $\sum w_i = 1$ for $i = 1, 2 \dots n$

However, the expression for the MSE will be quite messy and you will need to type in the formula very carefully.



Activity 4.11



Activity Establish trends

What will you do?

Consider the table of observed values y_t and forecast values F_t for six time steps t . The forecast values were found by the moving averages method.

t	1	2	3	4	5	6
y_t	6.2	11	13.02	15.68	18.8	23
F_t	7	10	13	16	19	22

1. Can you establish a trend in the data?
2. Expand the table by adding rows for e_t , $|e_t|$ and $(e_t)^2$.
3. Determine the ME.
4. Determine the MAD.
5. Determine the MSE.

Here's our feedback

1. The trend is one of rising values.
- 2.

t	1	2	3	4	5	6
y_t	6.2	11	13.02	15.68	18.8	23
F_t	7	10	13	16	19	22
e_t	0.8	1	0.02	-0.32	-0.2	1
$ e_t $	0.8	1	0.02	0.32	0.2	1
$(e_t)^2$	0.64	1	0.0004	0.1024	0.04	1

$$3. \text{ ME} = \frac{1}{6} [-0.8 + 1 + 0.02 - 0.32 - 0.2 + 1] = 0.1167$$

$$4. \text{ MAD} = \frac{1}{6} [0.8 + 1 + 0.02 + 0.32 + 0.2 + 1] = 0.5567$$

$$5. \text{ MSE} = \frac{1}{6} [(0.8)^2 + 1^2 + (0.02)^2 + (0.32)^2 + (0.2)^2 + 1^2] = 0.4638$$



Discussion

Interpretation of errors

The mean error can be problematic because rather than always showing how good the forecast is, it gives the bias.

t	1	2	3	4
y_t	10	20	30	40
F_t	2	2	2	94
e_t	8	18	28	-54

For example, for the values shown above, the mean error is zero ($\sum e_t = 8 + 18 + 28 - 54 = 0$), even though the forecast is clearly very poor. What you can say is if $ME > 0$, then the forecasts are, on average, too low, and if $ME < 0$, then the forecasts, are on average, too high.

The MAD gives the average distance each forecast is from the actual observation at that point.

MSE is useful in that it emphasises large errors and minimises small errors.

Exponential smoothing is very popular and based on the idea that old data is less important than recent data. A smoothing constant α , with a value between 0 and 1, is applied to the previous observation and factor $(1 - \alpha)$ is multiplied by the previous forecast.

This is done recursively:

Make a forecast F_1 for observation y_1 . Then $F_2 = \alpha y_1 + (1 - \alpha) F_1$, $F_3 = \alpha y_2 + (1 - \alpha) F_2$ and so on. In general:

$$F_t = (1 - \alpha)F_{t-1} + \alpha y_{t-1}$$

$$\text{or } F_t = F_{t-1} + \alpha (y_{t-1} - F_{t-1})$$

where: F_t = new forecast (for time period t)

$$F_{t-1} = \text{previous forecast (that is, for time period } t - 1)$$

$$y_{t-1} = \text{previous period's actual value}$$

The term $(y_{t-1} - F_{t-1})$ represents the forecast error.

The smoothing constant α can be changed to give more weight to recent data when the value is high or more weight to past data when it is low. For example, when $\alpha = 0.5$, it can be shown that the new forecast is based almost entirely on demand in the past three periods. When $\alpha = 0.1$, the forecast places little weight on any specific period, even the most recent, and it takes many periods (about 19) of historic values into account.

**Example 14**

Calculate F_2 , F_3 and F_4 for $\alpha = 0.2$.

t	y_t	F_t
1	6.2	6 (make a forecast for F_1)
2	11	
3	13.02	

1. $F_2 = 6 + 0.2(6.2 - 6)$
 $= 6.04$
2. $F_3 = 6.04 + 0.2(11 - 6.04)$
 $= 7.032$
3. $F_4 = 7.032 + 0.2(13.02 - 7.032)$
 $= 8.2296$
 $= 8.23$

Note 1: When the formula is used in the form $F_t = F_{t-1} + \alpha(y_{t-1} - F_{t-1})$, we see that the new forecast equals the previous forecast plus α times the forecast error.

Note 2: The recursive formula can be “unpacked” so that everything is expressed in terms of the given data only, plus the first forecast F_1 . This is useful if we only want future forecasts, where data is not available, and not forecast values for times where we already have historical data. For example, we have data for times $t = 1 \dots N$ and need a forecast for time $t = N + 1$. The formula is then:

$$F_{N+1} = \sum_{k=1}^N \alpha(1-\alpha)^{k-1} y_{N+1-k} + (1-\alpha)^N F_1$$

Activity 4.12



Activity Forecast values

What will you do?

1. Consider the table of observed values y_t for six monthly time steps t .

t	1	2	3	4	5	6
y_t	9.4	10	11.02	10.8	11	11.3

- Use the simple averages method to find a forecast for F_7 .
 - Use the two-month moving averages method to find a forecast for F_7 .
 - Use the four-month moving averages method to find a forecast for F_7 .
 - Apply exponential smoothing with $\alpha = 0.2$ to find a forecast for F_7 . Let $F_1 = 9$.
2. Consider the table of observed and forecast values.

t	1	2	3	4
y_t	2.1	2.3	2.6	2.4
F_t	2	2.4	2.5	2.6

Determine the ME, MAD and MSE.

