International Foundation Programme

Foundation course: Mathematics and Statistics

James Ward and James Abdey

FP0001

2013
Contents

Introduction 1
Route map to the guide ........................................... 1
Time management .................................................. 1
Recommendations for working through the units ............... 2
Overview of learning resources .................................... 2
The subject guide and textbooks .................................. 2
Online study resources ............................................. 3
Virtual Learning Environment (VLE) .............................. 3
Making use of the Online library .................................. 4
Examination advice .................................................. 4
Calculators ............................................................ 5

Part 1 Mathematics 6

Introduction to Mathematics ........................................ 7
Syllabus .................................................................. 7
Aims of the course ................................................... 7
Learning outcomes for the course (Mathematics) ............... 8
Textbook ............................................................... 8

1 Review I — A review of some basic mathematics 9
1.1 Arithmetic ....................................................... 9
1.1.1 Basic arithmetic ........................................... 9
1.1.2 Fractions ................................................... 14
1.1.3 Powers ..................................................... 17
1.2 Algebra .......................................................... 22
1.2.1 Algebraic expressions ..................................... 22
1.2.2 Equations, formulae and inequalities .................... 26
Learning outcomes ................................................... 31
Exercises .............................................................. 31

i
2 Review II — Linear equations and straight lines

2.1 Linear equations
2.1.1 Linear equations in one variable
2.1.2 Linear equations in two variables
2.1.3 Visualising the solutions of linear equations in two variables

2.2 Straight lines
2.2.1 Drawing straight lines given their equations
2.2.2 The intercepts of a straight line
2.2.3 The gradient of a straight line
2.2.4 Finding the equation of a straight line
2.2.5 Applications of straight lines

2.3 Simultaneous equations
2.3.1 Visualising the solution to a pair of simultaneous equations
2.3.2 Solving simultaneous equations algebraically
2.3.3 An application of simultaneous equations in economics

Learning outcomes
Exercises

3 Review III — Quadratic equations and parabolae

3.1 Quadratic equations
3.1.1 Factorising
3.1.2 Completing the square
3.1.3 Using the completed square form to solve quadratic equations
3.1.4 Warning!
3.1.5 The quadratic formula

3.2 Parabolae
3.2.1 Sketching parabolae
3.2.2 Where do a parabola and a straight line intersect?
3.2.3 Where do two parabolae intersect?

Learning outcomes
Exercises

4 Functions

4.1 Functions
4.1.1 What is a function?
4.1.2 Some common functions
13 Data exploration III — Descriptive statistics: measures of location, dispersion and skewness

13.1 Summation notation ................................................. 199
13.2 Measures of location ................................................. 199
  13.2.1 Which ‘average’ should be used? ......................... 202
  13.2.2 Frequency tables .............................................. 203
13.3 Measures of dispersion ............................................ 205
  13.3.1 Variance and standard deviation ......................... 207
  13.3.2 Variance using frequency distributions ................. 209
13.4 Skewness ............................................................. 209
13.5 Summary ............................................................. 212
13.6 Key terms and concepts .......................................... 212
Learning outcomes ...................................................... 212
Exercises ................................................................. 213

14 Probability I — Introduction to probability theory

14.1 Probability theory ................................................. 215
  14.1.1 Assigning probabilities ...................................... 216
  14.1.2 The classical method ....................................... 216
  14.1.3 The relative frequency approach ....................... 217
  14.1.4 Subjective probabilities .................................. 217
14.2 Terminology ........................................................ 217
14.3 Sets ................................................................. 218
14.4 Independence ....................................................... 220
14.5 Complementary events .......................................... 221
14.6 Additive laws ....................................................... 222
14.7 Multiplicative laws ................................................ 224
14.8 Bayes’ theorem ..................................................... 225
  14.8.1 Version 1 ....................................................... 226
  14.8.2 Version 2 ....................................................... 226
  14.8.3 Version 3 ....................................................... 227
14.9 Summary — a listing of probability results ................. 228
15 Probability II — Probability distributions
15.1 Random variables ................................................. 236
15.2 Discrete random variables ................................. 237
15.3 Continuous random variables ......................... 238
15.4 Mathematical expectation ................................. 238
15.5 Functions of a random variable ......................... 239
15.6 Variance .......................................................... 241
15.7 Discrete uniform distribution ............................. 243
15.8 Bernoulli distribution ......................................... 244
15.9 Binomial distribution ......................................... 245
15.10 Poisson distribution .......................................... 249
15.11 A word on calculators ........................................ 252
15.12 Summary .......................................................... 252
15.13 Key terms and concepts ..................................... 252
Learning outcomes ..................................................... 252
Exercises ................................................................ 253

16 Probability III — The Normal distribution and sampling distributions 255
16.1 The Normal distribution ......................................... 255
  16.1.1 Probabilities for any Normal distribution ............... 261
  16.1.2 Some probabilities around the mean ..................... 262
16.2 Sampling distributions ......................................... 263
  16.2.1 Estimation ....................................................... 264
  16.2.2 Sampling from a Normal population .................... 268
16.3 Central Limit Theorem ............................................ 269
  16.3.1 CLT examples .................................................. 271
16.4 Summary ............................................................. 272
16.5 Key terms and concepts ........................................ 273
Learning outcomes ..................................................... 273
Exercises ................................................................ 273

17 Sampling and experimentation I — Sampling techniques and contact
## Contents

### methods

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.1 Sampling</td>
<td>275</td>
</tr>
<tr>
<td>17.1.1 Non-probability sampling techniques</td>
<td>278</td>
</tr>
<tr>
<td>17.1.2 Probability sampling techniques</td>
<td>281</td>
</tr>
<tr>
<td>17.1.3 Method of contact</td>
<td>284</td>
</tr>
<tr>
<td>17.2 Summary</td>
<td>285</td>
</tr>
<tr>
<td>17.3 Key terms and concepts</td>
<td>286</td>
</tr>
<tr>
<td>Learning outcomes</td>
<td>286</td>
</tr>
<tr>
<td>Exercises</td>
<td>286</td>
</tr>
</tbody>
</table>

### 18 Sampling and experimentation II — Bias and the design of experiments

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.1 Introduction</td>
<td>288</td>
</tr>
<tr>
<td>18.2 Types of error</td>
<td>289</td>
</tr>
<tr>
<td>18.3 Bias</td>
<td>290</td>
</tr>
<tr>
<td>18.4 Adjusting for non-response</td>
<td>291</td>
</tr>
<tr>
<td>18.5 Experimental design in the social and medical sciences</td>
<td>293</td>
</tr>
<tr>
<td>18.5.1 Experimental versus observational studies</td>
<td>293</td>
</tr>
<tr>
<td>18.5.2 Randomised controlled clinical trials</td>
<td>293</td>
</tr>
<tr>
<td>18.5.3 Randomised blocks</td>
<td>294</td>
</tr>
<tr>
<td>18.5.4 Multi-factorial experimental designs</td>
<td>294</td>
</tr>
<tr>
<td>18.5.5 Quasi-experiments</td>
<td>294</td>
</tr>
<tr>
<td>18.5.6 Cluster randomised trials</td>
<td>295</td>
</tr>
<tr>
<td>18.5.7 Analysis and interpretation</td>
<td>295</td>
</tr>
<tr>
<td>18.6 Summary</td>
<td>295</td>
</tr>
<tr>
<td>18.7 Key terms and concepts</td>
<td>296</td>
</tr>
<tr>
<td>Learning outcomes</td>
<td>296</td>
</tr>
<tr>
<td>Exercises</td>
<td>296</td>
</tr>
</tbody>
</table>

### 19 Fundamentals of regression I — Correlation and the simple linear regression model

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.1 Introduction</td>
<td>298</td>
</tr>
<tr>
<td>19.2 Correlation</td>
<td>301</td>
</tr>
<tr>
<td>19.3 Simple linear regression</td>
<td>304</td>
</tr>
<tr>
<td>19.4 Parameter estimation</td>
<td>305</td>
</tr>
<tr>
<td>19.5 Prediction</td>
<td>306</td>
</tr>
<tr>
<td>19.6 Summary</td>
<td>307</td>
</tr>
<tr>
<td>19.7 Key terms and concepts</td>
<td>307</td>
</tr>
</tbody>
</table>
Welcome to the world of Mathematics and Statistics! These are disciplines which are widely applied in areas such as finance, business, management, economics and other fields in the social sciences. The following units will provide you with the opportunity to grasp the fundamentals of these subjects and will equip you with some of the vital quantitative skills and powers of analysis which are highly sought-after by employers in many sectors.

As Mathematics and Statistics has so many applications, it should not be surprising that it forms the compulsory component of the International Foundation Programme. The analytical skills which you will develop on this course will therefore serve you well in both your future studies and beyond in the real world of work. The material in this course is necessary as preparation for other courses you may study later on as part of a degree programme or diploma; indeed, in many cases a course in Mathematics or Statistics is a compulsory component on the University of London International Programmes’ degrees.

Route map to the guide

This subject guide provides you with a framework for covering the syllabus of the Mathematics and Statistics course in the International Foundation Programme and directs you to additional resources such as readings and the virtual learning environment (VLE).

The following 20 units will introduce you to these disciplines and equip you with the necessary quantitative skills to assist you in further programmes of study. Given the cumulative nature of Mathematics and Statistics, the units are not a series of self-contained topics, rather they build on each other sequentially. As such, you are strongly advised to follow the subject guide in unit order. There is little point in rushing past material which you have only partially understood in order to reach the final unit.

Once you have completed your work on all of the units, you will be ready for examination revision. A good place to start is the sample examination paper which you will find at the end of the subject guide.

Time management

About one-third of your private study time should be spent reading and the other two-thirds doing problems. (Note the emphasis on practising problems!)

To help your time management, each unit of this course should take a week to study and so you should be spending 10 weeks on Mathematics and 10 weeks on Statistics.
Introduction

**Recommendations for working through the units**

The following procedure is recommended for each unit.

i. Read the overview and the aims of the unit.

ii. Now work through each section of the unit making sure you can understand the examples given and try the activities as you encounter them. In parallel, watch the accompanying video tutorials for each section on the VLE.

iii. At the end of the unit, review the intended learning outcomes carefully, almost as a checklist. Do you think you have achieved these targets?

iv. Attempt the unit’s self-test quizzes on the VLE. You can treat these as additional activities. This time, though, you will have to think a little about which part of the new material you have learnt is appropriate to each question.

v. Attempt the exercises given at the end of the unit. The solutions can be found on the VLE, but you should only look at these after attempting them yourself!

vi. If you have problems at this point, go back to the subject guide and work through the area you find difficult again. Don’t worry — you will improve your understanding to the point where you can work confidently through the problems.

The last few steps are most important. It is easy to think that you have understood the text after reading it, but **working through problems is the crucial test of understanding**. Problem-solving should take most of your study time (refer to the ‘Time management’ section above). Note that we have given worked examples and activities to cover each substantive topic in the subject guide. The essential reading examples are added for further consolidation of the whole unit and also to help you work out exactly what the questions are about! One of the problems students sometimes have in an examination is that they waste time trying to understand which part of the syllabus particular questions relate to. These final questions, together with the further explanations on the VLE, aim to help with this before you tackle the sample examination questions at the end of each unit.

Try to be disciplined about this: don’t look up the answers until you have done your best. Some of the ideas you encounter may seem unfamiliar at first, but your attempts at the questions, however dissatisfied you feel with them, will help you understand the material far better than reading and re-reading the prepared answers — honest!

So to conclude, perseverance with problem-solving is your passport to a strong examination performance. Attempting (ideally successfully!) all the cited exercises is of paramount importance.

**Overview of learning resources**

**The subject guide and textbooks**

This subject guide for Mathematics and Statistics has been structured so that it is tailored to the specific requirements of the examinable material. It is ‘written to the
course’, unlike textbooks which may cover additional material which will not be examinable or may not cover some material that is! Therefore the subject guide should act as your principal resource.

However, a textbook may give an alternative explanation of a topic (which is useful if you have difficulty following something in the subject guide) and so you may want to consult one for further clarification. Additionally, a textbook will contain further examples and exercises which can be used to check and consolidate your understanding. For this course, a useful starting point is


as this will serve as useful background reading. But, many books are available covering the material frequently found in mathematics and statistics courses like this one and so, if you need a textbook for background reading, you should find one that is appropriate to your level and tastes.

