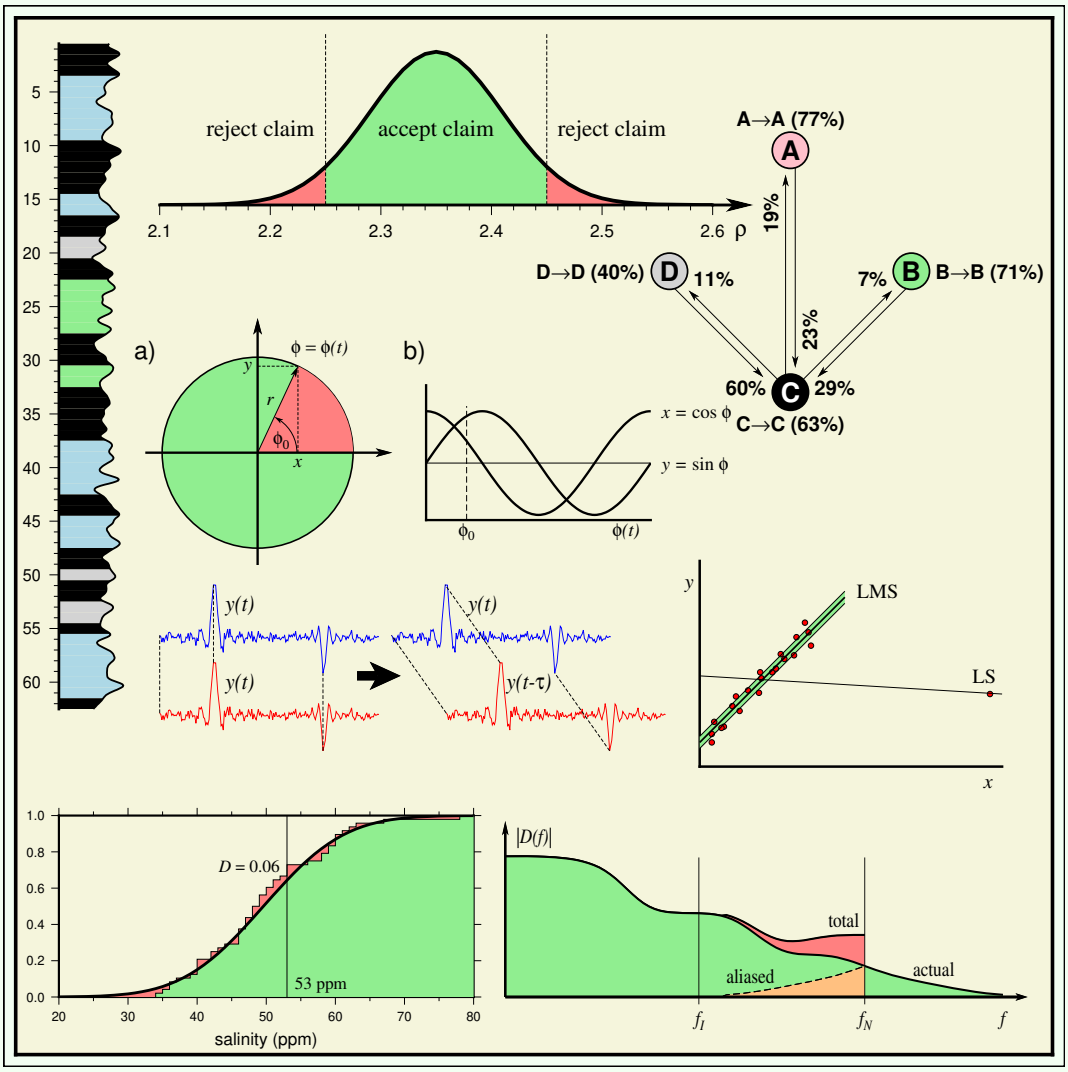
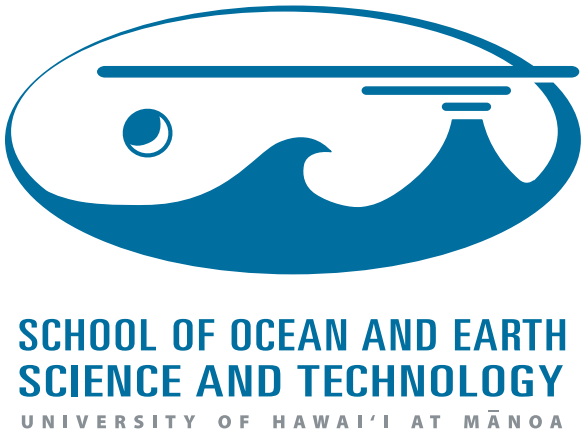
INTRODUCTION TO STATISTICS & DATA ANALYSIS



DEPT. OF EARTH SCIENCES

Department of Earth Sciences
School of Ocean and Earth Science and Technology
University of Hawai'i at Mānoa



Version 2.3 — Revised March 17, 2023

An Open Source project (`https://github.com/PaulWessel/DA1`) by UHM's Department of Earth Sciences.

Contents

Preface

This book was initially devloped by Paul Wessel in support of an introductory course in statistics and data analysis taken by many of our undergraduate majors and some graduate students in the Department of Earth Sciences in the School of Ocean and Earth Science and Technology at the University of Hawaii at Mānoa. Over the years, he expanded the material to support undergraduate students in the broad area of the natural sciences anywhere. There were several goals when in designing this book:

1. We wanted to introduce students of science and environmental engineering to some of the most common methods used in the examination and analysis of simple data sets encountered in the sciences. By learning these techniques your data analysis tool chest will expand, preparing you for further learning when you need it.