Time series models

A time series consists of a series of observations taken regularly over a period of time. For companies, these may be the records they keep of weekly demand, monthly sales or annual profits.

Time series models are aimed at forecasting in situations where there are trends or seasonal effects. The analysis of a time series depends on identifying patterns in the historical data, which is made easier by presenting the data as a scatter diagram or graph.

The observations form points (t, y_t) where t represents time along the horizontal axis and y_t are the actual observed values along the vertical axis. The t variable can be days, weeks, months, years or any suitable time period.

From the graph of data points, a pattern is identified and used to forecast future values. The three basic patterns for time series are **constant**, **trend** (rising or falling) and **seasonal**. A smooth line can be fitted to the data and used to forecast future values.



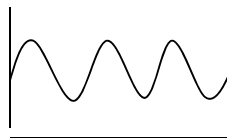
Examples of fitted lines for forecasting



Constant series

Series with rising trend

Figure 12



Seasonal series

Figure 13

The points on the smoothed graphs representing the patterns are (t, F_t) where the F_t values on the pattern form the values that will be used as forecasts. (The actual data points are not shown in Figures 12 and 13.)

Time series from actual data do not have such perfect shapes – there is random noise superimposed on the patterns shown above. The random noise is also called the “error” and denoted as e_t . So, as with moving averages methods, we have:

Actual observation = underlying pattern + noise (error)

$$y_t = F_t + e_t \text{ for } t = 1, 2, 3 \dots N$$

The amount of noise needs to be measured and evaluated to decide how valid the identified pattern has been. In turn, this affects the validity of the forecasting. If the noise is substantial, it will be difficult to identify a pattern and make forecasts.

Example 15

The average summer temperature for a town in Sri Lanka has been documented over 10 years. The scatter graph of actual data is denoted by the points on Series 1 and looks like this:

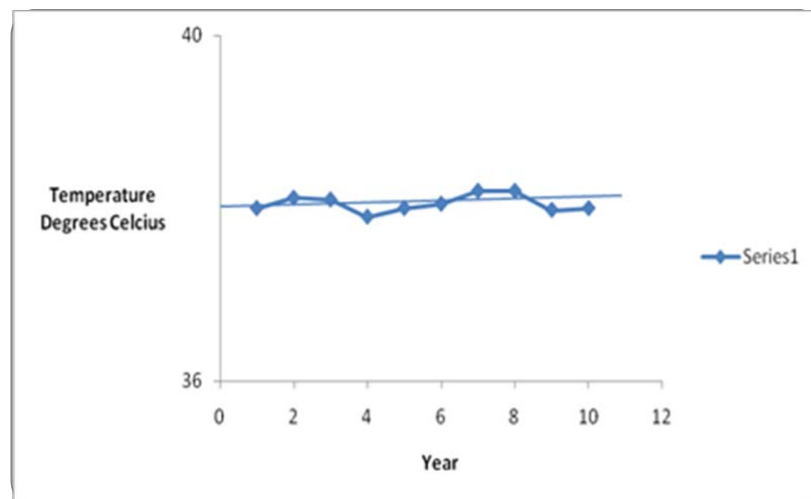


Figure 14

This is an example of a constant time series. The observations are more or less the same over time and a constant pattern can be identified. We have fitted a line to the data and this line can now be used to forecast future values.

There is some noise, so the line is not completely horizontal. The noise can be measured by finding the vertical distances from the data points to the straight line.

Example 16

Quarterly electricity consumption by a rapidly expanding industry over the past few years is displayed by the Series 1 points in Figure 15.

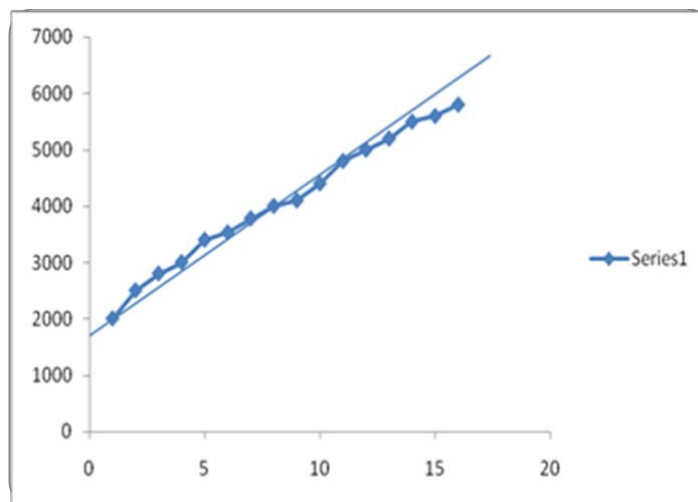


Figure 15

This is an example of a time series where a trend can be identified. We fit a rising straight line to the observations and use this line to forecast future values. (Unit 16, “Linear regression”, shows how to fit a straight line to data.)