Online study resources

In addition to the subject guide and the Essential reading, it is crucial that you take advantage of the study resources that are available online for this course, including the VLE and the Online Library.

You can access the VLE, the Online Library and your University of London email account via the Student Portal at http://my.londoninternational.ac.uk

You should have received your login details for the Student Portal with your official offer, which was emailed to the address that you gave on your application form. You have probably already logged in to the Student Portal in order to register. As soon as you registered, you will automatically have been granted access to the VLE, Online Library and your fully functional University of London email account.

If you have forgotten these login details, please click on the ‘Forgotten your password’ link on the login page.

Virtual Learning Environment (VLE)

The VLE, which complements this subject guide, has been designed to enhance your learning experience, providing additional support and a sense of community. In addition to making printed materials more accessible, the VLE provides an open space for you to discuss interests and to seek support from other students, working collaboratively to solve problems and discuss subject material. In a few cases, such discussions are driven and moderated by an academic who offers a form of feedback on all discussions. In other cases, video material, such as audio-visual tutorials, are available. These will typically focus on taking you through difficult concepts in the subject guide. For quantitative courses, such as Mathematics and Statistics, fully worked-through solutions of practice examination questions are available. For some qualitative courses, academic interviews and debates will provide you with advice on approaching the subject and examination questions, and will show you how to build an argument effectively.

Past examination papers and Examiners’ commentaries will be available to download in
due course (the first examination for this course will be sat in 2014) and these provide advice on how each examination question might best be answered. Self-testing activities allow you to test your knowledge and recall of the academic content of various courses. Finally, a section of the VLE has been dedicated to providing you with expert advice on practical study skills such as preparing for examinations and developing digital literacy skills.

Making use of the Online library

The Online library contains a huge array of journal articles and other resources to help you read widely and extensively.

Essential reading journal articles listed on a number of reading lists are available to download from the Online library.

The easiest way to locate relevant content and journal articles in the Online library is to use the Summon search engine.

If you are having trouble finding an article listed on the reading list, try:

1. removing any punctuation from the title, such as single quotation marks, question marks and colons, and/or
2. putting quotation marks around the title, for example “Why the banking system should be regulated”.

To access the majority of resources via the Online library you will either need to use your University of London Student Portal login details, or you will be required to register and use an Athens login: http://tinyurl.com/ollathens

Examination advice

Important: the information and advice given in the following section are based on the examination structure used at the time this subject guide was written. Please note that subject guides may be used for several years. Because of this, we strongly advise you to check both the current Regulations for relevant information about the examination, and the current Examiners’ commentaries, where you should be advised of any forthcoming changes. You should also carefully check the rubric/instructions on the paper you actually sit and follow those instructions.

The examination is by a two-hour, unseen, written paper. No books may be taken into the examination, but you will be provided with extracts of statistical tables (as reproduced in this subject guide). A calculator may be used when answering questions on this paper, see below, and it must comply in all respects with the specification given in the General Regulations.

The examination comprises two sections, each containing three compulsory questions. Section A covers the mathematics part of the course counting for 50% of the marks, and Section B covers the statistics part of the course for the remaining 50% of the marks. You are required to pass both Sections A and B to pass the examination.
In each section, the first question contains four short questions worth 5 marks each, followed by two longer questions worth 15 marks each. Since the examination will seek to assess a broad cross-section of the syllabus, we strongly advise you to study the whole syllabus. A sample examination paper is provided at the end of this subject guide along with a commentary providing extensive advice on how to answer each question.

Remember, it is important to check the VLE for:

- Up-to-date information on examination and assessment arrangements for this course.
- Where available, past examination papers and Examiners’ commentaries for the course which give advice on how each question might best be answered.

**Calculators**

You will need to provide yourself with a basic calculator. It should not be programmable, because such machines are not allowed in the examination by the University. The most important thing is that you should accustom yourself to using your chosen calculator and feel comfortable with it. Your calculator must comply in all respects with the specification given in the General Regulations.
Part 1
Mathematics
Introduction to Mathematics

Syllabus

This half of the course introduces some of the basic ideas and methods of Mathematics with an emphasis on their application. The Mathematics part of this course has the following syllabus.

- **Arithmetic and algebra:** A review of arithmetic (including the use of fractions and decimals) and the manipulation of algebraic expressions (including the use of brackets and the power laws). Solving linear equations and the relationship between linear expressions and straight lines (including the solution of simultaneous linear equations). Solving quadratic equations and the relationship between quadratic expressions and parabolae.

- **Functions:** An introduction to functions. Some common functions (including polynomials, exponentials, logarithms and trigonometric functions). The existence of inverse functions and how to find them. The laws of logarithms and their uses.

- **Calculus:** The meaning of the derivative and how to find it (including the product, quotient and chain rules). Using derivatives to find approximations and solve simple optimisation problems with economic applications. Curve sketching. Integration of simple functions and using integrals to find areas.

- **Financial mathematics:** Compound interest over different compounding intervals. Arithmetic and geometric sequences. The sum of arithmetic and geometric series. Investment schemes and some ways of assessing the value of an investment.

Aims of the course

The aims of the Mathematics part of this course are to provide:

- a grounding in arithmetic and algebra;

- an overview of functions and the fundamentals of calculus;

- an introduction to financial mathematics.

Throughout, the treatment is at an elementary mathematical level but, as you progress through this part of the course, you should develop some quite sophisticated mathematical skills.
Learning outcomes for the course (Mathematics)

At the end of the Mathematics part of the course, you should be able to:

- manipulate algebraic expressions;
- graph, differentiate and integrate simple functions;
- calculate basic quantities in financial mathematics.

Textbook

As previously mentioned in the main introduction, this subject guide has been designed to act as your principal resource. The textbook


may be useful as ‘background reading’ but it is not essential. However, you might benefit from reading parts of it if you find any of the material difficult to follow at first.
Unit 1: Review I
A review of some basic mathematics

Overview

In this unit we revise some material on arithmetic and algebra which you should have encountered before. Starting with arithmetic, this will involve revising the basic mathematical operations and how they can be combined with and without the use of brackets, how we can manipulate fractions and the use of powers. We then look at some basic algebra and see how to use and manipulate algebraic expressions.

Aims

The aims of this unit are as follows.

- To revise the basics of arithmetic, including the use of fractions and powers.
- To revise the most basic ideas behind algebra.

Specific learning outcomes can be found near the end of this unit.

1.1 Arithmetic

In this section we revise some material which could be called ‘arithmetic’. The idea behind this revision is to refresh our memories about how things like brackets, fractions and powers work so that our revision of ‘algebra’ in the next section will, hopefully, be easier.

1.1.1 Basic arithmetic

In mathematics we use four basic mathematical operations:

- **addition** denoted by ‘+’ gives us ‘sums’, e.g. $6 + 3 = 9$;

- **subtraction** denoted by ‘−’ gives us ‘differences’, e.g. $6 - 3 = 3$;

- **multiplication** denoted by ‘×’ or ‘·’ gives us ‘products’, e.g. $6 \times 3 = 18$ or $6 \cdot 3 = 18$;

- **division** denoted by ‘÷’ or a ‘horizontal line’ gives us ‘quotients’, e.g. $6 \div 2 = 3$ or $\frac{6}{2} = 3$.

In particular, notice that there are two common notations for multiplication and division. For multiplication, the reason for this is that a handwritten ‘×’ can be
1. Review I — A review of some basic mathematics

confused with a handwritten ‘x’ whereas, for division, the reason is that writing expressions that involve division (i.e. ‘÷’) as fractions enables us to manipulate them more easily using the laws of fractions.

**Combinations of operations**

Often, different mathematical operations will occur in the same expression. For example, we might be asked to work out the values of the expressions

1. \(22 - 7 + 12 - 26 + 1,\)
2. \(125 ÷ 25 \times 2 \times 3 ÷ 15,\)
3. \(22 - 20 \times 3 ÷ 4 - 5.\)

In such cases, we have the following rules.

1. If only addition and subtraction are involved: We work from left to right to get
   \[
   22 - 7 + 12 - 26 + 1 = 15 + 12 - 26 + 1 = 27 - 26 + 1 = 1 + 1 = 2.
   \]
2. If only multiplication and division are involved: We work from left to right to get
   \[
   125 ÷ 25 \times 2 \times 3 ÷ 15 = 5 \times 2 \times 3 ÷ 15 = 10 \times 3 ÷ 15 = 30 ÷ 15 = 2.
   \]
3. When addition/subtraction and multiplication/division are involved: We work out the multiplications and divisions first (working left to right as necessary) and then we do the additions and subtractions (working left to right as necessary) to get
   \[
   22 - 20 \times 3 ÷ 4 - 5 = 22 - 60 ÷ 4 - 5 = 22 - 15 - 5 = 7 - 5 = 2.
   \]

**Brackets I: Evaluating expressions that involve brackets**

If an expression involves brackets, then the operations within the brackets must be performed first. As such, brackets can be used to change the order in which operations are performed. For example, we might be asked to work out the values of the expressions

1. \(9 - (4 + 3)\) as opposed to \(9 - 4 + 3,\)
2. \(6 ÷ (2 \times 3)\) as opposed to \(6 ÷ 2 \times 3,\)
3. \((12 \times 3 - 8) \times 2\) as opposed to \(12 \times 3 - 8 \times 2.\)

In such cases, we work out the expression in brackets first, i.e. we get

1. working out the expression in brackets first we get
   \[
   9 - (4 + 3) = 9 - 7 = 2,
   \]
as opposed to
   \[
   9 - 4 + 3 = 5 + 3 = 8,
   \]
where we work from left to right.
2. working out the expression in brackets first we get
\[ 6 \div (2 \times 3) = 6 \div 6 = 1, \]
as opposed to
\[ 6 \div 2 \times 3 = 3 \times 3 = 9, \]
where we work from left to right.

3. working out the expression in brackets first, proceeding to the rules above as necessary, we have
\[ (12 \times 3 - 8) \times 2 = (36 - 8) \times 2 = 28 \times 2 = 56, \]
as opposed to
\[ 12 \times 3 - 8 \times 2 = 36 - 16 = 20, \]
where we multiply first and then subtract.

What if we have two or more sets of brackets? Well, if they are not ‘nested’, for example if we have
\[ (12 \times 3 - 8) \times (24 - 14), \]
then we need to work out what is in each of the brackets first, proceeding according to the rules above, i.e.
\[ (12 \times 3 - 8) \times (24 - 14) = (36 - 8) \times 10 = 28 \times 10 = 280. \]

And, if the brackets are ‘nested’, for example
\[ 6 + (9 - (4 + 3)), \]
then we start with the innermost set of brackets and work ‘outwards’, i.e.
\[ 6 + (9 - (4 + 3)) = 6 + (9 - 7) = 6 + 2 = 8. \]

These rules allow you to work out the values of simple mathematical expressions using brackets. In a moment we shall see another way of dealing with brackets which will be more useful to us in this course.

**Negative numbers**

Consider the following three expressions and their values.

1. \( 6 - 3 = +3, \)
2. \( 6 - 6 = 0, \)
3. \( 6 - 9 = -3. \)

In this case, we can see that subtracting larger and larger numbers from six, gives us a positive answer, zero and a negative answer respectively. For simplicity, we usually omit the ‘+’ sign and write ‘+3’, say, as 3.

When we have expressions involving negative numbers, we have the following handy rules.
1. **Adding a negative number:** This has the same effect as subtracting the corresponding positive number, e.g.

\[ 5 + (-3) = 5 - (+3) = 5 - 3 = 2, \]

and

\[ -5 + (-3) = -5 - (+3) = -5 - 3 = -8. \]

2. **Subtracting a negative number:** This has the same effect as adding the corresponding positive number, e.g.

\[ 5 - (-3) = 5 + (+3) = 5 + 3 = 8. \]

and

\[ -5 - (-3) = -5 + (+3) = -5 + 3 = -2. \]

3. **Multiplying a positive number by a negative number:** This gives us a negative number, e.g.

\[ (+5) \times (-3) = -(5 \times 3) = -15. \]

and

\[ (-5) \times (+3) = -(5 \times 3) = -15. \]

This is normally remembered as ‘positive times negative is negative’.