2. We sought to fill in many of the intermediate steps in derivations that traditional textbooks will skip. Thus, it should be much easier to trace these derivations since there are no annoying messages of the type "the derivation has been left as an exercise for the reader." The more elaborate derivations also help potential instructors using this book to present such material at the level of detail they require.

3. We wished the book to be affordable. Publishing it as a digital book directly by the authors means it costs a fraction of a traditional, professionally produced textbook. This method also enables timely updates and easy corrections, at no extra cost to the purchaser. Of course, the flip-side of this decision is that the book has not benefited from the help of professional editors and graphics artists (but it *has* been extensively reviewed by other scientists familiar with the methodologies).

4. We hoped to make the book suitable for introductory college level courses in data analysis and statistics, with examples of analysis using real data sets. Thus, there is a variety of assignments at the end of each chapter, for which data sets are available in the repository.

As a science student or curious bystander, why should you consider this book? While many general answers may be given, including the relatively low cost, some are particularly relevant:

1. The natural and environmental sciences are very data intensive, with huge quatitative data sets being collected by improved instrumentation and made readily accessible over the Internet.

2. The job market requires quantitative skills, in particular for those pursuing a career in the natural or environmental sciences. You will be dead in the water if you cannot juggle data to some extent.

3. All sciences need reproducibility of analyses for the testing of relevant hypotheses. This concept goes to the heart of what science and the scientific method are all about.

4. The increased sophistication of newer instruments is transforming visual characterizations into numbers that need to be analyzed quantitatively, hence new methods of analyses continue to be developed.

What is the target audience for this book? There are numerous candidates, and we recommend the book specifically to:

1. Anyone ever planning to analyze data of any kind. This is of course a pretty broad statement but we believe it is true. Once you get into the habit of thinking and working quantitatively, there is no going back.

2. Anyone who wants to be prepared for a changing job market, considering that you are competing for opportunities with other students who likely have been exposed to similar material.

3. Budding scientists, engineers and technical personnel, especially those fearful of mathematics. Many students tend to avoid courses and treatises that expose them to mathematics, thus limiting what they can achieve as scientists. We hope, by showing all the steps in the derivations and presenting the coupling of mathematics to data sets, that this book will help alleviate such fears.

The main purpose of this introductory data analysis book is to prepare you for facing and dealing with data, their limitations and the methods of analysis that may be most suitable for different types of data. We hope to achieve that goal by a multiprong series of attacks:

1. Expose you to many different data analysis techniques and thus broaden your horizon. No book can cover everything but being aware of other approaches allows you to pursue alternative methodologies when the need arises.

2. Make you appreciate why you should fully understand a technique's nuts and bolts before running a "black box" operator, such as most typical software packages, on your data.

3. Make you comfortable with applied mathematics at an intermediate level. The mathematics we employ in this book is mostly algebra, trigonometry and calculus, with an introduction to matrix algebra. There are numerous data analysis techniques that are simply exercises in applied matrix algebra.

4. Make you comfortable with the tools of the trade, such as MATLAB, R, Python, or Octave for your data analysis needs rather than depending on "business software", such a spreadsheets. However, this book is not a recipe collection of algorithms in these languages, but rather presents equations and the assumptions used to pursue specific analysis goals or tests. The author's data analysis website (`http://www.soest.hawaii.edu/pwessel/DA`) contains links to all the data sets used in this book as well as any MATLAB example code discussed.

Finally, data analysis is a very broad subject and the authors are certainly not experts in all the available techniques. Our goal is to give you a flavor of what is available, show you how to find out more, and enable you to make sensible decisions on how to approach and analyze your data.

This book is based on Paul Wessel's collection of course notes that was inspired by several sources. Here are the ones he had found most useful:

1. Course notes from a class in data analysis at Columbia University during his graduate school days, developed by Doug Martinson at Lamont-Doherty Earth Observatory, provided a clear introduction to spectral analysis. His willingness to let us use some of his early material in this book is gratefully acknowledged.

2. John C. Davis' textbook, *Statistics and Data Analysis in Geology, 3rd edition* available from John Wiley and Sons. His textbook has numerous useful problems complete with data sets that are now in the public domain (`http://www.kgs.ku.edu/Mathgeo/Books/Stat`). A few of these are used in this book.

3. *Exploratory Data Analysis* by John W. Tukey, available from Addison-Wesley Publishing Company, outlines the basics of exploratory data analysis.

4. *An Introduction to Error Analysis* by John R. Taylor, available from University Science Books, reviews the various rules for the propagation of errors in compound expressions.

5. *Numerical Recipes* by Press et al., available from Cambridge University Press in multiple language flavors, clarifies many statistical procedures with a dose of nerd humor.

6. *Robust Regression and Outlier Detection* by Peter. J. Rousseeuw and Annick M. Leroy, available from John Wiley and Sons, is a classic for understanding why using robust methods is imperative when dealing with actual data.

7. *Applied Regression Analysis* by Norman R. Draper and Harry Smith, also available from John Wiley and Sons, deals with conventional least squares regressions.

8. *Developments in Geomathematics* by Frederik P. Agterberg, available from Elsevier, deals with a variety of topics, including multiple regressions.

9. Paul Wessel's cumulative experience with analyzing marine geophysical and plate kinematic data since 1985.

Depending on the time available, this book may cover more material than can be presented in a single course. However, many of the topics may be of interest to you at a later stage, hence we hope the book may serve a valuable purpose as a reference.