Once again, we can see the noise component and measure it by looking at the vertical distances from the data points to the straight line.

Example 17

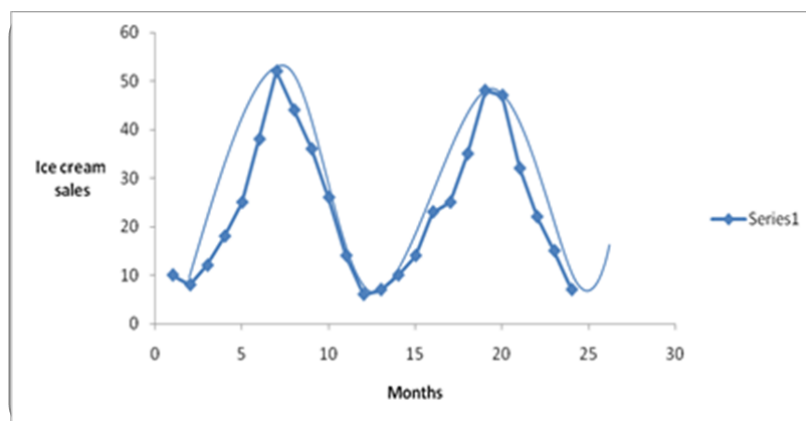


Figure 16



This time series shows a seasonal trend. It becomes more difficult to fit a smooth curve to the observations to capture the pattern.

Decomposing a time series

A time series consists of three components: trend, seasonality and noise (random elements or error). The trend is the long-term behaviour: rising, falling or flat. The seasonality is the existence of a repeating pattern. A time series can combine seasonal patterns with a trend – for example, a wave pattern that slopes upwards or downwards. The noise (error) can reflect unpredictable events that have affected the series of values. In this module, we will accept a fixed seasonal component and only forecast the general underlying trend by a straight line.

Measuring noise in patterns

We will assume a trend has been discovered and a line has been fitted. Exactly as in Unit 16, we analyse the overall effect of errors e_t . In Figure 14, a rising trend line has been fitted to the data. The length of the distance between each pair of points y_t and F_t is denoted by the bracket } and denotes the error e_t . Data points y_t are denoted by \blacktriangle and forecast points F_t are shown on the line by \bullet .

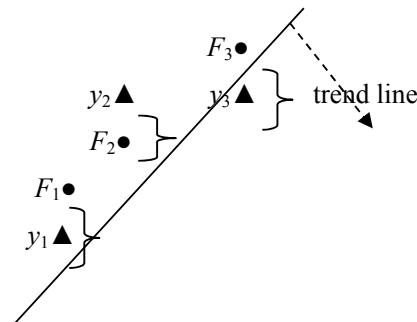


Figure 17

The “fit” of the pattern will depend on the size of the overall error. You can use mean error, mean absolute deviation or mean squared error, exactly as in Unit 16.

In the next unit, you are shown how to actually fit a straight line to data to capture the trend and forecast future values.

Linear regression

Previously in this module, you looked at cases where data was in the form (t, y_t) where quantity y is observed over time t . In general, data can be any observations that are in the form of pairs of values (x_t, y_t) and collected over time where t (time) = 1, 2, 3... N . You have already considered how having a model to describe the relations between data allows you to make predictions and decisions concerning the situation.

Now you will see how to specifically fit a straight line to data. Look at the following scatter graph of data points (x_t, y_t) , $t = 1, 2, 3... 10$. Variables x and y represent two quantities you have been observing, such as sales figures y and staff motivation levels x over 10 months. So point

(x_3, y_3) is the point (6, 160) and indicates that at motivation level 6 the sales figure was 160.

There appears to be an approximate linear relation between the quantities x and y .

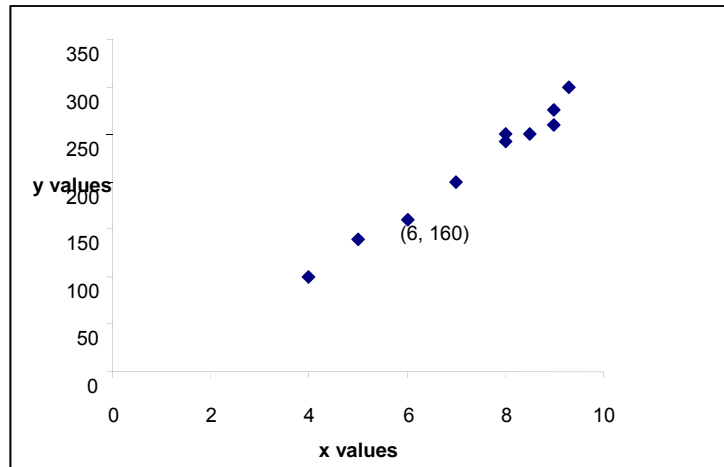


Figure 18

How do you actually obtain the straight line that best fits the observations (x_i, y_i) and use it to forecast other values? And what is meant by “best fit”? Having a good model to describe the relations between data allows you to make predictions and decisions about the situation.