4. **Multiplying a negative number by a negative number:** This gives us a positive number, e.g.

\[ (-5) \times (-3) = +(5 \times 3) = +15 = 15. \]

This is normally remembered as ‘negative times negative is positive’.

5. **Dividing a positive number by a negative number (or vice versa):** This gives us a negative number, e.g.

\[ (+6) \div (-3) = -(6 \div 3) = -2. \]

and

\[ (-6) \div (+3) = -(6 \div 3) = -2. \]

This is normally remembered as ‘positive divided by negative is negative’ (or vice versa).

6. **Dividing a negative number by a negative number:** This gives us a positive number, e.g.

\[ (-6) \div (-3) = +(6 \div 3) = +2 = 2. \]

This is normally remembered as ‘negative divided by negative is positive’.

Indeed, notice the similarity between (3) and (5) which can be remembered as ‘multiplying (or dividing) a positive and a negative yields a negative’ and (4) and (6) which can be remembered as ‘multiplying (or dividing) a negative and a negative yields a positive’.
Brackets II: Removing brackets from expressions

A more useful way of thinking about brackets involves being able to ‘remove’ the brackets from an expression. For example, consider the expression

\[3 + 2 \times (9 - 4)\].

Using the rules above we could work this out by thinking of it as

\[3 + 2 \times (9 - 4) = 3 + 2 \times 5 = 3 + 10 = 13\].

Alternatively, we can ‘remove’ the brackets by thinking of the ‘2’ in ‘2 \times (9 - 4)’ as multiplying everything in the bracket, i.e.

\[2 \times (9 - 4) = (2 \times 9) - (2 \times 4)\].

Using this method we get

\[3 + 2 \times (9 - 4) = 3 + ((2 \times 9) - (2 \times 4)) = 3 + (18 - 8) = 3 + 10 = 13\],

which is the same answer as before.

| Activity 1.1 | Show that if we worked out \(3 + (9 - 4) \times 2\), we would also get 13. |

What if we had to work out \(3 - (9 - 4)\)? We adopt the convention that a minus sign in front of a bracket is the same as adding something that has been multiplied by \(-1\). Using this, and what we saw above, gives us

\[3 - (9 - 4) = 3 + (-1) \times (9 - 4) = 3 + ((-1 \times 9) - (-1 \times 4)) = 3 + (-9 + 4) = 3 + (-5) = -2\].

Of course, this is what we should expect as \(3 - (9 - 4) = 3 - 5 = -2\).

Absolute values

The magnitude (or absolute value) of a number is found by ignoring the minus sign (if there is one). For example, the magnitude of 6, written \(|6|\), is 6 and the magnitude of \(-6\), written \(|-6|\), is also 6, i.e. we have

\[|6| = 6 \text{ and } |-6| = 6\].

In a way, the magnitude operation acts like a bracket as we need to evaluate the magnitude of the number inside it before we use it in calculations, e.g.

- \(4 - |2 - 3| = 4 - 1 = 3\) as \(|2 - 3| = |-1| = 1\), and
- \(|4 - 2| - 3 = 2 - 3 = -1\) as \(|4 - 2| = |2| = 2\).
Inequalities

We use the symbols ‘<’ and ‘>’ to show that one number is ‘less than’ or ‘greater than’ another number respectively. So, for example, $2 < 3$ and $5 > 1$. Zero is less than any positive number and greater than any negative number, e.g. $0 < 5$ and $0 > -5$. As such, any negative number is less than any positive number, e.g. $-3 < 2$. Negative numbers are larger when they have smaller magnitudes (i.e. when they are closer to zero), e.g. $-3 < -2$ and $-1 > -5$. As such, we can say that smaller negative numbers (like $-100$ compared to $-1$) have larger absolute values (like $100$ compared to $1$).

1.1.2 Fractions

A fraction such as $\frac{3}{2}$ is, using our ‘horizontal line’ notation for division, the same as dividing the number above the line (i.e. 3) by the number below the line (i.e. 2). We call the number above the line the numerator and the number below the line the denominator. If we have two fractions, say

$$\frac{3}{5} \text{ and } \frac{4}{2},$$

the number we get by multiplying their denominators together is called the common denominator of these fractions, and this will be $5 \times 2 = 10$ in this case.

Manipulating fractions

Sometimes we want to manipulate fractions in order to simplify them or to put them in a form where their denominator is the common denominator. The two basic procedures we use to do these two manipulations are as follows.

- To simplify a fraction we want to write it in lowest terms, e.g. $\frac{6}{10}$ can be written as

$$\frac{6}{10} = \frac{2 \times 3}{2 \times 5} = \frac{3}{5},$$

by dividing through on top and bottom by the common factor of 2.

- Conversely, to write a fraction so that its denominator is a common denominator, e.g. to write $\frac{3}{5}$ so that its denominator is, as above, the common denominator of 10 we note that it can be written as

$$\frac{3}{5} = \frac{2 \times 3}{2 \times 5} = \frac{6}{10},$$

by multiplying top and bottom by 2.

This second technique is especially useful when we add and subtract fractions as we shall now see.

*That is, so that the numerator and denominator have no common divisors.
Adding and subtracting fractions

To add or subtract fractions, we first put them over a common denominator, e.g.

\[
\frac{4}{5} + \frac{2}{3} = \frac{4 \times 3}{5 \times 3} + \frac{2 \times 5}{3 \times 5} = \frac{12}{15} + \frac{10}{15} = \frac{12 + 10}{15} = \frac{22}{15},
\]

and

\[
\frac{4}{5} - \frac{2}{3} = \frac{4 \times 3}{5 \times 3} - \frac{2 \times 5}{3 \times 5} = \frac{12}{15} - \frac{10}{15} = \frac{12 - 10}{15} = \frac{2}{15}.
\]

Multiplying fractions

To multiply fractions, we just multiply the numerators and denominators together, e.g.

\[
\frac{4}{5} \times \frac{2}{3} = \frac{4 \times 2}{5 \times 3} = \frac{8}{15}.
\]

Reciprocals

The reciprocal of a fraction is what we get when we swap the numerator and denominator around, e.g. the reciprocal of \(\frac{3}{5}\) is \(\frac{5}{3}\). The reciprocal is useful when we come to divide fractions as we shall now see.

Dividing fractions

To divide fractions, we multiply the first fraction by the reciprocal of the second, e.g. if we want to evaluate

\[
\frac{4}{5} \div \frac{2}{3},
\]

the rule tells us that this is the same as multiplying \(\frac{4}{5}\) by the reciprocal of \(\frac{2}{3}\), which is \(\frac{3}{2}\), and so we have

\[
\frac{4}{5} \div \frac{2}{3} = \frac{4 \times 3}{5 \times 2} = \frac{12}{10}.
\]

This can now be worked out using the multiplication rule, i.e.

\[
\frac{4}{5} \div \frac{2}{3} = \frac{4 \times 3}{5 \times 2} = \frac{12}{10}.
\]

Of course, we can simplify this by noting that the numerator and denominator have a common factor of 2, i.e. the answer is \(\frac{6}{5}\) in lowest terms.

It is, perhaps, also interesting to note that the reciprocal of a fraction is just one divided by that fraction, e.g. as

\[
1 \div \frac{3}{2} = 1 \times \frac{2}{3} = \frac{2}{3},
\]

we can see that the reciprocal of \(\frac{3}{2}\), i.e. \(\frac{2}{3}\), is just one divided by \(\frac{3}{2}\).
Improper and proper fractions

An improper fraction is one where the numerator is greater in magnitude than the denominator and a proper fraction is one where the numerator is less in magnitude than the denominator, e.g. \( \frac{22}{5} \) is an improper fraction and \( \frac{4}{5} \) is a proper fraction.

Sometimes it is convenient to be able to write improper fractions as proper fractions, e.g. we can write
\[
\frac{22}{5} = \frac{20 + 2}{5} = \frac{20}{5} + \frac{2}{5} = 4 + \frac{2}{5},
\]
as 5 goes into 20 four times. This can be written as \( 4\frac{2}{5} \) and we read it as ‘four and two fifths’ to indicate that \( \frac{22}{5} \) is the same as four ‘wholes’ and two fifths of a ‘whole’.

However, in this course, we will usually not use this way of writing fractions as, using our convention of writing \( 4 \times \frac{2}{5} \) as \( 4 \cdot \frac{2}{5} \), we can easily get confused between ‘four and two fifths’ and ‘four times two fifths’. As such, when the need arises, we will normally stick to improper fractions.

Decimals

Often, you will see fractions written as decimals and vice versa, e.g. the fraction \( \frac{1}{4} \) is exactly the same as the decimal 0.25. But, be aware that some fractions do not have a nice finite decimal expansion, e.g.
\[
\frac{1}{3} \text{ is the decimal } 0.33333\ldots,
\]
i.e. there is an infinite number of threes after the decimal point. The problem with this is that, in such cases, using decimals instead of fractions can lead to rounding errors, e.g.
\[
3 \times \frac{1}{3} = 1,
\]
exactly. But, just keeping the first four threes of the decimal expansion for \( \frac{1}{3} \), i.e. rounding \( \frac{1}{3} \) to four decimal places, written 4dp, we have 0.3333 and this gives us
\[
3 \times \frac{1}{3} \simeq 3 \times 0.3333 = 0.9999,
\]
where ‘\( \simeq \)’ means ‘approximately equal to’. That is, using the decimal rounded to four decimal places gives us an answer which is not exactly one, i.e. there is a rounding error in our calculation, and this is why we generally use fractions instead of decimals.

Percentages

The percentage sign, i.e. ‘\%’, means ‘divide by 100’, e.g. 20\% is the same as \( \frac{20}{100} \) as a fraction, or 0.2 as a decimal. As such, 20\% of 150 is
\[
150 \times \frac{20}{100} = \frac{3,000}{100} = 30.
\]
Knowing this, we can see what it means to increase 150 by 20\% or decrease 150 by 20\%, i.e.
to increase 150 by 20%, we get

\[ 150 + 30 = 180, \]

as 30 is 20% of 150. Notice that an increase by 20% can also be seen as 120% of the original, i.e.

\[ 150 \times \frac{120}{100} = \frac{18,000}{100} = 180, \]

as before.

to decrease 150 by 20%, we get

\[ 150 - 30 = 120, \]

as 30 is 20% of 150. Notice that a decrease by 20% can also be seen as 80% of the original, i.e.

\[ 150 \times \frac{80}{100} = \frac{12,000}{100} = 120, \]

as before.

These ideas will be particularly useful when we come to consider compound interest in Unit 9.

1.1.3 Powers

Another operation that you will have come across before is the idea of ‘raising a number to a certain power’. The number which represents the power can also be called the exponent and the number which is being raised to that power is called the base. For example, we could have \(4^2\), \(4^{-2}\) or \(4^{\frac{1}{2}}\) and, in each case, ‘4’ is the base and the other number, i.e. ‘2’, ‘–2’ or ‘\( \frac{1}{2} \)’ respectively, is the exponent or power. We often refer to expressions of this form as ‘powers’.

Positive integer powers

The simplest powers to work out are those where the power is a positive integer such as 1, 2, 3, … . In such cases, the power just means ‘multiply the base by itself that many times’, e.g.

\[ 4^1 = 4, \quad 4^2 = 4 \times 4 = 16, \quad 4^3 = 4 \times 4 \times 4 = 64, \ldots . \]

One application of this is standard index form (or scientific notation) where we are able to write large numbers in terms of powers of 10, e.g. we can write three million as

\[ 3,000,000 = 3 \times 1,000,000 = 3 \times 10^6, \]

as 1,000,000 is the same as \(10^6\).

Powers and other operations

In terms of combinations of operations, evaluating the effect of a power comes before multiplying and dividing, e.g. we can see that

\[ 2 \times (4^2) + 3 = 2 \times 16 + 3 = 32 + 3 = 35. \]
1. Review I — A review of some basic mathematics

Of course, as before, we can also use brackets to change the order in which we do the operations, e.g.

\[(2 \times 4)^2 + 3 = \overline{8^2 + 3} = 64 + 3 = 67,\]

and

\[2 \times (4^2 + 3) = 2 \times (16 + 3) = 2 \times 19 = 38.\]

In particular, when writing out expressions involving brackets, take care to distinguish between, e.g. \(2^3 + 5\) and \(2^{3+5}\), as the former is 13 whilst the latter is 256!