Paul is grateful to his SOEST colleagues, both past and present. In the early years, the book benefited from feedback from faculty who used the book in their courses (Fred Duennebier, Neil Frazier, Julia Morgan, Rob Dunn, Garrett Apuzen-Ito, and Cecily Wolfe). They provided valuable suggestions and error corrections. Likewise, numerous former students have also contributed by pointing out typographical errors in the equations, odd or foul language, stale jokes, or confusing illustrations. Starting with the 2023 edition these notes will be maintained by the faculty of the Department of Earth Sciences in SOEST, University of Hawaii, and maintenance of the open repository on GitHub will be a community effort. Because the repository is public, anyone may use it and all we ask is to be notified if the reader finds errors, unclear sections, and similar problems.

We are also grateful to our former secretary, Evelyn Norris, who taught herself enough LaTeX to help typeset large portions of this manuscript from printouts of clumsy Microsoft Word documents with obsolete equation editing plug-ins. As software come and go, LaTeX remains solid and post-processing tools allow for new output formats, such as the digital version you are reading. All illustrations are data-driven, original creations using the Generic Mapping Tools (GMT; `https://www.generic-mapping-tools.org`) via individual scripts that produce publication-quality PDF illustrations. The entire processing workflow from LaTeX and GMT source code to PDF is automated via a UNIX makefile.

Because it is relatively easy to update digital books as frequently as required, we hope to hear from you should you find any remaining errors in the text or graphics. By bringing them to our attention (`mailto:earth-da-book@soest.hawaii.edu`) they can be corrected and updated as soon as possible. Thank you for your interest in this book.

Dept. of Earth Sciences, SOEST, UHM, March 2023.

Chapter 1

EXPLORING DATA

> "I never guess. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."
>
> *Sherlock Holmes, Consulting Detective*

Data come in all types and amounts, from a handful of hard-earned analytical quantities obtained after days of meticulous work in a laboratory to terabytes of remotely sensed data simply gushing in from satellites and remotely operated vehicles. It is therefore desirable to have a common language to describe data and to make initial explorations of trends.

1.1 Classification of Data

All observational sciences require data that may be analyzed and explored, which give rise to new ideas for how the natural world works. Such ideas may be developed into simple hypotheses that ultimately can be tested against new data and either be rejected or live to fight another day. New data crush hypotheses every day, hence we have to be careful with and respectful of our data to a much greater extent than our hypotheses and models. We start our exploration of data by considering the various ways we can categorize data, discussing a few basic data properties, and introducing typical steps taken in exploratory data analysis.

1.1.1 Data types

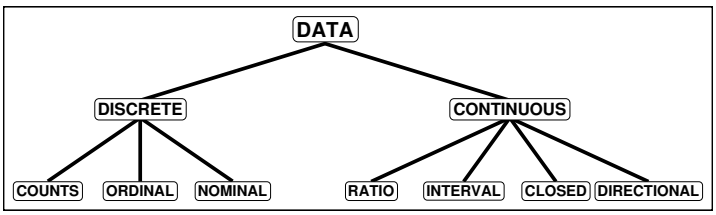


Figure 1.1: Classification of data types. All data we end up analyzing in a computer program are necessarily digitized and hence discrete, but they may represent a *phenomenon* that produced *continuous* output.

Data represent measurements of either discrete or continuous quantities, often called *variables*. *Discrete* variables are those having discontinuous or individually distinct possible values (Figure 1.1). Examples of such data include:

- *Counts*: The flipping of a coin or rolling of dice, or the enumeration of individual items or groups of items. Such data can always be manipulated numerically.

- *Ordinal* data: These data can be ranked, but the intervals between consecutive items are not constant. For instance, consider Moh's hardness scale for minerals (Table 1.1): While *topaz* (hardness 8) is harder than *fluorite* (hardness 4), the values do not imply a doubling in hardness.
- *Nominal* data: These data cannot even be ranked. Examples of nominal data include categorizations or classifications, e.g., color of items (red *vs* blue marbles), lithology of rocks (sandstone, limestone, granite, etc.), and similar categorical data.

Hardness	Mineral	Chemical Formula
1	Talc	$Mg_3Si_4O_{10}(OH)_2$
2	Gypsum	$CaSO_4 \cdot 2H_2O$
3	Calcite	$CaCO_3$
4	Fluorite	CaF_2
5	Apatite	$Ca_5(PO_4)_3(OH^-,Cl^-,F^-)$
6	Feldspar	$KAlSi_3O_8$
7	Quartz	SiO_2
8	Topaz	$Al_2SiO_4(OH^-,F^-)_2$
9	Corundum	Al_2O_3
10	Diamond	C

Table 1.1: Traditional Moh's hardness scale.

We will see in the following chapters that discrete data will often require specialized handling and testing.

Continuous variables are those that have an uninterrupted range of possible values (i.e., with no breaks). Consequently, they have an infinite number of possible values over a given range. Our instruments, however, necessarily have finite precision and thus yield a finite number of recorded values. Examples of continuous data include:

- Seafloor depths, wingspans of birds, weights of specimens, and thickness of sedimentary layers.
- Fault strikes, directions of wind and ocean currents.
- The Earth's geopotential fields, temperature anomalies, and dimensions of objects.
- Percentages of components (such data, which are closed and forced to a constant sum, sometimes require special care and attention).

In addition, much data of interest to scientists and engineers vary as a function of *time* and/or *space* (the independent variables). Since time and space vary continuously themselves, our discrete or continuous variables will most often vary continuously as a function of one or more of these independent variables. Such data represent continuous *time series* data. They may also be referred to as *signals*, *traces*, *records*, and other names. We may further subdivide continuous data into sub-categories:

- *Ratio scale* data: These data have a fixed zero point (e.g., weights, temperatures in degrees Kelvin).
- *Interval scale* data: These data have an arbitrary zero point (e.g., temperatures in degrees Celsius or Fahrenheit).
- *Closed* data: These data are forced to attain a constant sum (e.g., percentages).
- *Directional* data: These data have components (e.g., vectors or orientations in two or higher dimensions).