Note: Relationships can be causal when changes in one quantity (the independent variable) actually cause the changes in the second quantity (the dependent variable). This could be for a situation where staff motivation causes an increase in sales or productivity. Other relations between data (x_i, y_i) , or in time series data (t, y_t) need not be causal.



Note It?

Causal relations

Unit 1 of Module 1 discussed how the human brain seems to be programmed to see patterns and relationships in sequences of numbers. A similar phenomenon arises when we observe quantities and see a relationship between them.

In the previous example, we observed sales figures and staff motivation levels over 10 months. There seems to be a linear relationship and it is possible the relation is causal: the more that staff are motivated, the higher sales figures will be. However, a similar scatter graph may apply to data (x_t, y_t) where x denotes sales of eggs and y denotes sales of raincoats. The linear relationship will be valid but not necessarily *causal*. We cannot use the linear relation to imply that buying more eggs causes people to buy more raincoats.

Always be aware that relations between quantities may show a pattern, but may not always be causal.

Read and discuss the following statement by David Brooks, which appeared in an article published in *The New York Times* on 28 October, 2008:

Looking at and perceiving the world is an active process of meaning-making that shapes and biases the rest of the decision-making chain... We should be aware that humans tend to see data that confirm their prejudices more vividly than the data that contradict ideas...

In this section, we ignore questions of cause and effect and concentrate on the method of linear regression. You are encouraged to review the sections in Module 1, Unit 2 that deal with straight lines and linear relations. We also do not consider cases where the relation appears to be non-linear.

The method of simple linear regression

Remember that the equation of a straight line is $y = mx + c$ where m is the slope and c the intercept on the vertical (y) axis. If the quantities are productivity P and bonuses R , then we may represent the perceived linear relation by either $P = mR + c$ or by $R = m^*P + c^*$. That is, we can decide which variable to represent on the vertical axis. (The two slopes m and m^* and intercepts c and c^* are, of course, not the same.)

The first step in fitting a line to data would be to draw a scatter graph of the data (x_t, y_t) , $t = 1, 2, \dots, N$ (as in the figure above) and see whether a linear relationship is likely. The data values x_t will be taken as the given x -values of the relation we are seeking to find. The relation we want to construct is of the form $\tilde{y} = f(x_t)$ where \tilde{y} now will lie on a line and approximate the actual y -values of the data. Denote the approximate values $\tilde{y}(x_t)$ as \tilde{y}_t .

Next, we must realise that if we fit straight line $\tilde{y} = mx + c$, then the data y_t will not fit the equation exactly — there will be an error e_t so that at each value of x_t we have:

$$e_t = y_t - \tilde{y}(x_t) = y_t - \tilde{y}_t = y_t - mx_t - c \text{ and } t = 1, 2, 3, \dots, N$$

In other words, for each of the N data points we have:

$$y_t = mx_t + c + e_t$$

The error e_t is also called the **residual**. It tells us how far above or below the fitted line the actual data y_t will lie.

The final step is to find the best fit (best choice of m and c) by minimising the error term. For linear regression, we choose to minimise the mean squared error (MSE). For this reason, linear regression analysis is also called the method of least squares.

$$\text{Minimising MSE} = \frac{\sum_{t=1}^N (e_t)^2}{N}$$

$$\text{means minimising } \frac{\sum (y_t - mx_t - c)^2}{N}$$

(Summations are for t from 1 to N .) The error MSE is a quadratic function in m and c because y_t and x_t will be known numbers. We know that the minimum of a function is obtained by differentiating. It can be shown mathematically that if this is done, we can solve for m and c . In fact, the straight line of best fit has the slope:

$$m = \frac{N \sum x_t y_t - \sum x_t \sum y_t}{N \sum x_t^2 - (\sum x_t)^2} \quad \text{and intercept } c = \bar{y} - m\bar{x}$$

Remember from Module 2, Unit 6 that \bar{y} and \bar{x} are the mean values of the y_t and x_t data values. The formulas are:

$$\bar{y} = \frac{1}{N} \sum_{t=1}^N y_t \quad \text{and} \quad \bar{x} = \frac{1}{N} \sum_{t=1}^N x_t$$

The calculations are time consuming if done by hand. We show you how to do regression on your computer after Activity 4.14.

The value of having a straight line as a model for the situation is that you can then use the model to forecast future values. Linear regression is used for forecasting values for variables.

In Activity 4.13, the regression is used to forecast profits for the anticipated future numbers of units sold. This is an actual causal relation – number of units sold is the independent variable and profit is the dependent variable. In other cases, the relationships need not be one of direct cause and effect.

We can also use regression for **interpolation**. Interpolation means finding in-between values.



Activity 4.13



Activity

Consider production data

What will you do?

Consider data (x_t, y_t) , $t = 1, 2, 3, 4$ for production in your factory.

Year (t)	Units manufactured (x_t)	Profit (y_t)
1	20	120
2	40	170
3	60	210
4	80	230

Profit is in thousands of dollars. Answer these questions.

1. Use linear regression to find the line of best fit for the data in the table.
2. Give a forecast for the profit that will be made if you anticipate that you will manufacture 114 units in the following year.
3. Use interpolation to find the approximate profit if 50 items are manufactured.
4. Draw the line of best fit for the data.
5. For a profit of 300,000 dollars, how many items should your company manufacture?