Also, similar to what we saw earlier, it is possible to remove the brackets from expressions involving powers by applying the power to all of the numbers in the bracket. For example,

- \((2 \times 3)^4 = 2^4 \times 3^4 = 16 \times 81 = 1,296.\)

- \(\left(\frac{2}{3}\right)^4 = \frac{2^4}{3^4} = \frac{16}{81}.\)

The power laws

If we have the same base, then the power laws can allow us to simplify expressions that involve multiplying powers, dividing powers and raising to powers. These laws are as follows.

- **Multiplying powers:** If we multiply two powers, we add the powers. For example, if we have \(2^4 \times 2^3\), we can write,

  \[2^4 \times 2^3 = 2^{4+3},\]

  as \(2^4 \times 2^3 = 16 \times 8 = 128\) and \(2^{4+3} = 2^7 = 128.\)

- **Dividing powers:** If we divide two powers, we subtract the power in the denominator from the power in the numerator. For example, if we have \(2^4/2^3\), we can write,

  \[\frac{2^4}{2^3} = 2^{4-3},\]

  as \(\frac{2^4}{2^3} = \frac{16}{8} = 2\) and \(2^{4-3} = 2^1 = 2.\)

- **Raising to powers:** If we raise a power to another power, we multiply the powers. For example, if we have \((2^4)^3\), we can write,

  \[(2^4)^3 = 2^{4\times3},\]

  as \((2^4)^3 = 16^3 = 4,096\) and \(2^{4\times3} = 2^{12} = 4,096.\)

Notice that, if the bases of the powers are not the same, then we can not use the power laws. For example, to calculate

- \(3^4 \times 2^5\) we could use \(3^4 \times 2^5 = 81 \times 32 = 2,592\), but we could not use the power law.

- \(\frac{3^4}{2^5}\) we could use \(\frac{3^4}{2^5} = \frac{81}{32}\), but we could not use the power law.
Negative integer powers

Negative integer powers, such as \(-1, -2, -3, \ldots\), mean ‘take the reciprocal of the base raised to the corresponding positive power’. For example,

\[
4^{-1} = \frac{1}{4^1} = \frac{1}{4}, \quad 4^{-2} = \frac{1}{4^2} = \frac{1}{16}, \quad 4^{-3} = \frac{1}{4^3} = \frac{1}{64}, \ldots
\]

In particular, note that a power of \(-1\) is the same as the reciprocal, e.g. \(4^{-1} = \frac{1}{4}\) which is the reciprocal of 4. Similarly, this means that

\[
\left(\frac{3}{5}\right)^{-1} = \frac{5}{3},
\]

which is the reciprocal of \(\frac{3}{5}\).

Zero powers

We now observe that any number raised to the power zero is one. For example, as

\[
4^1 \times 4^{-1} = 4^{1-1} = 4^0,
\]

by the power law, and

\[
4^1 \times 4^{-1} = 4 \times \frac{1}{4} = 1,
\]

we can see that \(4^0 = 1\).

Fractional powers I: Square roots

A square root of a number, say 64, is a number which, when multiplied by itself, gives us 64. So, as

\[
8 \times 8 = 64,
\]

we can see that 8 is a square root of 64. Indeed, since a negative number times a negative number is positive, we can see that

\[
(-8) \times (-8) = 64,
\]

and so \(-8\) is also a square root of 64. Thus, we can see that the square roots of 64 are 8 and \(-8\). We often express this by saying ‘the square roots of 64 are ±8’ where the ‘±’ is read ‘plus or minus’. Thus, we can see, by repeating this argument, that every positive number has two square roots, one positive and one negative, and both of the same magnitude.

What about other numbers? Well, since \(0 \times 0 = 0\), we can see that the square root of zero is zero and, moreover, zero is the only square root of zero. And, if we consider negative numbers, say \(-64\), we can see that there are no square roots since there is no way of multiplying a number by itself to get \(-64\).

We often denote the positive square root of a number, say 64, by ‘\(\sqrt{64}\)’ and so, from the above we can see that \(\sqrt{64} = 8\) and \(\sqrt{0} = 0\). Of course, as negative numbers have no square roots, something like \(\sqrt{-64}\) does not exist.
Going back to our earlier example, as the square root of 64 is a number which, when multiplied by itself, gives us 64 we can see that

\[
(\sqrt{64})^2 = \sqrt{64} \times \sqrt{64} = 64,
\]

and this is why the square root is so called: *squaring* the square root gives us the original number. Now, if we think of raising the number 64 to the power \( \frac{1}{2} \), we can see that

\[
(64^{\frac{1}{2}})^2 = 64^{\frac{1}{2} \times 2} = 64^1 = 64,
\]

using the power laws. And, comparing these two expressions, it is natural to think of \( 64^{\frac{1}{2}} \) as exactly the same thing as \( \sqrt{64} \), i.e.

\[\sqrt{64} = 64,\]

and so we identify square roots with powers of \( \frac{1}{2} \).

**Activity 1.2** Find the square roots of 4, 9, 16, 25, 36 and 49.

**Fractional powers II: \( n \)th roots**

More generally, if \( n \) is a positive integer greater than 2, we say that an \( n \)th root of a number, say 64, is a number which gives us 64 when raised to the power \( n \). We often denote the \( n \)th root of a number, say 64, by \( n\sqrt{64} \). For example,

- the *cube root* of 64, denoted by \( 3\sqrt{64} \), is 4 as four cubed is 64, i.e.
  
  \[4^3 = 64 \text{ and so } 3\sqrt{64} = 4.\]

  Notice that 64 has no negative cube root since \((-4)^3 = -64\) and not 64, as such 64 only has one cube root, i.e. 4. Repeating this argument, we can see that all positive numbers only have one cube root.

  In terms of powers, as \((3\sqrt{64})^3 = 4^3 = 64\) and

  \[
  (64^{\frac{1}{3}})^3 = 64^{\frac{1}{3} \times 3} = 64^1 = 64,
  \]

  comparing these two expressions it is natural to think of \( 64^{\frac{1}{3}} \) as exactly the same thing as \( 3\sqrt{64} \), i.e.

  \[64^{\frac{1}{3}} = 3\sqrt{64},\]

  and so we identify cube roots with powers of \( \frac{1}{3} \).

- the *sixth root* of 64, denoted by \( 6\sqrt{64} \), is 2 as two to the power six is 64, i.e.
  
  \[2^6 = 64 \text{ and so } 6\sqrt{64} = 2.\]

  Notice that 64 also has a negative sixth root since \((-2)^6 = 64\) and so 64 has two sixth roots, i.e. \( \pm 2 \). Repeating this argument, we can see that all positive numbers will have two sixth roots.
In terms of powers, as \((\sqrt[6]{64})^6 = 2^6 = 64\) and 
\[ (64^{\frac{1}{6}})^6 = 64^{\frac{1}{6} \times 6} = 64^1 = 64, \]
comparing these two expressions it is natural to think of \(64^{\frac{1}{6}}\) as exactly the same thing as \(\sqrt[6]{64}\), i.e.
\[ 64^{\frac{1}{6}} = \sqrt[6]{64}, \]
and so we identify sixth roots with powers of \(\frac{1}{6}\).
And, more generally, we can write the positive \(n\)th root of a number \(a\), or \(\sqrt[n]{a}\), as \(a\) to the power of \(\frac{1}{n}\), i.e. \(a^{\frac{1}{n}}\).

**Activity 1.3** Find the cube root of 27 and the fourth roots of 81.

**Fractional powers III: powers of \(n\)th roots**

Other fractional powers can be evaluated using the rules above, e.g. to evaluate \(8^{\frac{2}{3}}\) we can think of it as
\[ 8^{\frac{2}{3}} = 8^{2 \times \frac{1}{3}} = (8^2)^\frac{1}{3} = 64^{\frac{1}{3}} = 4, \]
or as
\[ 8^{\frac{2}{3}} = 8^{\frac{1}{3} \times 2} = (8^{\frac{1}{3}})^2 = 2^2 = 4, \]
using the power laws. Other examples involving fractional roots would be

- \((3^{\frac{1}{2}})^4 = 3^{\frac{1}{2} \times 4} = 3^2 = 9\), and
- \(\frac{4^2}{4^{\frac{1}{2}}} = 4^{\frac{2}{2} - \frac{1}{2}} = 4^{\frac{4-1}{2}} = 4^{\frac{3}{2}} = 4^{\frac{1}{2}} = 2\),

using the power laws.

**Fractional powers IV: Warnings**

When using the above ideas you should also bear the following in mind.

- When using the square root and \(n\)th root sign, i.e. \(\sqrt{}\) and \(\sqrt[n]{\cdot}\), always be clear about what parts of the expression are included in the root. For example,

  \[ \sqrt{4 \times 16} \quad \text{and} \quad \sqrt[6]{4 \times 16}, \]

  are different expressions (the former is equal to 8 whilst the latter is equal to 32). Generally speaking, you can make your expressions clear by extending the ‘tail’ of the root sign or using brackets.

- Be careful when working with powers of negative numbers since even roots of negative numbers do not exist. For example,

  \[ ((-2)^2)^\frac{1}{2} = 4^{\frac{1}{2}} = 2, \]

  is fine, but \((-2)^{\frac{1}{2}}\) does not exist and, as such, nor does \(((-2)^{\frac{1}{2}})^2\).
Recap on combinations of operations

To summarise everything we have seen above about this, operations are done in ‘BEDMAS’ order, i.e.

Brackets, Exponents, Division, Multiplication, Addition, Subtraction.

Otherwise, we work from left to right.

1.2 Algebra

We use algebra to express and manipulate information about unknown quantities. These unknown quantities are called variables and these are normally represented by letters such as $x$, $y$ and $z$. One way to think of this is that numbers are constants, i.e. they always have the same value, whereas variables can take different values depending on the context.

1.2.1 Algebraic expressions

An algebraic expression is a sequence of numbers, variables and operations, e.g. $4x + 3y - 7$. In expressions such as this, $4x$ means $4 \times x$, i.e. four lots of $x$. As such, we can see that, for any value of $x$, we have things like

$$4x + 3x = 7x,$$

as four lots of $x$ plus three lots of $x$ is seven lots of $x$. Note that all of the mathematical operations that we have seen so far can be used in algebraic expressions.

Attributing meaning to algebraic expressions

Often, we use mathematical expressions to represent the value of some quantity. For instance, we can consider the following examples.

1. If you have a job which pays £10 per hour and you work $x$ hours, then your income is given by the algebraic expression £10$x$.

2. If a firm has a revenue of £$x$ and costs of £$y$, then its profit is £($x - y$).

3. If a firm prices a product at £$x$ per unit and sells $x$ units of this product, then the revenue is £$x^2$. If the costs are £$x$, then its profit is £($x^2 - x$).

As the above examples show, some algebraic expressions contain one variable, such as $4x + 3x$, some contain two variables, such as $4x + 3y - 7$, and some can contain one variable used several times, such as $x^2 - x$ where $x$ is used twice (i.e. once in an $x$ term and once in an $x^2$ term). Of course, the quantities represented may be more complicated than those given in these examples.
1. Review I — A review of some basic mathematics

Example 1.1  Suppose that you heat your house with gas for \(d\) days per year and on each day you use \(m\) cubic metres of gas. This means that you use \(dm\) cubic metres of gas each year.

If gas costs \(L P\) per cubic metre, this means that the cost of heating your house for a year is \(LdmP\).

Suppose that you must also pay a fixed amount of \(L81\) per year to the gas company. This means that the cost of heating your house for a year is now \(L(dmP + 81)\).

Suppose that you pay your gas bill in twelve equal monthly instalments, this means that you must pay

\[
\frac{LdmP + 81}{12}
\]

every month.

Activity 1.4  What will the annual payment be if the gas company raises the price of gas by \(Lp\) per cubic metre? What will the corresponding monthly repayments be?

Evaluating algebraic expressions

Given an algebraic expression, we are sometimes given specific values for each of the variables involved and asked to evaluate it, i.e. find a value for the whole algebraic expression given the values of the variables. So, for example, using our examples above we have the following.

1. With \(x = 5\), you have a job which pays \(L10\) per hour and you work 5 hours, then your income is given by \(L(10 \times 5) = L50\).