Data can also be classified according to how they are recorded for use: Analog signals are those signals which have been recorded continuously (even though one might argue that this is impossible due to instrument limitations). Discrete data are those which have been recorded at discrete intervals of the independent variable. In either case, data must be discretized before they can be analyzed by a computer. Consequently, all data which are represented in computers are necessarily discrete.

1.1.2 Data limits

For a variety of reasons, such as lack of time or funds, our data tend to be limited in one or more ways. In particular, limits typically will apply to three important aspects of any data set:

Domain : No phenomena can be observed over all time or over all space, hence data have a limited *domain*. The domain may be one-dimensional for scalar quantities and N-dimensional for data in N dimensions (e.g., a spatial vector data set such as the Earth's magnetic field is three-dimensional).

Range : It is equally true that no measuring technique can record (or transmit) values that are arbitrarily large or small. The lower limit on very small quantities is often set by the noise level of the measuring instrument (because matter is quantized, all instruments will have internal noise). The *Dynamic Range* (DR) is the range over which the data can be measured (or exists). This range is usually given on a logarithmic scale measured in *decibels* (dB), i.e.,

$$DR = 10\log_{10}\left(\frac{\text{maximum power}}{\text{minimum power}}\right) \text{ dB}. \tag{1.1}$$

One can see that every time DR increases by 10, the ratio of the maximum to minimum values has increased by an order of magnitude. Strictly speaking, (1.1) is to be applied to data represented as a *power* measurement (a squared quantity, such as variance or square of the signal amplitude). If the data instead represent *amplitudes*, then the formula should be

$$DR = 20\log_{10}\left(\frac{\text{maximum amplitude}}{\text{minimum amplitude}}\right) \text{ dB}. \tag{1.2}$$

For example, if the ratio between highest and lowest voltage (or current) is 10, then $DR = 20\log_{10}(10) = 20$ dB. If these same data were represented in watts (power), and since power is proportional to the voltage squared, the ratio would be 100, and $DR = 10\log_{10}(100) = 20$ dB. Thus, regardless of the manner in which we express our data we get the same result (provided we are careful). In most cases though (except for electrical data) the first formula given is the one to be used (so the data extrema should be expressed as powers). Few instruments have a dynamic range greater than 100 dB. In any case, because of the limited range and domain of data, any data set, say $f(t)$, can be enclosed as

$$t_0 < t < t_1 \text{ and } |f(t)| < M, \tag{1.3}$$

for some constant M. Such functions are always integrable and manageable and can be subjected to further analysis without any special treatment.

Frequency : Finally, most measuring devices cannot respond instantly to sudden change. The resulting data are thus said to be *band limited*. This means the data will not contain frequency information higher (or lower) than the signal representing the fastest (or slowest) response of the recording device.

1.1.3 Noise

In almost all cases, real data contain information other than what is strictly desired ("desired" is a key word here since we all know the saying that "one scientist's signal is another scientist's noise"). We say such data consist of *samples* of *random variables*. This statement does not mean that the data are totally random, but instead imply that the value of any future observation can only be predicted in a *probabilistic* sense — it cannot be exactly predicted as is the case for a *deterministic* variable, which is completely predictable by a known law. In other words, because of inherent variability in natural systems and the imprecision of experiments or measuring devices, if we were to give an instrument the same input at two different times we would likely get two different outputs due to the difference in *noise* at the two different times. In analyzing data, one must never overlook or ignore the role of noise, as understanding the noise level provides the key to how subtle signals we can reliably resolve in our data. One of the main goals in data analysis is to detect the signal in the presence of noise or to reduce the degree of noise contamination. We therefore try to enhance the *signal-to-noise ratio* (S/N), defined (in decibels) as

$$\text{S/N} = 10\log_{10}\left(\frac{\text{power of signal}}{\text{power of noise}}\right) \text{ dB}. \tag{1.4}$$

Minimizing the influence of noise during data acquisition and stacking co-registered data are some of the approaches used to enhance the S/N ratio.

1.1.4 Accuracy versus precision

An *accurate* measurement is one that is very close to the true value of the phenomenon we are observing. A *precise* measurement is one that has very little scatter: Repeat measurements will give more or less the same values (Figure 1.2). If the measured data have high precision but poor accuracy, one may suspect that a systematic *bias* has been introduced, e.g., we are using an instrument whose zero position has not been calibrated properly. If we do not know the expected value of a phenomenon but are trying to determine just that, it is obviously better to have accurate observations with poor precision than very precise, but inaccurate values, since the former will give a correct, but imprecise estimate while the latter will give a wrong, but very precise result! Fortunately, we will often have a good idea of what a result should be and can use that prior knowledge to detect any bias in the measurements.

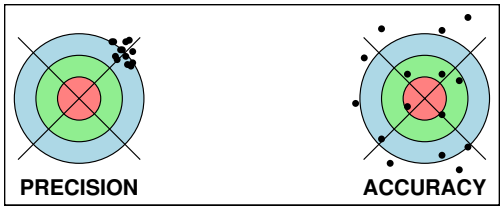


Figure 1.2: Precision is a measure of repeatability, while accuracy refers to how close the average observed value is to the "true" value.