Here's our feedback

$$\begin{aligned}
 1. \quad m &= \frac{N \sum_{t=1}^4 x_t y_t - \sum_{t=1}^4 x_t \sum_{t=1}^4 y_t}{N \sum_{t=1}^4 x_t^2 - (\sum_{t=1}^4 x_t)^2} \\
 &= \frac{4 \times 40,200 - 200 \times 730}{4 \times 12,000 - 200 \times 200} \\
 &= 1.85 \\
 c &= \bar{y} - m \bar{x} \\
 &= 182.5 - 1.85 \times 50 \\
 &= 90
 \end{aligned}$$

The line of best fit is $y = 1.85x + 90$.

2. Profit $y = 1.85(114) + 90$
 $= 300.9$ thousand dollars
 $= 300,900$ dollars
3. Interpolation means to find the values in between given data.
 Profit $y = 1.85(50) + 90$
 $= 182.50$ thousand dollars

4.

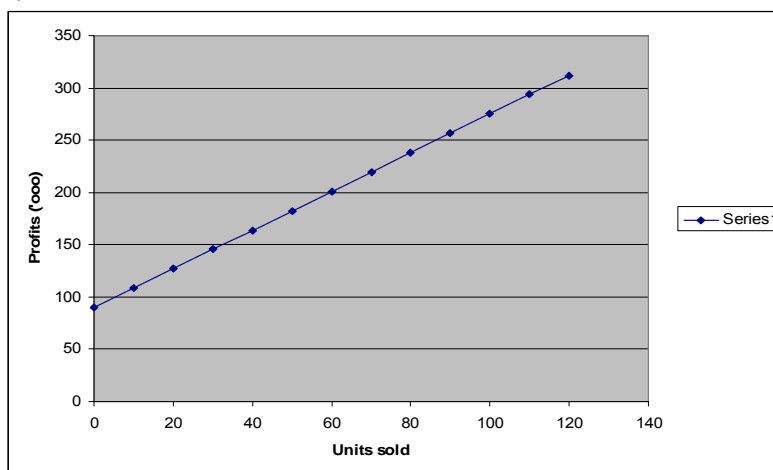


Figure 19

5. From Figure 19, we see that if Profit = 300, then items manufactured is approximately 115.

We can also use the regression equation $y = 1.85x + 90$ with $y = 300$.

$$\text{Solve for } x: x = \frac{210}{1.85} = 113.5$$

Measuring goodness of fit

There are ways to measure how well the linear regression fits the data.

1. The co-efficient of determination, denoted by r^2 , is one such measure.

$$r^2 = \frac{[N \sum x_t y_t - \sum x_t \sum y_t]^2}{[N \sum x_t^2 - (\sum x_t)^2] \times [N \sum y_t^2 - (\sum y_t)^2]}$$

The coefficient r^2 has a value between 0 and 1. The closer it is to 1, the more the variation in data is explained by the linear regression and we accept that the line is a good fit. A value greater than 0.5 indicates a reasonably strong linear relationship between x and y values.

Different areas of management require different values of r^2 to indicate a good fit of the data. For example, in economic models we want r^2 to be much larger than 0.5 and close to 1. In areas of job performance, we are satisfied with values of r^2 close to 0.5 or above.

We can also use the co-efficient of correlation r , which is just the square root of r^2 . Its values range between -1 and $+1$. A value of 1



indicates a perfect increasing linear relationship with no error. A value of -1 indicates a perfect decreasing linear relationship with no error. A value of 0 indicates no linear relationship.

Other important statistics in analysing the goodness of the regression are the R^2 , adjusted R^2 , multiple R and standard error. We define the total sum of squares (TSS), which describes the total variation of y

around its mean, as $TSS = \sum_{t=1}^N (y_t - \bar{y})^2$

This can be decomposed as:

$$\sum_{t=1}^N (y_t - \bar{y})^2 = \sum_{t=1}^N (y_t - \tilde{y}_t)^2 + \sum_{t=1}^N (\tilde{y}_t - \bar{y})^2$$

That is, $TSS = ESS + RSS$ (regression sum of squares).

ESS represents the variation of data y around its mean that the regression line cannot account for. RSS represents the amount of variation in y around its mean that the regression function can explain.

$$2. \text{ We then define } R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS}$$

If the regression line fits the data perfectly, then $ESS = 0$ and $R^2 = 1$.

If $RSS = 0$, then regression explains nothing and $R^2 = 0$. (R^2 is just another way of expressing r^2 above.)

3. Multiple R represents the strength of the linear relationship between actual and estimated values y_t and \tilde{y}_t . It is equivalent to the square root of the R square in cases (such as the above) where there is only one independent variable.
4. Standard error S_e measures the standard deviation of the error and gives the accuracy of the predictions obtained from the regression model.

$$S_e = \sqrt{\frac{\sum (y_t - \tilde{y}_t)^2}{N - k - 1}}$$

In practice, we calculate these co-efficients using spreadsheet packages such as Excel.