2. With \(x = 40\) and \(y = 30\), the firm has a revenue of \(L40\) and costs of \(L30\), and so its profit is \(L(40 - 30) = L10\).

3. With \(x = 10\), the firm prices the product at \(L10\) per unit and sells 10 units, i.e. the revenue will be \(L10^2\). The costs will be \(L10\), and so its profit is \(L(10^2 - 10) = L(100 - 10) = L90\).

Indeed, we can also look at how this works in our more complicated example.

Example 1.2  Following on from Example 1.1, suppose that when heating your house, gas costs \(L0.12\) per cubic metre and that you use 13 cubic metres of gas per day for 125 days. This means that we have to pay

\[
\frac{L13 \times 125 \times 0.12 + 81}{12} = \frac{L195 + 81}{12} = \frac{L276}{12} = L23
\]
every month.
Activity 1.5 What is the cost of heating your house for a year?

What will the annual payment be if the gas company raises the price of gas by 8p per cubic metre? What will the corresponding monthly repayments be?

Simplifying algebraic expressions

As long as we take care to combine ‘like with like’, an algebraic expression can sometimes be simplified, i.e. it can be changed into a form that is easier to evaluate without altering what we will get from an evaluation. For example, we saw earlier that

\[ 4x + 3x = 7x, \]

and so we can write \( 4x + 3x \) as \( 7x \), which is simpler. In particular, we can often simplify expressions by removing brackets from an expression and simplifying what remains, e.g. if we have an algebraic expression like \( 3(2x) \) we can think of this as ‘three lots of \( 2x \)’ which gives us \( 6x \), i.e.

\[ 3(2x) = 6x. \]

However, if we have an algebraic expression like \( 3(x + 2) \), which we can think of as ‘three lots of \( x + 2 \)’, we can remove the brackets by multiplying everything inside the brackets by 3, i.e.

\[ 3(x + 2) = 3x + 6, \]

whereas if we have an algebraic expression like \( -(2x - 1) \), we can think of the minus as telling us to multiply everything inside the brackets by \(-1\), i.e.

\[ -(2x - 1) = -2x + 1. \]

Indeed, we may be able to do some simplifying after we have multiplied out the brackets, e.g.

\[ 2(x + 3) + x = 2x + 6 + x = 3x + 6, \]

where, here, we have multiplied out the brackets and collected ‘like’ terms to get a simpler expression. Some other examples of simplifying algebraic expressions are:

- \( 4x - 3x = x, \)
- \( 4(2x) - x = 8x - x = 7x, \)
- \( 3(x + y) = 3x + 3y, \)
- \( 3(x + 1) + 4(x - 1) = 3x + 3 + 4x - 4 = 7x - 1, \) and
- \( 3(x + 1) - 4(x - 1) = 3x + 3 - 4x + 4 = -x + 7. \)

Notice that none of these simplifications changes the outcome of any evaluation which we may want to perform, i.e. whatever we get if we evaluate the expression at the start we will also get if we evaluate the expression at the end. In this sense, the expressions may look different, but algebraically they are the same throughout.
Multiplying out two pairs of brackets

Sometimes we will want to multiply out the brackets in more complicated expressions. For example, how would you remove the brackets from \((x + 3)(y - 2)\)? We can think of this in two ways:

- Multiplying out the first bracket, everything in the first bracket needs to be multiplied by the second bracket, i.e.
  \[(x + 3)(y - 2) = x(y - 2) + 3(y - 2),\]
  and then simplifying this as before we get
  \[(x + 3)(y - 2) = x(y - 2) + 3(y - 2) = xy - 2x + 3y - 6.\]

- Multiplying out the second bracket, everything in the second bracket needs to be multiplied by the first bracket, i.e.
  \[(x + 3)(y - 2) = (x + 3)y + (x + 3)(-2),\]
  and then simplifying this as before we get
  \[(x + 3)(y - 2) = (x + 3)y + (x + 3)(-2) = xy + 3y - 2x - 6.\]

But, notice that these are the same expression, and so we can multiply out in either way as long as we make sure that every term in a bracket is multiplied by every term in the other bracket.

**Activity 1.6** We can write \((x + 3)^2\) as \((x + 3)(x + 3)\). Use this to remove the brackets from the expression \((x + 3)^2\). In a similar manner, remove the brackets from the expression \((2x + 3)^2\).

Factorising

Sometimes we can simplify expressions even further by putting brackets in, e.g. going back to an earlier example, we could write

\[2(x + 3) + x = 2x + 6 + x = 3x + 6 = 3(x + 2),\]

as \(3(x + 2) = 3x + 6\) if we multiply out the brackets. The process of putting brackets into an expression is called factorisation. For our current purposes, we just need to note that we can factorise when every term in our expression has a common factor, such as 3 in the example above. Some other examples, which can be verified by multiplying out the brackets, are:

- \(2x - 6 = 2(x - 3),\)
- \(-2x - 10 = -2(x + 5),\) and
- \(3xy - 12y = 3y(x - 4).\)

We will return to factorisation in Unit 3.
1. Review I — A review of some basic mathematics

1.2.2 Equations, formulae and inequalities

So far, we have considered how to manipulate algebraic expressions and what they may be used to express. We now look at the ways in which a pair of algebraic expressions may be related to one another.

Equations

An *equation* is a mathematical statement which sets two algebraic expressions equal to one another. For example, \( a = b \), \( x^2 = 4 \) and \( x + 3 = -2x + 4 \) are all equations.

A *solution* to an equation is a value for each variable in the equation which is such that, when we evaluate both expressions with these values substituted for the variables, the expressions are equal. For example, \( x = 3 \) is a solution of the equation \( x^2 - 3 = 2x \) as, substituting \( x = 3 \) into both sides we get the same number, i.e. 6. Sometimes, an equation can have more than one solution. For example, \( x = -1 \) is also a solution of \( x^2 - 3 = 2x \) as, substituting \( x = -1 \) into both sides we get the same number, i.e. \(-2\).

Generally speaking, as we shall see in Units 2 and 3, a given equation may have no solutions, one solution or many solutions.

Solving an equation is to find all of its solutions. Sometimes this is easy and sometimes it is not so easy to do this. In the simplest case, we just have to simplify both sides to see the solution. For example, to solve the equation \( 4x - 3x = 2 + 5 \), we simplify both sides to see that \( x = 7 \).

If this doesn’t work, we can *rearrange* the equation into a simpler equation that has the same solution(s). To do this, we proceed by performing some well-chosen mathematical operation on both sides at the same time so that the equation is unchanged, but simplified. The mathematical operations that we can use in such cases are:

- add (or subtract) an expression from both sides;
- multiply (or divide) by a non-zero expression on both sides.

But, raising both sides to a power can cause problems as if we were squaring both sides, say, we know that a positive expression has two square roots. For example, the equation \( 4x - 8 = 2x + 4 \) has the same solutions as the equations

\[
4x - 8 - (4x + 4) = 2x + 4 - (4x + 4),
\]

and

\[
\frac{4x - 8}{9} = \frac{2x + 4}{9},
\]

but, it has different solutions to the equation

\[
(4x - 8)^2 = (2x + 4)^2.
\]

Bearing this in mind, let’s see how we would actually solve this equation.
Example 1.3  Solve the equation $4x - 8 = 2x + 4$.

We solve this by rearranging it, i.e. performing some well chosen mathematical operations on both sides at the same time:

\[
egin{align*}
4x - 8 &= 2x + 4 & \text{our equation} \\
4x - 8 - 2x &= 2x + 4 - 2x & \text{subtracting } 2x \text{ from both sides} \\
2x - 8 &= 4 & \text{simplifying} \\
2x - 8 + 8 &= 4 + 8 & \text{adding 8 to both sides} \\
2x &= 12 & \text{simplifying} \\
x &= 6 & \text{dividing both sides by 2}
\end{align*}
\]

Thus, the solution to our equation is $x = 6$.

Lastly, always check that any solution you find is a solution by using it to evaluate both sides of the original equation.

| Activity 1.7 | Check that $x = 6$ is a solution to the original equation. |

Example 1.4  Solve the equation $3x + 6 = 5x - 10$.

We again proceed by rearranging the equation:

\[
egin{align*}
3x + 6 &= 5x - 10 & \text{our equation} \\
3x + 6 - 3x &= 5x - 10 - 3x & \text{subtracting } 3x \text{ from both sides} \\
6 &= 2x - 10 & \text{simplifying} \\
6 + 10 &= 2x - 10 + 10 & \text{adding 10 to both sides} \\
16 &= 2x & \text{simplifying} \\
8 &= x & \text{dividing both sides by 2}
\end{align*}
\]

Thus, the solution to our equation is $x = 8$.

| Activity 1.8 | Check that $x = 8$ is a solution to the equation $3x + 6 = 5x - 10$. |

The equations in the last two examples are linear equations and they will be the starting point for a more detailed discussion of equations that will start in Unit 2.

Inequalities

An inequality is a mathematical statement where two algebraic expressions are related by an inequality, such as ‘$>$’ or ‘$<$’, so that we know that one of the expressions is greater than or less than the other. Inequalities can be solved by finding the range of values, for each variable, that make it true. For example, the inequality $x < 2$ is true precisely when $x < 2$. 
As with equations, inequalities can be solved by rearranging them into simpler inequalities that are true for the same range of values. Generally, given an inequality, this means that we can:

- add (or subtract) an expression from both sides, or
- multiply (or divide) by a positive expression on both sides,

to simplify, but not change, the inequality. For example,

- \(x + 4 > -1\) can be simplified to give \(x > -5\) by subtracting 4 from both sides.
- \(3x > 6\) can be simplified to give \(x > 2\) by dividing both sides by 3 (as 3 is positive).

However, if we multiply (or divide) by a negative expression, we must ‘reverse the direction’ of the inequality. For example,

- \(-3x > 6\) can be simplified to give \(x < -2\) by dividing both sides by \(-3\) and reversing the direction of the inequality (as \(-3\) is negative).

To see why we need to do this, consider the inequality \(2 < 3\) which is true. If we multiply by 2 (which is positive) we get \(4 < 6\) which is still true, but if we multiply by \(-2\) (which is negative) we get \(-8 < -12\) which is not true. However, if we reverse the direction of the inequality as well, we get \(-8 > -12\) which is now true.

**Example 1.5** Solve the inequality \(4x - 6 < 6x - 2\).

We solve this by rearranging it, i.e. performing some well chosen mathematical operations on both sides at the same time:

\[
\begin{align*}
4x - 6 &< 6x - 2 & \text{our inequality} \\
4x - 6 - 4x &< 6x - 2 - 4x & \text{subtracting } 4x \text{ from both sides} \\
-6 &< 2x - 2 & \text{simplifying} \\
-6 + 2 &< 2x - 2 + 2 & \text{adding } 2 \text{ to both sides} \\
-4 &< 2x & \text{simplifying} \\
-2 &< x & \text{dividing both sides by } 2
\end{align*}
\]

Thus, the solution to our inequality is \(-2 < x\), or rewriting this, \(x > -2\).

Alternatively, we could have rearranged it by doing some slightly different operations:

\[
\begin{align*}
4x - 6 &< 6x - 2 & \text{our inequality} \\
4x - 6 - 6x &< 6x - 2 - 6x & \text{subtracting } 6x \text{ from both sides} \\
-2x - 6 &< -2 & \text{simplifying} \\
-2x - 6 + 6 &< -2 + 6 & \text{adding } 6 \text{ to both sides} \\
-2x &< 4 & \text{simplifying} \\
x &> -2 & \text{dividing both sides by } -2 \text{ and reversing the inequality}
\end{align*}
\]

Thus, the solution to our inequality is, again, \(x > -2\).
1. Review I — A review of some basic mathematics

Formulae

A formula is an algebraic expression where a single variable, the subject, is equated to an expression involving other variables. For example, the area, $A$, of a circle is given in terms of its radius, $r$, by the well-known formula $A = \pi r^2$. Sometimes we will want to rearrange a formula so that a different variable is the subject. The procedure for doing this is the same as the one we used to solve an equation, but the ‘solution’ will be an algebraic expression rather than a number.

**Example 1.6** Following on from Example 1.1, let $S$ denote the amount, in pounds, of our monthly gas payments so that

$$ S = \frac{dmP + 81}{12}. $$

If our monthly repayment, $S$, is now given, for how many days, $d$, can we heat our house?