1.1.5 Randomness

Phenomena, and hence the data that represent them, may range from those that are completely determined by a natural law (*deterministic* data) to those that seem to have no inherent structure or patterns (*chaotic* or *random*). Most phenomena, and hence the data we collect, lie in-between these two extremes. Here, there is a more-or-less clear pattern or structure to the data, but each individual measurement also exhibits a random component that makes it impossible to predict the outcome of a future observation with 100% certainty. Such data are called *probabilistic* or *stochastic*. We will often fit simple, deterministic models to such data and hope that the residuals reflect insignificant, uncorrelated noise. We will also want to know if that hope is warranted using specific statistical tests.

1.1.6 Analysis

Analysis means to separate something into its fundamental components in order to identify, interpret and/or study the underlying structure. In order to do this properly we should have some idea of what the components are likely to be. Therefore, we should have some concept of a model of the data in mind (whether this is a conceptual, physical, intuitive, or some other type of model is not important). We essentially need guidelines to aid our analysis. For example, it is not a good idea to take a data set and simply compute its Fourier series because you happen to know something about Fourier analysis. One needs to have an idea as to what to look for in the data. Often, this knowledge will grow with a set of well planned ongoing analyses, whose techniques and uses are the essence of this book.

The following steps are parts of most data analysis schemes:

1. *Collect* or obtain the data.
2. Perform *exploratory data analysis*.
3. *Reduce* the data to a few quantities (*statistics*) that *summarize* their bulk properties.
4. Compare data to various *hypotheses* using statistical *tests*.

We will briefly discuss step (1) while reviewing error analysis, which is the study and evaluation of uncertainty in measurements and how these propagate into our final statistical estimates. The main point of step (2) is to familiarize ourselves with the data set and its major structure. This acquaintance is almost always best done by graphing the data. Only an inexperienced analyst will use a sophisticated "black-box" technique to compile statistics from data and accept the validity of these statistics without actually looking at the data. Step (3) will usually include a model

(simple or complicated) where the purpose is to extract a few representative parameters out of possibly millions of data points. These statistics can then be used in various tests (4) to help us decide which hypothesis the data favor, or rather *not* favor. That is the curse of statistics: You can never prove anything, just disprove! By disproving all possible hypotheses other than your pet theory, other scientists will eventually either grudgingly accept your views or they die of old age and then your theory will be accepted! Hence, persistence and longevity are important characteristics of a successful researcher. Joking aside, it is important to listen to your data as it is disturbingly easy to be convinced that your pet model or theory is right, data be damned. This phenomenon is called *conviction bias*.

1.2 Exploratory Data Analysis

As mentioned, the main objective of exploratory data analysis (EDA) is to familiarize yourself with your observations and determine their main structure. Since simply staring at a table or computer printout of numbers will eventually lead to premature blindness or debilitating insanity, there are several standard techniques that we will classify under the broad EDA heading:

1. Scatter plots — show it all.
2. Schematic plots — simplify the sample.
3. Histograms — explore the distribution.
4. "Smoothing" of data — reduce the noise.
5. Residual plots — determine the trends.

We will briefly discuss each of these five categories of exploratory techniques. For a complete treatment on EDA, see John Tukey's classic *Exploratory Data Analysis* book; the reference is listed in the Preface.

1.2.1 Scatter plots

Figure 1.3: Scatter plots showing all individual data points — the "raw" data — are invaluable in identifying outlying data points and other potential problems.

If practical, consider plotting every individual data value on a single graph. Such "scatter" plots show graphically the correlation between points, the orientation of the data, bad outliers, and the spread of clusters (e.g., Figure 1.3). We will later (in Chapter 3) provide a more rigorous definition for what correlation is; at this stage it is just a visual appearance of a trend. Scatter plots are particularly useful in two dimensions, but even three-dimensional data are fairly easy to visualize. For higher dimensions we may choose to view *projections* of the data onto lower-dimensional spaces and thus examine only 2–3 components at the time.

1.2.2 Schematic plots

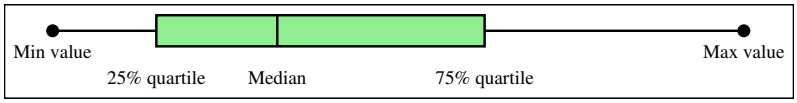


Figure 1.4: An example of a "box-and-whisker" diagram. The five quartiles give a visual representation of how one-dimensional data (or a single component of higher-dimensional data) are distributed.

The main objective here is to summarize a one-dimensional data distribution using a simple graph. One very common method is the *box-and-whisker* diagram, which graphically presents five informative measures of the sample. These five quantities are the *range* of the data (represented by the minimum and maximum values), the *median* (a line at the half-way point), and the *concentration* (represented by the 25% and 75% *quartiles*) of a data distribution. Schematically, these statistics can be conveniently illustrated as shown in Figure 1.4.

As an example of the successful use of box-and-whisker diagrams, we shall return to the winter of 1893–94 when Lord Rayleigh was investigating the density of nitrogen from various sources. Some of his previous experiments had indicated that there seemed to be a discrepancy between the densities of nitrogen produced by removing the oxygen from a sample of air and the nitrogen produced by decomposition of different chemical compounds. The 1893–94 results clearly established this difference and prompted further investigations into the composition of air, which eventually led him to the discovery of the inert gas *argon*. Part of his success in convincing his peers has been attributed to his use of box-and-whisker diagrams to emphasize the difference between the two data sets he was investigating. We will use Lord Rayleigh's data (reproduced in Table 1.2) to make a scatter plot and two schematic plots: The already mentioned box-and-whisker diagram and the *bar* graph.