Activity 4.14



Activity Consider sales data

What will you do?

Consider data (t, y_t) , $t = 1, 2, 3$ for sales in your shop. Sales are in thousands of items.

Year (t)	Sales (y)
1	10
2	8
3	5

1. Use linear regression to find the equation of the line of best fit for the data in the table.
2. Draw the graph of the regression model for sales.
3. Determine the value of one of the measures of goodness of fit. Discuss its meaning.
4. Forecast the values of y at $t = 4$ and $t = 4.5$.
5. After how many months will you have no sales?

Multiple linear regression

If we assume that a quantity such as sales figure y is linearly related to the price x of the product only, then we have simple linear regression as discussed above: $y = mx + c$. However, if we assume that sales figures are also probably linearly dependent on variables such as spending on advertising, inflation, unemployment, and so on, then we enter the field of multiple linear regression.

The independent variable y (sales) is affected by independent variables x^I (price), x^{II} (advertising spending), x^{III} (inflation), and so on.

If there are k independent variables, then we are looking for a relation of the form $y = m_1x^I + m_2x^{II} + m_kx^k + c$. That is, we try to find the values of m_1, m_2, \dots, m_k and c so that the mean square error is again minimised.

Multiple regression becomes too complicated to do by hand and computer packages such as Excel are used instead.

Goodness of fit is measured by adjusted R square (R_a^2).

Adjusted R square (R_a^2) takes into account the number of independent variables in a regression model. $R_a^2 = 1 - \frac{ESS}{TSS} \left(\frac{N-1}{N-k-1} \right)$ where N is the number of observations and k is the number of independent variables.

We should only add more variables to a model if we can decrease ESS to get R_a^2 closer to 1. R_a^2 will decrease (which is bad) if we add independent variables that do not decrease the ESS enough.

Multiple regression results should be interpreted by looking at the regression statistics and the ANOVA table.



Linear regression with Excel

We saw that the formulas for determining the slope m and the measures of goodness of fit are cumbersome to apply when you do the calculations manually or using a pocket calculator.

In Excel, the Data Analysis Regression tool can do a regression analysis within seconds. When there is one independent variable x , it is called simple regression; when there are more independent variables x_i , it is called multiple regression.

Remember: In OpenOffice, you have to download macros to perform statistical analysis. These macros can be downloaded from <http://www.oocomacros.org/user.php>, which is linked to <http://documentation.openoffice.org/thirdparty.html>.

On the webpage for oocomacros, scroll down the list of macros to OOoStatistics and click on <Download>.

Example 18

Consider the trading expenses and sales of a business for the period 2000 to 2010 in this table. We expect a linear relation between the quantities, with trading expenses dependent on the sales figures.

Year	Sales	Trading expenses
2000	234	175
2001	346	250
2002	455	300
2003	610	455
2004	724	500
2005	778	678
2006	875	720
2007	1,024	900
2008	1,250	945
2009	1,300	1,040
2010	1,632	1,490

Enter the data in a spreadsheet and perform a regression analysis using Excel's Regression Data Analysis tool:

- Click <Tools>, <Data Analysis>, <Regression> and <OK>.
- Enter in the "Input Y Range" box: C2:C12.
- Enter in the "Input X Range" box: B2:B12.
- Enter in the "Output Range" box: A20.
- Click <OK>.

The table from Figure 20 will appear:

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.98784
R Square	0.975827
Adjusted R Square	0.973141
Standard Error	65.00366
Observations	11

<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	1535183	1535183	363.31606	1.3871E-08
Residual	9	38029.29	4225.476		
Total	10	1573213			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	-82.2095	44.41745	-1.85084	0.09722109	182.6887911	18.2697065	182.6887911	18.26970655
X Variable 1	0.905646	0.047513	19.06085	1.3871E-08	0.798163557	1.01312924	0.798163557	1.01312924

Figure 20

Interpreting the output

The first box contains the regression statistics.

Multiple R represents the strength of the linear relationship between actual and estimated values y_t and \hat{y}_t . It is equivalent to the square root of the R square in cases (such as the above) where there is only one independent variable.

R square (R^2) is the coefficient of determination r^2 that we introduced in Module 2, Section 6.3. It measures the proportion of the total variation in dependent variable y_t around its mean that is accounted for by the regression function. Multiple R and R^2 are close to 1 in this case, showing a strong linear relationship.

Adjusted R square (R_a^2) takes into account the number of independent variables in a regression model. $R_a^2 = 1 - \left(\frac{ESS}{TSS}\right)\left(\frac{N-1}{N-k-1}\right)$

In this case, $k = 1$, $N = 11$ and $R_a^2 < R^2$

Standard error $S \approx 65$ gives the variation of data around the fitted line.

ANOVA

ANOVA or **A**nalysis **O**f **V**ariance was discussed in Module 3.

The Excel printout for regression includes an ANOVA table. It has three rows: Regression (R), Residual (E) and Total (T).

The SS column is the “Sum of Squares”, which contains the values RSS, ESS and TSS which were introduced in Module 2, Unit 6.