We proceed by rearranging the formula:

$$ S = \frac{dmP + 81}{12} \quad \text{our formula} $$

$$ 12S = dmP + 81 \quad \text{multiplying both sides by 12} $$

$$ 12S - 81 = dmP \quad \text{subtracting 81 from both sides} $$

$$ \frac{12S - 81}{mP} = d \quad \text{dividing both sides by } mP $$

Thus we can see that the number of days is given by

$$ d = \frac{12S - 81}{mP}. $$

**Activity 1.9** In a similar manner, find the price, $P$, per cubic metre of gas.

Identities

An identity is a special kind of mathematical formula that allows us to rewrite one mathematical expression in another way. For instance,

$$ x(x + 1) = x^2 + x, $$

is an identity because reading it from left to right tells us how to multiply out the brackets in ‘$x(x + 1)$’ and reading it from right to left tells us how to factorise the quadratic $x^2 + x$. In particular, notice that although this looks like an equation, it isn’t really because it is true for all values of $x$! In fact, throughout this unit we have been reviewing how certain mathematical operations work and, as you have probably realised, many of these can be usefully summarised by using identities. For instance, the following identities allow us to summarise some of the ideas that we encountered when we discussed fractions.
1. **Review I — A review of some basic mathematics**

### Arithmetic with fractions

To add and subtract fractions, we use the rules

\[
\frac{a}{b} + \frac{c}{d} = \frac{a + c}{bd} \quad \text{and} \quad \frac{a}{b} - \frac{c}{d} = \frac{a - c}{bd},
\]

where \(bd\) is called the common denominator. To multiply fractions we use the rule

\[
\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd},
\]

and we divide fractions by using the rule

\[
\frac{a}{b} \div \frac{c}{d} = \frac{a}{b} \times \frac{d}{c},
\]

where \(d/c\) is called the reciprocal of \(c/d\).

At this stage, we can also usefully summarise some of the ways in which powers work as follows.

### Power laws

The power laws state that

\[
a^n \cdot a^m = a^{n+m} \quad \frac{a^n}{a^m} = a^{n-m} \quad (a^n)^m = a^{nm}
\]

provided that both sides of these expressions exist. In particular, we have

\[
a^0 = 1 \quad \text{and} \quad a^{-n} = \frac{1}{a^n}.
\]

If it exists, we also define the *positive* \(n\)th root of \(a\), written \(\sqrt[n]{a}\), to be \(a^{1/n}\).

We can also summarise some of our results concerning brackets by using identities as you can see in the next activity.

**Activity 1.10** Write out the identities that arise when you remove the brackets from the following algebraic expressions.

i. \(a(bc)\),

ii. \(a(b + c)\),

iii. \((a + b)^2\),

iv. \((a + b)(c + d)\).

And, just to be sure that we understand what is going on, try the next activity.

**Activity 1.11** Use these identities to simplify the following algebraic expressions.

i. \((x + y)^2 - x(x + y) - y(x + y)\),

ii. \(\frac{(x + y)^2 - (x - y)^2}{4xy}\),

iii. \(\sqrt{x + y} - (\sqrt{x} + \sqrt{y})\).
Learning outcomes

At the end of this unit, you should be able to:

- simplify and evaluate arithmetic expressions including those that involve brackets and powers;
- manipulate algebraic expressions including those that involve brackets and powers;
- solve simple equations and inequalities;
- model certain situations using formulae and be able to rearrange such formulae;
- use identities to manipulate arithmetic and algebraic expressions.

Exercises

Exercise 1.1
Evaluate the expressions $3 \cdot 2 + 6 \cdot 7 + 4$ and $\frac{3 - (3 - (4 - 5) - 2) - 6}{-(-(-1)) - 1}$.

Exercise 1.2
Evaluate the following expressions.

i. $|-3| + |-2|$,  
ii. $|-3| - |-2|$,  
iii. $-|3| + |-2|$,  
iv. $-|3| - |-2|$,  
v. $-|3| - |2|$.

Exercise 1.3
Write the proper fractions $\frac{4}{7}$, $\frac{2}{3}$ and $\frac{1}{4}$ as improper fractions.

Exercise 1.4
Evaluate the following expressions, writing your answers in lowest terms.

i. $\frac{1}{3} + \frac{1}{2}$,  
ii. $\frac{30}{7} - \frac{5}{3}$,  
iii. $\frac{2}{5} \cdot \frac{25}{4}$,  
v. $\frac{13}{8} + \frac{9}{4}$.

Exercise 1.5
You deposit £1000 in a bank account that pays 10% interest. What will the balance be after one year? Two years?

After two years, what is the increase in the balance as a percentage of the original deposit?

Exercise 1.6
Evaluate the following expressions.

i. $9^2 - 9 \frac{1}{2}$,  
ii. $16^{-\frac{1}{4}} + 16^{\frac{1}{2}}$,  
iii. $7^{\frac{1}{2}} \cdot 7^{\frac{2}{3}}$,  
v. $\frac{10^{-2}}{2^{-10}}$.
1. Review I — A review of some basic mathematics

Exercise 1.7
Express the following in the simplest form possible.

i. \( \frac{x^2y}{2xz} + \frac{x^3z}{xy} \),  
ii. \( x(y^2z^3)^{\frac{1}{2}}(xz)^{-2} \),  
iii. \( x(xy)^{-2}(x + z)^{\frac{1}{2}} \).

Exercise 1.8
Multiply out the brackets in the following expressions simplifying your answers as far as possible.

i. \( (x + 1)(x - 1) \),  
ii. \( (2y + 3)(y - 2) \),  
iii. \( (x + 3y)(2x - y) \),  
iv. \( (2x - 3y)(x + z) \).

Exercise 1.9
Solve the following equations.

i. \(-3p = 21\),  
ii. \(4q - 1 = 15\),  
iii. \(5z + 4(z - 2) = 1\),  
iv. \(\frac{5}{6}k - 2k + \frac{1}{3} = \frac{2}{3}\),  
v. \(5m - 3(m - 2) = 11(m + 2)\),  
vi. \(83(w - 1,996) + 17(w - 1,996) = 600\).

Exercise 1.10
You hire a car for £20 plus the cost of petrol used. Let \(x\) be the distance you travel in miles and \(p\) be the price, in pence, of petrol per gallon. If petrol consumption is 30 miles per gallon, write down expressions, in pence, for the amount you spend on petrol and the cost per mile.

Exercise 1.11
Rearrange the formula \(y = \frac{z}{2 + x} - 3\) to make \(x\) the subject.

Exercise 1.12
Solve the inequality \(5 - x > 2x - 1\).
Part 2
Statistics
Introduction to Statistics

Syllabus

This half of the course introduces some of the basic ideas of theoretical statistics, emphasising the applications of these methods and the interpretation of tables and results. The Statistics part of this course has the following syllabus.

- **Data exploration:** The statistics part of the course begins with basic data analysis through the interpretation of graphical displays of data. Univariate, bivariate and categorical situations are considered, including time series plots. Distributions are summarised and compared and their patterns discussed. Descriptive statistics are introduced to explore measures of location and dispersion.

- **Probability:** The world is an uncertain place and probability allows this uncertainty to be modelled. Probability distributions are explored to describe how likely different values of a random variable are expected to be. The Normal distribution is introduced and its importance in statistics is discussed. The concept of a sampling distribution is explored.

- **Sampling and experimentation:** An overview of data collection methods is followed by how to design and conduct surveys and experiments in the social sciences. Particular attention is given to sources of bias and conclusions which can be drawn from observational studies and experiments.

- **Fundamentals of regression:** An introduction to modelling a linear relationship between variables. Interpretation of computer output to assess model adequacy.

Aims of the course

The aims of the Statistics part of this course are to provide:

- a basic knowledge of how to summarise, analyse and interpret data
- an insight into the concepts of probability and the Normal distribution
- an overview of sampling and experimentation in the social sciences
- an introduction to modelling a linear relationship between variables.

Treatment is at an elementary mathematical level throughout, so you should be comfortable with the material covered in ‘Unit 1: Review I — A review of some basic mathematics’ of the Mathematics part of the subject guide.
Learning outcomes for the course (Statistics)

At the end of the Statistics part of the course, you should be able to:

- interpret and summarise raw data on social science variables graphically and numerically
- appreciate the concepts of a probability distribution, modelling uncertainty and the Normal distribution
- design and conduct surveys and experiments in a social science context
- model a linear relationship between variables and interpret computer output to assess model adequacy.

Textbook

As previously mentioned in the main introduction, this subject guide has been designed to act as your principal resource. The following textbook is referenced throughout the Statistics part of the course.


This has been indicated as ‘background reading’ meaning it is not essential, but you could benefit from reading it if you find any of the material in the guide difficult to follow.
Unit 11: Data exploration I
The nature of statistics

Overview

We begin the Statistics section of the course with data exploration, arguably the single most important part of any data analysis. To make sense of any data, we must first ‘understand’ the basic features of each variable under consideration. Visualising data communicates a wealth of information to even non-technical audiences. Data exploration presents different ways of presenting data graphically depending on the type of variable(s) being explored. We then move on to descriptive statistics (measures of location and measures of dispersion) which are commonly-used statistics in the social sciences whose roles are to ‘describe’ or ‘summarise’ data numerically.

Aims

This unit explains the nature of statistics providing a gentle introduction to the discipline. The concept of ‘data’ is explored including the different types of data which may be obtained. The role of statistics in the research process is also discussed. Particular aims are:

- to demonstrate how social scientists familiarise themselves with datasets prior to further analysis
- to introduce the different types of data that can occur
- to explain how statistics can be used to conduct social research.

Background reading


11.1 Introduction

So just what is ‘Statistics’? Well, there are several possible definitions. A good working one is:

‘the study of data, involving the collection, classification, summary, display, analysis and interpretation of numerical information’.

We consider each of these briefly.
Statistics is largely concerned with **data**. This is a plural noun meaning ‘given things’ or more loosely ‘information’ or ‘facts’.

Sometimes we look at non-numerical data such as sex (‘gender’) or social class, but usually we are concerned with **numerical** information. The primary objective is to determine what the data tell us about the underlying context in business, economics, society, medicine etc.

### 11.1.1 Data collection

We can do this in several ways:

- **Direct observation**, for example driver behaviour on a motorway.
- **Simulation** of data, by computer, using certain assumptions. For example, what is the likely effect on traffic flow if the speed limit is changed?
- An **experiment**, for example, some patients are given an active drug and others a placebo.
- A **survey** to find out more about consumers or voters (or computers or cars).

The main distinction between an experiment and a survey is that in the former case there is some sort of intervention by the researcher. Most, although not all, of the statistics you may go on to carry out (in finance, politics etc.) are likely to be based on survey data.

### 11.1.2 Data classification

We mention the types of data shortly — this will have an important impact on how they should be analysed. It is a very good idea to check and ‘clean’ data in practice to make sure there are no obvious outliers (anomalous values — more on these later in the unit) which may need to be excluded, and to ensure there are no recording errors. Of course, computers play a vital role in all areas of statistics, although they are not used explicitly in this course.

### 11.1.3 Data summary

This is discussed in more detail later on in this unit. The idea is to get a quick picture of what the ‘typical’ data value is, as measured by an average such as the mean; to assess the spread of the data, as measured typically by the variance; and to see if the data are symmetric, as measured by the skewness.

### 11.1.4 Data display

This refers to tables, graphs and charts. The purpose is not to produce a pretty picture but to gain insight into the data and their context. A simple display is often clearest and best. In some cases, the display alone is sufficient; there is no need for any formal mathematical or statistical study.
11.5 Analysis

This is the heavy part of Statistics. Most of the time, the methods used are well-established, so it is only necessary to learn the relevant technique. It is important to understand that most methods depend on certain assumptions about the data. If these assumptions fail to hold, the conclusions are likely to be invalid.

11.6 Interpretation

Outside a few universities and research institutes, clear interpretation is vital! Interpretation should be understandable by managers and others without formal statistical training. For example, do not say ‘the $p$-value of 0.02 shows that the result of the $t$-test is significant’, but rather ‘there is evidence that men and women differ in their attitude to a policy of lowering taxes’.

11.7 Uncertainty

In general, what is being measured is subject to uncertainty, or random variation. For example, two randomly chosen groups of 100 voters will not give exactly the same outcomes.