Date	Origin	Purifying Agent	Weight
29 Nov. 1893	NO	Hot iron	2.30143
5 Dec. 1893	”	”	2.29816
6 Dec. 1893	”	”	2.30182
8 Dec. 1893	”	”	2.29890
12 Dec. 1893	Air	”	2.31017
14 Dec. 1893	”	”	2.30986
19 Dec. 1893	”	”	2.31010
22 Dec. 1893	”	”	2.31001
26 Dec. 1893	N_2O	”	2.29889
28 Dec. 1893	”	”	2.29940
9 Jan. 1894	NH_4NO_2	”	2.29849
13 Jan. 1894	”	”	2.29889
27 Jan. 1894	Air	Ferrous hydrate	2.31024
30 Jan. 1894	”	”	2.31030
1 Feb. 1894	”	”	2.31028

Table 1.2: Lord Rayleigh's density measurements of nitrogen (Lord Rayleigh, On an anomaly encountered in determinations of the density of nitrogen gas, *Proc. Roy. Soc. Lond., 55*, 340–344, 1894).

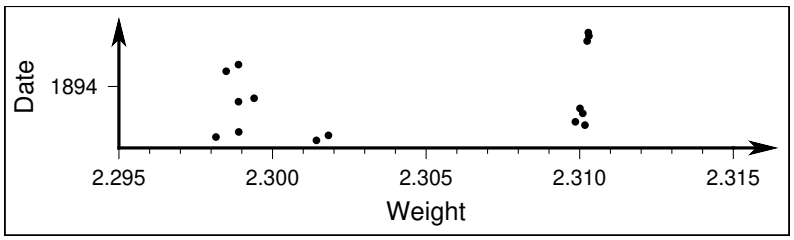


Figure 1.5: Scatter plot of Lord Rayleigh's data on nitrogen. The plot reveals distinct groups.

We will first look at all the data using a scatter plot. It may look something like Figure 1.5. It is immediately clear that we probably have two different data groupings here. Note that this is only apparent if you plot the "raw" data points. Plotting all the values as one data set using the box-and-whisker approach would result in a confusing graph (Figure 1.6), which tells us very little that is meaningful. Even the median, traditionally a stable indicator of "average" value, is questionable since it lies between the two data clusters. Clearly, it is important to find out if our data consist of a single population or if it contains a mix of two (or even more) data components.

Fortunately, in this example we know how to separate the two data sets based on their origins. It appears that we are better off plotting the data sets separately instead of treating them as a single population. However, the choice of

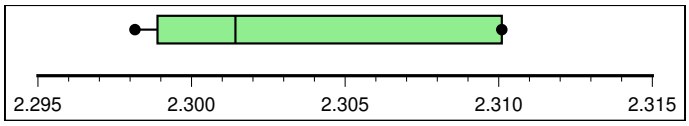


Figure 1.6: Schematic box-and-whisker plot of Lord Rayleigh's nitrogen data. While representing the entire sample the graph obscures the pattern so clearly revealed by the scatter plot.

diagram is also important. Consider a simple bar graph (here indicating the average value) summarizing the data given in Table 1.2. It would simply look as shown in Figure 1.7. In this presentation, it is just barely visible that the weight

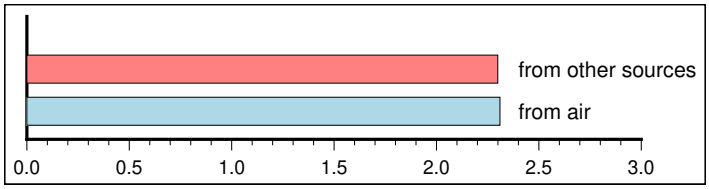


Figure 1.7: Bar graph of the average values from Lord Rayleigh's nitrogen data. Because the average values are very similar the two bars look very similar and do not tell us much.

of "nitrogen" extracted from the air is slightly heavier than nitrogen extracted from the chemical compounds. Given the way they are shown, the data present no clear indication that the two data sets are *significantly* different. Part of the problem here is the fact that we are drawing the bars from an origin at zero, whereas all the variation actually takes place in the 2.29–2.32 interval (again evident from the scatter plot in Figure 1.5). By expanding the scale and choosing a box-and-whisker plot we concentrate on the differences and produce an illustration as the one shown in Figure 1.8.

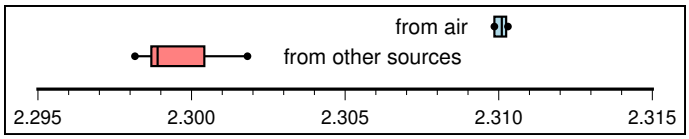


Figure 1.8: Separate box-and-whisker diagrams of nitrogen weight given in Table 1.2 based on origin. Their separation and extent clearly convey the finding of two separate sources.

It is obvious that the second box-and-whisker diagram allows a clearer interpretation than the bar graph. The diagram also benefits from the stretched scale, which highlights the different ranges of the data groupings. In Lord Rayleigh's situation the convincing diagram was accepted as strong evidence for the existence of a new element (later determined to be *argon*), and a Nobel prize in physics followed in 1904.

1.2.3 Histograms

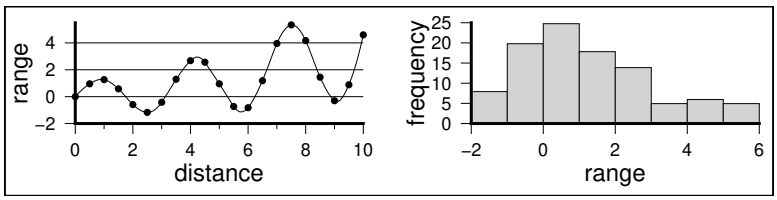


Figure 1.9: A data set, here a function of distance, can be converted to a histogram by counting the frequency of occurrence within each sub-range. The histogram only uses the y-values of the (x, y) points shown in the left diagram.