The df column gives the degrees of freedom of RSS, ESS and TSS. Degrees of freedom refers to the number of values in the calculation of a statistic that are free to vary. You can also think of it as the number of independent pieces of information available to estimate another piece of information.

RSS contains one piece of information: \bar{y} , ESS contains $N-2 = 9$ pieces and TSS contains $N-2 = 10$ pieces of information.

The MS column is the mean square column, with $MSR = RSS/df$ and $MSE = ESS/df$.

The F statistic comes from the quotient MSR/MSE .

The F-value = 363.316 and its significance is the P-value.

The bottom box contains the co-efficients of the straight-line equation: the y-intercept is $c = -82,095$ and the slope is 0,9056. There is a linear relation: trading expense = $0.9056 \times \text{Sales} - 82,095$. Trading expenses increase with one unit approximately for every unit increase in sales.

The other values are the standard errors, t -statistics and P -values for intercept and slope. The standard error for the intercept is 44.4 and the standard error for the slope is 0.0475. This means a new set of data could give a slope that differs by 0.0475 and an intercept that differs by 44. The t -ratio is the ratio of the co-efficient to its standard error.



The analysis of statistics for regression can become very technical and we will not go into further details.

Activity 4.15



Activity

Do a regression analysis

What will you do?

- 1 Repeat Activity 4.14 using Excel. Discuss as many regression statistics as you can.
- 2 Do a regression analysis for the following data from five companies and interpret the output.

Bonus of CEO (millions of CAD)	Profit of company (millions of CAD)
1.40	110.890
2.00	148.567
4.58	200.55
6.45	185.86
10.1	260.380

Activity 4.16



Activity

Understand the terminology

What will you do?

Use this terminology table to record any terms or words you're uncertain about.

This activity is an opportunity to consolidate your understanding of new terminology and concepts you encountered in Unit 16. Fill in the terms you have learned and then write your own descriptions of them.



Terminology

Term	Description
------	-------------

:

:

:

:

:

:

:

:

:



Remember these key points

- A future value or trend can be forecast based on the average of historical values. When we average, we tend to smooth out irregularities in the past data. We establish the trend in the data by using either a simple average or a moving average method.
- We also need to measure the errors in forecasting to decide on the accuracy of our forecast and to compare forecasting methods. The error e_t is the difference between actual data y_t and forecast value F_t :

$$e_t = y_t - F_t \text{ and } t = 1, 2, \dots, N$$

- Errors can be measured in terms of mean error (ME), mean absolute deviation (MAD), or mean squared error (MSE).
- Linear regression is a tool for displaying data (x_t, y_t) , $t = 1, 2, \dots, N$ when there appears to be a linear relation between the quantities. A straight line $\hat{y} = mx_t + c$ is fitted to the data, with error $e_t = y_t - \hat{y} = y_t - mx_t - c$ and $t = 1, 2, 3, \dots, N$.

- The slope and y -intercept are determined so that the mean square error is minimised. Minimising $\frac{\sum(e_t)^2}{N}$ gives a straight line of best fit with slope

$$m = \frac{N \sum x_t y_t - \sum x_t \sum y_t}{N \sum x_t^2 - (\sum x_t)^2} \quad \text{and } c = \bar{y} - m\bar{x}$$

- The co-efficient of determination, denoted by r^2 , is a measure of how well the linear relation describes the data.

$$r^2 = \frac{[N \sum x_t y_t - \sum x_t \sum y_t]^2}{[N \sum x_t^2 - (\sum x_t)^2] \times [N \sum y_t^2 - (\sum y_t)^2]}$$

- The coefficient has a value between 0 and 1. The closer it is to 1, the more the variation in data is explained by the linear regression and we accept the line is a good fit.
- Other measures are R^2 , adjusted R^2 , multiple R and standard error. The regression analysis can be done efficiently using Excel.
- Regression can be used to make forecasts or to interpolate between data points.

Unit summary

You have successfully completed this unit if you can:



Summary

- **describe** different methods for forecasting in management;
- **calculate** trends with moving averages and exponential smoothing;
- **apply** basic linear regression in forecasting;
- **use** Excel to analyse regression statistics;
- **understand** forecasting with time series; and
- **explain** the limitations of forecasting methods.

Assessment



Assessment — Module 4

- Draw a binomial tree to represent the Markov price process $\{S_t : t = 0, 1, 2, 3\}$ for a share that starts with initial price $S_0 = 10$ and then either increases by 10 per cent per time step t or decreases by 10 per cent per time step t . The probability of an increase is 0.5 (50 per cent).
 - Is the tree recombining?
 - Calculate $E[S_1]$ as well as $E[S_2]$ and compare the values with S_0 .
 - Repeat the calculations in (b) above for the case where the probability of an increase is 0.7 (70 per cent). What conclusions can you draw?

- A fast food company wants to study the arrival of customers at the counter. It has estimated the probabilities of certain numbers of customers arriving at the counter during any two-minute period. The data is:

Number of customers	0	1	2	3
Probability	0.1	0.45	0.25	0.2

The company wants to simulate the number of customers arriving over 10 two-minute periods by generating strings of random numbers. The aim is to try to find out how many customers may arrive in a 20-minute period. This is important, because it cannot cope if more than 18 people arrive during a typical 20-minute period.