We often wish to establish whether a change or a difference (between men and women, left-wing and right-wing voters etc.) can be put down to chance, or whether it is the result of some real effect. We study probability largely in order to measure this uncertainty.

11.8 Descriptive and inferential statistics

It is convenient to distinguish two approaches:

- **Descriptive statistics** comprises those methods concerned with describing a set of data so as to yield meaningful interpretations.

- **Statistical inference** comprises those methods concerned with analysing a subset of data so as to draw conclusions about the entire set of data.

While we are defining things, let us formalise a little. The **population** is the collection of all individuals or items under consideration. A **sample** is that part of the population from which information is obtained when inference is used.

**Example 11.1**

- A manufacturer of tyres wants to estimate the average life of a tyre. This is an inferential study: the population consists of all tyres produced, the sample consists of 100 (or 50, or 500, or 5,000) tyres that are examined.

- A sports writer wants to list the times taken to run 100m in Olympic Games over 60 years. This is a descriptive study.

*p*-values occur in hypothesis testing, which is a form of statistical inference which is not covered in this course.
A politician wants to know how many votes were cast for her party in her region at a recent election. This is a descriptive study.

An economist estimates the average income of all California residents. This is an inferential study: the population consists of all residents, the sample consists of the subset examined.

Notice that in an inferential study, it is the properties of the population that we wish to determine. You could argue that it would be better to examine all population members. This is known as conducting a census. But this will usually be slower, more costly and may sometimes be impossible. Consider a census of all the trees in the UK — or all the fish in the Atlantic Ocean!

The main thing to ensure is that the sample is representative of the population. This is most easily done using a simple random sample, where each population member has an equal chance of inclusion in the sample, although there are alternatives. We will explore this further in the ‘Sampling and experimentation’ section of the course.

It may not come as a surprise that, generally speaking, descriptive statistics are more easily carried out than inferential statistics. Sadly, descriptive statistics are often poorly done, or even omitted completely in practical contexts, as well as student work. This is a shame because they can tell us a great deal about the data, and can even render inferential statistics unnecessary. As a rule, any data analysis should start with descriptive statistics.

### 11.2 Types of data

There are several types of data, and it is important to know which one we are dealing with, so that the correct statistical procedure is used.

- **Categorical** data (also known as qualitative data) give information about the discrete groups into which a population or sample is divided. These may be nominal or ordinal.
  - **Nominal** data are unranked. For example, a group of individuals may be classified by gender, eye colour, blood type (A, B, AB, O), or religion etc.
  - **Ordinal** data are ranked. They give information about order or rank on a scale. For example, a group of students may be classified by the grades they receive in an examination (A, B, C etc.). So-called Likert scales are ordinal (this course is ‘very interesting’, ‘interesting’, ‘quite interesting’, ‘not very interesting’, ‘boring’). Investments can be graded by risk on an ordinal scale.

- **Metric** data are numerical values on some continuous scale. They may be interval or ratio data.
  - **Interval** data are measured on a continuous scale and have the property that the differences between numbers have a meaning. For example, centigrade temperatures are interval data: the difference between 150 and 160 is the same as the difference between 250 and 260, but both are different from the
difference between 150 and 200. The current time (for example, 19:34) is also measured on an interval scale.

- **Ratio** data are similar to interval data but now there is an *absolute zero*, and therefore the ratio of two numbers can be given a meaning. For example, height, weight and the length of time an individual has been alive all constitute ratio data. In each of these cases, there is a fixed zero; nobody can have a negative height or weight, or have lived a negative amount of time. In contrast, the zero for centigrade temperatures or the current time is merely a matter of convention. [Notice that Kelvin temperatures do have an absolute zero and are therefore measured on a ratio scale.]

**Example 11.2** In a household survey:

- Sex (gender) of the head of household — nominal.
- Type of heating used — nominal.
- Age of head of household — ratio.
- Thermostat setting in winter — interval.
- Household income — ratio.
- Average monthly electricity bill — ratio.
- Time when heating is switched on — interval.
- Rating of electricity provider on a 10-point scale — ordinal.

Finally, we mention that many datasets considered are on a single attribute, for example weight. Such data are called **univariate** data. Sometimes, we wish to consider two variables together, say the height and weight of a group of individuals. Such data are called **bivariate** data. **Multivariate** data arise when we consider three or more variables together — perhaps height, weight, age and pulse rate.

There are other ways to classify data and the classification is not always precise. However, in most cases it is fairly clear and is sufficient for most applications — in particular, the choice of the correct statistical method.

### 11.3 The role of statistics in the research process

First some definitions:

- **Research**: trying to answer questions about the world in a systematic (scientific) way.
- **Empirical** research: doing research by first collecting relevant information (*data*) about the world.

Research may be about almost any topic: physics, biology, medicine, economics, history, literature etc. Most of our examples will be from the social sciences: economics,
management, finance, sociology, political science, psychology etc. Research in this sense is not just what universities do. Government, business, and all of us as individuals do it too. Statistics is used in essentially the same way for all of these.

**Example 11.3**

It all starts with a question:

- Can labour regulation hinder economic performance?
- Understanding the gender pay gap: what has competition got to do with it?
- Does racism affect health?
- Children and online risk: powerless victims or resourceful participants?
- Refugee protection as a collective action problem: is the EU shirking its responsibilities?
- Do directors perform for pay?
- Heeding the push from below: how do social movements persuade the rich to listen to the poor?
- Does devolution lead to regional inequalities in welfare activity?
- The childhood origins of adult socio-economic disadvantage: do cohort and gender matter?
- Parent care as unpaid family labour: how do spouses share?

We can think of the empirical research process as having five key stages:

1. Formulating the *research question*.
2. *Research design*: deciding what kinds of data to collect, how and from where.
3. *Collecting* the data.
4. *Analysis* of the data to answer the research question.
5. *Reporting* the answer and how it was obtained.

We conclude this section with an example of how statistics can be used to help answer a research question.

**Example 11.4  CCTV, crime and fear of crime**

Our research question is: What is the effect of closed-circuit television (CCTV) surveillance on

- the number of recorded crimes?
the fear of crime felt by individuals?

We illustrate this using part of the following study

Gill and Spriggs (2005): Assessing the impact of CCTV. *Home Office Research Study 292*. Available at

The research design of the study comprised:

- **Target area**: a housing estate in northern England.
- **Control area**: a second, comparable housing estate.
- **Intervention**: CCTV cameras installed in the target area but not in the control area.
- **Compare** measures of crime and fear of crime in the target and control areas, in the 12 months before and 12 months after the intervention.

The data and data collection:

- **Level of crime**: number of crimes recorded by the police, in the 12 months before and 12 months after the intervention.
- **Fear of crime**: a survey of residents of the areas.
  - Respondents: random samples of residents in each of the areas.
  - In each area, one sample before the intervention date and one about 12 months after.
  - Sample sizes:

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>172</td>
<td>168</td>
</tr>
<tr>
<td>Control</td>
<td>215</td>
<td>242</td>
</tr>
</tbody>
</table>

- Question considered here: ‘In general, how much, if at all, do you worry that you or other people in your household will be victims of crime?’ (from 1 = ‘worry all the time’ to 5 = ‘never worry’).

Statistical analysis of the data:

<table>
<thead>
<tr>
<th></th>
<th>Target</th>
<th>Control</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[a]</td>
<td>[b]</td>
<td>[c]</td>
</tr>
<tr>
<td>Before</td>
<td>26</td>
<td>53</td>
<td>0.98</td>
</tr>
<tr>
<td>After</td>
<td>23</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Change</td>
<td>−3</td>
<td>−7</td>
<td>0.98</td>
</tr>
</tbody>
</table>

It is possible to calculate various statistics, for example the *Relative Effect Size* RES = ([d]/[c])/([b]/[a]) = 0.98 is a summary measure which compares the changes in the two areas.
RES < 1, which means that the observed change in reported fear of crime has been a bit less good in the target area.

However, there is uncertainty because of sampling: only 168 and 242 individuals were actually interviewed at each time in each area.

*Confidence interval* for RES includes 1, which means that changes in self-reported fear of crime in the two areas are not statistically significantly different from each other.

### Number of (any kind of) recorded crimes:

<table>
<thead>
<tr>
<th>Target area</th>
<th>Control area</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>[a] Before</td>
<td>[c] Before</td>
<td>1.34 0.79–1.89</td>
</tr>
<tr>
<td>[b] After</td>
<td>[d] After</td>
<td></td>
</tr>
<tr>
<td>112</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>−11</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

Now RES = 1.34 > 1, which means that the observed change in the number of crimes has been worse in the control area than in the target area.

However, the numbers of crimes in each area are fairly small, which means that these estimates of the changes in crime rates are fairly uncertain.

*Confidence interval* for RES again includes 1, which means that the changes in crime rates in the two areas are not statistically significantly different from each other.

In summary, this study did *not* support the claim that introduction of CCTV reduces crime or fear of crime.

(If you want to read more about research of this question, see Welsh and Farrington (2008). Effects of closed circuit television surveillance on crime. *Campbell Systematic Reviews* 2008:17. (See [http://www.campbellcollaboration.org/library.php](http://www.campbellcollaboration.org/library.php))

Many of the statistical terms and concepts mentioned above were not explained. However, it serves as an interesting example of how statistics can be employed in the social sciences to investigate research questions.

Activities 11.1, 11.2 and 11.3 are not concerned with any technicalities of statistics, and they do not ask you to do any calculations yourself (except, perhaps, a little bit in 11.2). Instead, these exercises invite you to think about various topics related to the use of statistics, and to research design more generally. These include such issues as definition and measurement of variables, selection of subjects for studies, and justifiability of claims about causes and effects.

You are asked to think of answers to the questions, using your own reasoning and common sense. You are welcome to discuss the questions with friends. You do not need to worry about getting the answers right or wrong; the only point is to start thinking!

**Activity 11.1** Consider the following statements. Do you think the conclusions are valid? If so, say why. If not, indicate why not: because the logic used is faulty, because any assumptions made are dubious, because the data collection method is...
inappropriate, or for any other reason.

(a) ‘10% of drivers involved in 100 car accidents had previously taken substance X. A parallel study of drivers not involved in accidents showed that only 1% had taken X. Therefore X is a contributory cause of car accidents.’

(b) ‘Five years ago, the average stay of patients in this hospital was 21 days. Now it is 16 days. We now cure our patients more quickly.’

(c) ‘We wanted to see if the public approved of our plans to transfer resources to elderly patients. We carried out a large-scale survey based on 800 daytime city centre interviews. We found 79% of respondents approved our plans. Therefore we have public backing.’

(d) ‘Nugro is the revolutionary hair restorer for men. A sample of 100 men with thinning hair was selected to apply Nugro lotion every day for a month. Of these, 77 reported new hair growth. Nugro is proven to be effective in the treatment of male baldness.’

Activity 11.2 The following cross-tabulation shows data on the 3,593 people who applied to graduate study at the University of California, Berkeley, in 1973. The table classifies the applicants according to their sex, and whether or not they were admitted to the university.

<table>
<thead>
<tr>
<th>Sex</th>
<th>Admitted</th>
<th>% Yes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1,180</td>
<td>686</td>
<td>1,866</td>
</tr>
<tr>
<td>Female</td>
<td>1,259</td>
<td>468</td>
<td>1,727</td>
</tr>
<tr>
<td>Total</td>
<td>2,439</td>
<td>1,154</td>
<td>3,593</td>
</tr>
</tbody>
</table>

The table shows that 36.8% of male applicants, but only 27.1% of female applicants, were admitted.

Bob observes this and concludes that in that year Berkeley practised discrimination against female applicants. Amy, however, decides to take another look at the statistics. She adds one more piece of data, the department to which each person applied, and creates cross-tabulations separately for each department (which are labelled A, B, C, D and E). These tables are shown below. For example, the first table cross-classifies the sex and admission status of just those 585 people who applied to department A, and so on.

Amy examines her tables and states that she disagrees with Bob: there is no evidence of discrimination. Why does she conclude this? Why do Amy and Bob come to different conclusions? Which one do you agree with?
### Activity 11.3

Each of the statements below mentions a piece of statistical evidence, and a claim based on it. Do you agree with the claims? Why or why not? Are there any fallacies in the claims, or complications which are being glossed over? The questions marked with (†) are a bit more subtle and complex than the rest.