Histograms convey an accurate impression of the data *distribution* even if it is multimodal. One simply breaks the data range into equidistant sub-ranges and plots the frequency or occurrences for each range (e.g., Figure 1.9).

Obviously, the width of the sub-range determines the level of detail you will see in the final histogram. Because of this, it is usually a good idea to plot the discrete values as individual points since the "binning" throws away some information about the distribution. If the amount of data is moderate, then one can plot the individual values inside the histogram bars. Furthermore, you should explore how the *shape* of the distribution changes as you vary the *bin width*. Clearly, as the histogram bin width approaches zero you will end up with one point (or none) per bin, while at the other extreme (a very wide bin width) you will simply have a single bin with all your data. Try to select a width that yields a representative distribution, but at the same time try to understand what is going on when your widths give different shapes.

1.2.4 Smoothing

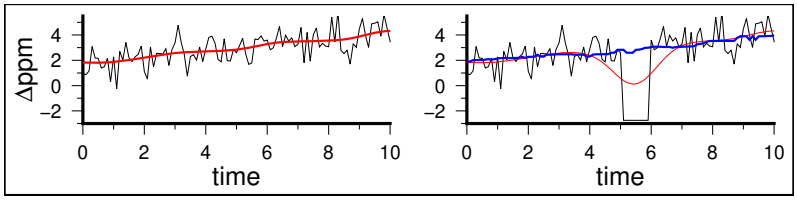


Figure 1.10: Smoothing of data is usually done by filtering. The left panel shows a noisy data set and a smooth Gaussian filtered curve (red), while the right panel shows the same data with a glitch between $x = 5$ and $x = 6$. For such data the Gaussian filter is unduly influenced by the outlying data whereas a median filter (blue) is much more tolerant.

The purpose of *smoothing* is to highlight the general trend of the data and suppress high-frequency oscillations. We will briefly look at two types of smoothing: The *median* filter and the *Hanning* filter. The median filter is typically a three-point filter and simply replaces a point with the median value of the point and its two immediate neighbors. The filter then shifts one step further to the right and the process repeats. This technique is very efficient at removing isolated spikes or *outliers* in the data since the bad points will be completely ignored as they will never occupy the median position, unless they appear in groups of two or more (in that case a wider median filter, say 5-point, would be required). In contrast, the Hanning filter is simply a moving average of three points where the center point is given twice the weight of the neighbor points, i.e.,

$$y'_i = \frac{y_{i-1} + 2y_i + y_{i+1}}{4}. \tag{1.5}$$

Note that while such a filter works well for data that have random high-frequency noise, it gives disastrous results for spiky data since the outliers are averaged into the filtered value and never simply ignored. For noisy data with occasional outliers one might consider running the data first through a median filter, followed by a treatment of the Hanning filter. In building automated data procedures one should always have the worst-case scenario in mind and seek to protect the analysis by preprocessing with a median filter. Figure 1.10 illustrates the use of simple smoothing with Gaussian filters, with or without protection from outliers by a median filter. We will have more to say about filtering in Chapter 8.

1.2.5 Residual plots

We can always make the assumption that our data can be decomposed into two parts: A smooth trend plus noisier residuals. This separation forms the basis for *regional-residual* analysis. The simplest trend (or regional) is just a straight line. One can easily define such a line by picking two representative points (x_1, y_1) and (x_2, y_2) and then compute the trend as

$$T(x) = y_1 + \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1). \tag{1.6}$$

We then can form residuals $r_i = y_i - T(x_i)$ (in Chapter 6 we will learn more rigorous ways to find linear trends in $x - y$ data). If a significant trend still exists, one can try several standard transformations to determine the nature of the

"smooth" trend, such as the family of trends represented by

$$y^n, ..., y^1, y^{1/2}, \log(y), y^{-1/2}, y^{-1}, ..., y^{-n}. \tag{1.7}$$

The residual plot procedure we will follow is:

1. Take $\log(y)$ of the data and plot the values.
2. If the line is concave then choose a transformation closer to y^{-n}.
3. If the data is convex then choose a transformation closer to y^n.

While approximate, this method gives you a feel for how the data vary. Note that if your y-values contain or straddle zero then you cannot explore a logarithmic relationship for that axis.

We will conclude this section with a quote from J. Tukey's book "Exploratory Data Analysis", which is worth contemplating:

> *"Many people would think that plotting y against x for simple data is something requiring little thought. If one wishes to learn but little, it is true that a very little thought is enough, but if we want to learn more, we must think more."*

The moral of it all is: *Always* plot your data. *Always*! *Never* trust output from software performing statistical analyses without comparing the results to your data. Very often, such software implements statistical methods that are based on certain assumptions about the data distribution, which may or may not be appropriate in your case. Plotting your data may be the easiest thing to do (for dimensionality less than or equal to three) or it may be quite challenging. Exploring sub-spaces of just a few dimensions is a simple starting point, while more sophisticated methods will examine the dimensionality of the data to see if some dimensions simply provide redundant information.

1.3 Problems for Chapter 1

Note: Problems that ask you to comment on what you observe in the data do not require any special knowledge about geology, seismology, or any other discipline.

Problem 1.1. Perform exploratory data analysis on the data set *sludge.txt*, which contains annual average concentrations of total phosphorous in municipal sewage sludge from an unnamed city, given in percent of dry weight solids. Make a few EDA plots to determine any trends or patterns captured by the data.

Problem 1.2. Perform exploratory data analysis on the data set *hypso.txt*, which contains topography bins in km and frequency in percent, in other words the *hypsometric* curve for the Earth. Your EDA should include a few graphs such as a scatter plot, a box-and-whisker, and a histogram). Based on the graphical results, make some preliminary conclusions about the data set (e.g., are there outliers? Are the distributions unimodal? Asymmetrical?).