- Generate a column of 10 random numbers using function `RANDBETWEEN(0,99)` in Excel. Discuss your method.
 - Determine how many people arrived in a 20-minute period in this simulation.
 - Repeat the process 20 times (a sample of 20 runs) and determine the average number of people that will arrive in a 20-minute period. Discuss your results and compare with those of fellow students.
- A car manufacturer makes USD 1,000 profit on each model A car and USD 1,350 profit on each model B car. The assembly team can assemble a model A car in 30 minutes and a model B car in 45 minutes. A working day consists of eight hours.
 - Write down an expression for the per-day profit function P . Use a spreadsheet package to represent this by a plane in 3-D space.
 - Write down the daily time constraint. Represent the feasible region in a spreadsheet using Chart Wizard.
 - Give a graphical representation of the feasible region showing the profit function contour line that gives maximum daily profit.



- d) Find the optimal solution graphically.
- e) Verify the solution with the corner-point method.
- f) Use the Solver tool to find the optimal solution.
4. A profit function has the form $P(Q) = Q^3 - 2R^2 + 4QR - 3R$. The following constraints apply: $Q + R \geq 100$, $Q \leq 150$ and $R \leq 300$.
- a) Find the values of Q and R that maximise the profit. Calculate the profit in this case.
- b) If there were no constraints, what would the maximum profit be? Use both analytical (calculus) methods and the Solver tool.
5. Your country's Internal Revenue Service (IRS) asks you to help develop a model to detect whether companies inflate the figures of their contributions to charity in order to gain greater deductions on their tax returns. You have the following data for eight companies whose tax returns have been audited and accepted as truthful.
- a) Draw a scatter graph for the data and decide whether a linear relationship is likely.
- b) Do a regression analysis and derive a linear equation for the relation between gross income (G) and charitable contribution (C).
- c) Interpret the value of R^2 and evaluate your model.
- d) Write a report to explain to the IRS how it could use your model to detect instances where companies claim unusually high contributions to charity.

Gross annual income (millions)	Contribution to charity
66.7	100,000
72.84	115,000
95.1	135,800
110.25	140,000
129.56	165,500
166.4	175,000
182.3	200,000
200.75	225,000

6. Consider the table of observed values y_t for six annual time steps t .

t	1	2	3	4	5	6
y_t	352.5	340	341.5	329.8	320	311.3

- Use the simple averages method to find a forecast for F_7 .
 - Use the two-year moving averages method to find a forecast for F_7 .
 - Use the four-year moving averages method to find a forecast for F_7 .
 - Apply exponential smoothing with $\alpha = 0.1$ to find a forecast for F_7 . Let $F_1 = 350$.
7. Consider the table of observed values y_t and forecast values $F1_t$ and $F2_t$. The forecasts F1 were obtained by Method 1 and the forecasts F2 were obtained by Method 2. Calculate some measures of error and interpret them. Which method is best?

t	1	2	3
y_t	30	34	38
$F1_t$	28	33	35
$F2_t$	28	35	40



References



References

- Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D. & Martin, R. K. (1988). *An introduction to management science*. Eagan, MN: West Publishing Company.
- Brooks, D. (2008, October 28). Stimulus for skeptics. *The New York Times*. Retrieved from <http://www.nytimes.com>
- Derman, E. & Wilmott, P. (2009). The financial modeler's manifesto [website]. Retrieved from <http://www.wilmott.com/blogs/>
- Hillier, F. S. & Lieberman, G. J. (2005). *Introduction to operations research*. New York, NY: McGraw-Hill.
- Ragsdale, C. T. (2004). *Spreadsheet modelling and decision analysis*. Mason, OH: South-Western Thomson Learning.
- Soros, G. (2006). *The age of fallibility*. London, England: Weidenfeld & Nicolson.
- Waters, D. (1997). *Quantitative methods for business*. Boston, MA: Addison-Wesley.
- Wisniewski, M. (2006). *Quantitative methods for decision makers*. Upper Saddle River, NJ: Prentice Hall.
- Wood, D. (2008, September). Reassuringly expensive. *Risk Magazine*.

Further reading



Reading

Where else can I look?

These texts and websites can be consulted for additional information:

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D. & Martin, R. K. (1988). *An introduction to management science*. Eagan, MN: West Publishing Company.

Hillier, F. S. & Lieberman, G. J. (2005). *Introduction to operations research*. New York, NY: McGraw-Hill.

Mind Tools. (1996). MindTools [website]. Retrieved from <http://www.mindtools.com>

Ragsdale, C. T. (2004). *Spreadsheet modelling and decision analysis*. Mason, OH: South-Western Thomson Learning.

Soros, G. (2006). *The age of fallibility*. London, England: Weidenfeld & Nicolson.

Waters, D. (1997). *Quantitative methods for business*. Boston, MA: Addison-Wesley.

Wikipedia. (n.d.). Wikipedia: The free encyclopedia. Retrieved from <http://en.wikipedia.org>

Wisniewski, M. (2006). *Quantitative methods for decision makers*. Upper Saddle River, NJ: Prentice Hall.