(a) A public consultation exercise on attitudes to genetically modified (GM) food was carried out in the UK in 2002–3. This involved various events where interested members of the public could come and take part in discussions about GM food. After the events, the participants were asked to complete a questionnaire, which was also available on a website. Around 37,000 people completed the questionnaire, and 90% of those expressed opposition to GM foods. Therefore a very large majority of the people in the UK oppose GM foods.

(b) In a study of the ages and professions of people who had died, it was found that the profession with the lowest average age of death was ‘student’. Therefore being a student is the most dangerous of professions.

(c) In 2007, the officially recorded suicide rate in Sweden was 15.8 per 100,000 people per year. This was much higher than in many other countries, some of which even had a rate of 0.0. This indicates that suicide is a much more serious problem in Sweden than in those other countries.
(d) Data on the past 10 years in a country show that the number of deaths from drowning tends to be higher in months when total consumption of ice cream is high. Therefore eating ice cream before going swimming increases the risk of drowning.

(e) A country has two kinds of secondary schools, private schools and state-owned schools. Statistics show that 40% of students graduating from private schools, but only 20% of those graduating from state schools, go on to study at a university. Therefore, private schools are twice as good as state schools.

(f) Sociologists conduct a study where they select a random sample of people and ask these people for a list of their close friends. A random sample of the people named as friends is then contacted and the survey is repeated. The people sampled at the second stage have, on average, many more friends than do the people in the original sample. Therefore, your friends have more friends than you do.

11.4 Summary

This introductory unit has outlined the purpose of statistics and the role the discipline plays in the research process. Preliminary considerations of issues relating to data collection and analysis were discussed, as well as the different types of data which exist. Having spent some time thinking about the nature of statistics, you are now ready to start doing statistics, beginning with data visualisation in the next unit.

11.5 Key terms and concepts

- Bivariate data
- Census
- Descriptive statistics
- Experiment
- Metric data
- Nominal data
- Population
- Ratio data
- Sample
- Statistical inference
- Univariate data
- Categorical data
- Data
- Direct observation
- Interval data
- Multivariate data
- Ordinal data
- Probability
- Research
- Simulation
- Survey
11. Data exploration I — The nature of statistics

Learning outcomes

At the end of this unit, you should be able to:

- outline issues relating to data collection and analysis
- describe the different types of data
- explain the role of statistics in the research process
- discuss the key terms and concepts introduced in this unit.

Exercises

Exercise 11.1
The given working definition of ‘Statistics’ was:

‘the study of data, involving the collection, classification, summary, display, analysis and interpretation of numerical information’.

What does this mean?

Exercise 11.2
Briefly discuss the distinction between descriptive statistics and inferential statistics.

Exercise 11.3
Explain the different types of data that can occur.

Exercise 11.4
What is the measurement level for each of the following variables?

(a) The quality ranking of a newspaper

(b) The classification of an examination result as ‘Distinction’, ‘Merit’, ‘Pass’, or ‘Fail’

(c) Country of birth

(d) Favourite music

(e) Income measured by percentiles (for example, if someone’s income is above the 20-th percentile, this means 20% of the population earn less).
Exercise 11.5

In 2009 the UK government reclassified cannabis from a class C drug to a class B drug, thereby introducing the threat of arrest for possession of the drug. The following table cross-classifies age and agreement with the reclassification.

<table>
<thead>
<tr>
<th>Age</th>
<th>Agree with reclassification (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. No</td>
</tr>
<tr>
<td>18–39</td>
<td>50</td>
</tr>
<tr>
<td>40–59</td>
<td>?</td>
</tr>
<tr>
<td>60 and over</td>
<td>?</td>
</tr>
</tbody>
</table>

Complete the table in such a way that there is a weak positive association between Age and Agreement. (Assume the measurement scale of Agreement as given in the table is an ordinal one.)
Important note: This Sample examination paper reflects the examination and assessment arrangements for this course in the academic year 2013–2014. The format and structure of the examination may have changed since the publication of this subject guide. You can find the most recent examination papers on the VLE where all changes to the format of the examination are posted.

Mathematics and Statistics

Time allowed: 2 hours.

Candidates should answer ALL questions. Section A (50 marks) covers the Mathematics part of the course, Section B (50 marks) covers the Statistics part of the course. Candidates are required to pass BOTH sections to pass the examination.

Candidates are strongly advised to divide their time accordingly.

A list of formulae and the table of cumulative Normal probabilities is provided at the end of this paper.*

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

*The table is provided at the back of this subject guide in Appendix C.
Section A: Mathematics

Answer ALL questions (50 marks in total).

1. (a) The demand for a product, \( q \), is related to its price, \( p \), by the equation

\[
q = 10 - p,
\]

while suppliers respond to a price of \( p \) by supplying an amount, \( q \), given by the equation

\[
q = 4p - 30.
\]

Find the equilibrium price and the corresponding level of production.

Write down the supply function. For which values of \( p \) and \( q \) is it economically meaningful? (5 marks)

(b) Solve the equation \( \log_2(8) - \log_3(9) = \log_{10}(x) \). (5 marks)

(c) Find the derivatives of the following functions. (5 marks)

i. \( \cos(x^2) \).

ii. \( e^x \cos(x^2) \).

(d) Suppose that you buy a car for £10,000 and its value depreciates continuously at a rate of 25% per year. What is its value after three years? Explain why the car’s value is halved after \( 4 \ln(2) \) years. (5 marks)

[You may use the fact that, to 5dp, \( e^{0.25} = 1.28403 \).]

2. Consider the function \( f(x) = x^3 - 2x^2 - 15x \).

(a) Find and classify the stationary points of \( f(x) \). (5 marks)

(b) Sketch the curve \( y = f(x) \). (5 marks)

(c) Find the area of the region bounded by the curve \( y = f(x) \), the \( x \)-axis and the vertical lines \( x = -1 \) and \( x = 1 \). (5 marks)

3. Consider an annuity which pays £100 every year. The first payment is to be made now and further payments will be made at the end of each year for the next \( n \) years.

(a) Find the present value of this annuity, simplifying your answer as far as possible, given that an interest rate of 5% per annum compounded annually is available to you. (5 marks)

(b) If the annuity is to make eleven payments, what is the smallest lump sum payment that will be worth more to you than the annuity? (3 marks)

(c) How many payments are needed if the annuity is to be worth more than a lump sum of £2,000? (5 marks)

(d) If the annuity was a perpetuity, what would be its present value? (2 marks)

[You may use the facts that, to 5dp, \( 1.09^{11} = 1.71034 \) and \( \log_{1.05}(21) = 62.40033 \).]
A. A sample examination paper

Section B: Statistics

Answer **ALL** questions (50 marks in total).

4. **(a)** Do you think the distribution of income (pounds per year) in the UK is most likely to be symmetrically distributed, skewed to the right, or skewed to the left? Briefly explain why. Which measure of central tendency would you use to describe income? Justify your choice. (5 marks)

**(b)** Given events $A$ and $B$ where $P(A) = 0.5$ and $P(A \cup B) = 0.7$, find $P(B)$ in the following three cases.

i. $A$ and $B$ are mutually exclusive.

ii. $A$ and $B$ are independent.

iii. $P(A|B) = 0.5$. (5 marks)

(c) In an examination, the scores of students who attend schools of type A are approximately normally distributed about a mean of 61 with a standard deviation of 5. The scores of students who attend type B schools are approximately normally distributed about a mean of 64 with a standard deviation of 4. Which type of school would have a higher proportion of students with marks above 70? (5 marks)

(d) You randomly select 1,000 names from the subscription list of a magazine designed for hunters. You mail a questionnaire about gun control to these readers and receive 700 responses. You randomly select 200 of the 700 responses for inclusion in your study. What forms of bias are evident in this design? (5 marks)

5. Three members of an exclusive country club, Mr Adams, Miss Brown and Dr Cooper, have been nominated for the office of president. The probabilities of Mr Adams and Miss Brown being elected are 0.3 and 0.5, respectively. If Mr Adams is elected, the probability of an increase in membership fees is 0.8. If Miss Brown or Dr Cooper is elected, the corresponding probabilities of an increase in membership fees are 0.1 and 0.4, respectively.

**(a)** What is the probability that Dr Cooper is elected president? (5 marks)

**(b)** What is the probability that there will be an increase in membership fees? (5 marks)

**(c)** Given that membership fees have been increased, what is the probability that Dr Cooper was elected president? (5 marks)

Continued overleaf
6. For a group of 15 students, the following table shows the average number of hours per week spent on study and their final results in the corresponding examination.

<table>
<thead>
<tr>
<th>No. of hours studied, $x$</th>
<th>16</th>
<th>17.5</th>
<th>11.5</th>
<th>13.5</th>
<th>15</th>
<th>12.5</th>
<th>20.5</th>
<th>14.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examination mark, $y$</td>
<td>77</td>
<td>85</td>
<td>48</td>
<td>59</td>
<td>75</td>
<td>41</td>
<td>95</td>
<td>72</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of hours studied, $x$</th>
<th>16.5</th>
<th>13.5</th>
<th>22</th>
<th>18.5</th>
<th>17</th>
<th>19.5</th>
<th>19.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examination mark, $y$</td>
<td>80</td>
<td>70</td>
<td>99</td>
<td>85</td>
<td>83</td>
<td>97</td>
<td>89</td>
</tr>
</tbody>
</table>

Summary statistics for these data are:

\[
\sum x_i = 247.5, \quad \sum x_i^2 = 4218.75, \quad \sum y_i = 1155,
\]
\[
\sum y_i^2 = 92999, \quad \sum x_i y_i = 19750.5
\]

(a) Calculate the sample correlation coefficient for these data and comment. 

(b) Calculate the least squares regression line of $y$ on $x$. 

(c) Use the calculated line to predict examination marks for students who studied for 16 hours. Would you consider a prediction based on 20 hours to be more accurate? Explain why/why not.
A. A sample examination paper

Formula sheet

Section A: Mathematics

The chain rule: If $f(x) = f(g)$ for some function $g(x)$, then \( \frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx} \).

The product rule: \( \frac{d}{dx} \left( f(x)g(x) \right) = \frac{df}{dx}g(x) + f(x)\frac{dg}{dx} \).

The quotient rule: \( \frac{d}{dx} \left( \frac{f(x)}{g(x)} \right) = \frac{1}{\left[ g(x) \right]^2} \left( \frac{df}{dx}g(x) - f(x)\frac{dg}{dx} \right) \).

The sum of a finite geometric series is given by

\[
a + ar + ar^2 + \cdots + ar^{n-1} = a \frac{1 - r^n}{1 - r}.
\]

Section B: Statistics

The variances for a population and a sample are

\[
\sigma^2 = \frac{\sum_{i=1}^{N} x_i^2}{N} - \mu^2 \quad \text{and} \quad s^2 = \frac{\left( \sum_{i=1}^{n} x_i^2 \right) - n\bar{x}^2}{n - 1}.
\]

For events $A$ and $B$,

\[
P(A \cup B) = P(A) + P(B) - P(A \cap B),
P(A \cap B) = P(A)P(B \mid A),
P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}.
\]

For a discrete random variable $X$,

\[
\mu = E(X) = \sum x \cdot P(X = x),
E(g(X)) = \sum g(x) \cdot P(X = x),
\sigma^2 = \text{Var}(X) = E(X^2) - (E(X))^2.
\]
If \( X \sim \text{Bin}(n, \pi) \), then
\[
P(X = x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x} \quad \text{where} \quad \binom{n}{x} = \frac{n!}{x!(n-x)!},
\]
\( E(X) = n\pi \) and \( \text{Var}(X) = n\pi(1 - \pi) \).

If \( X \sim \text{Poisson}(\lambda) \), then
\[
P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad E(X) = \lambda \text{ and } \text{Var}(X) = \lambda.
\]

The sample correlation coefficient is
\[
r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}},
\]
where
\[
S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2,
\]
\[
S_{yy} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2,
\]
\[
S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}.
\]

The regression slope and intercept are given by
\[
\hat{b}_1 = \frac{\sum (x_i y_i - n\bar{x}\bar{y})}{\sum x_i^2 - n\bar{x}^2} \quad \text{and} \quad \hat{b}_0 = \bar{y} - \hat{b}_1\bar{x}.
\]