Problem 1.3. Perform exploratory data analysis on the data set *porosity.txt* (porosity in percent for a few sandstone samples). Again, generate several typical EDA plots and reach some preliminary conclusions about the data (e.g., any outliers? Unimodal vs. multimodal? Asymmetrical vs. symmetrical?).

Problem 1.4. Scientists obtained free-air gravity anomalies along the track of the *R/V Conrad* cruise 2308, which surveyed an area around Oahu in the Hawaiian Islands. Ideally, when the ship reoccupies the same location the free-air anomaly should be the same, but instrument drift, inadequate correction for ship motion and erroneous navigation data all lead to discrepancies called *cross-over errors* (COE); here *c2308_xover.txt* contains these data. This mismatch in values at the intersections of the track gives information on the uncertainty in the measurements. Examine the CEOs using both histograms and box-and-whisker techniques.

Problem 1.5. Make both a box-and-whisker diagram and a histogram of the bathymetry depths in *pac_bathy.txt* (depths in m for a 1° radius region in the Western Pacific). Try different binning intervals and discuss the important aspects of the distribution.

Problem 1.6. Examine the variations in the magnetic field (in nT) at the Honolulu station for September, 1998 reproduced in *honmag_1998_09.txt*. Show a scatter plot and a box-and-whisker diagram of the total field (HONF).

Problem 1.7. Perform EDA on the data given in *seismicity.txt*, which contains records with longitude, latitude, depth (in km), and magnitude for significant earthquake hypocenters from the Tonga-Kermadec region. Try various presentations and decide on two plots that best highlight the features of the data set. Provide commentary on what you see as significant for each plot and especially note any outliers or peculiarities about the data set (even if they do not make it onto the two plots you select).

Problem 1.8. Repeat the exercise in Problem 1.7 on the data set *hi_ages.txt*, which contains a list of longitude, latitude, radiometric age (in Myr), and distance from Kilauea (in km) along the Hawaiian seamount chain.

Problem 1.9. Repeat the exercise in Problem 1.7 on the data set *v3312.txt*, which documents the distance (in km) and depth (in m) along the track of *R/V Vema* cruise 3312 from Japan to Guam.

Problem 1.10. Explore the record of Mississippi river daily discharge (m^3/s) from 1930–1941 listed in *mississippi.txt*. Select two graphs that tell the most complete story about this data set.

Problem 1.11. The GISP-2 ice core $\delta^{18}O$ variations in parts per thousand (ppt) from the last $\sim$10,000 years can be found in the file *icecore.txt*. Select two graphs that reveal the most about this data set. Note the times are relative to an origin at 1950.

Problem 1.12. The data set *trend.txt* contains (x, y) pairs exhibiting a trend. Use the residual plot technique to determine the likely trend exhibited by the majority of the data. Are there any outliers?

Problem 1.13. The Earth's magnetic field is known to have reversed direction over geologic time. The file *GK2007.txt* contains data from the Gee and Kent (2007) magnetic time scale. It lists all normal and reversely magnetized chrons and gives the duration of each interval (in Myr). Note: Examine the last letter in the chron: "n" means normalized and "r" stands for reversed polarity.

a) Make box-and-whisker plots for the entire data set as well as for normal and reverse polarities separately.

b) Make a histogram of all intervals using a bin width of 0.1 Myr.

c) The distribution seems to have a long tail. Convert your data to logarithmic data ($\log_{10}$) and make another histogram with a log-increment of 0.1. Does the transformation reduce the tail of the data?

Chapter 2

REVIEW OF ERROR ANALYSIS

> "An error does not become truth by reason of multiplied propagation, nor does the truth become error because nobody will see it."
>
> *Mahatma Gandhi, India Independence Leader*

Error analysis is the study and evaluation of uncertainty in measurements of continuous data (discrete data may have no errors). We know that no measurement, however carefully made, can be completely free of uncertainties. Since science depends on measurements, it is crucially important to be able to evaluate these uncertainties and to keep them to a minimum. Thus errors are not mistakes and you cannot avoid them by being very careful, but you should strive to make them as small as possible. Understanding uncertainties is critical for determining the *significance* of models that you may construct to explain your data. In this chapter, we will review the basic rules that govern the reporting of measurements with uncertainties and how these uncertainties propagate when used to obtain derived quantities (such as sums and products) as well as how they affect the uncertainties in more complicated situations (such as being arguments to nonlinear functions).

2.1 Reporting Uncertainties

When reporting the value of a measurement, care should be taken to give the best possible estimate of the uncertainty or error in the measurement. Values read off scales or measured with mechanical instruments can usually be bracketed between lower and upper limits. E.g., we may know that a temperature measurement T is not less than 23° nor larger than 24°. Hence, we shall report the value of T as

$$T = 23.5 \pm 0.5^\circ, \tag{2.1}$$

or, in general, $x \pm \delta x$. For many types of measurements we can state with absolute certainty that x must be within these bounds. Unfortunately, very often we cannot make such a categorical statement. To do so, we would have to specify unreasonably large values for δx to be absolutely confident that the actual quantity lies within the stated interval. In most cases, we will lower our confidence to, say, 90% and use a smaller δx. We need more detailed knowledge of the statistical laws that govern the process of measurement for finding a suitable δx, so we will return to this issue later in this book.

2.1.1 Significant figures

In general, the last significant figure of any stated answer should be of the same order of magnitude (i.e., same decimal position) as the uncertainty. For intermediate calculations you should use one extra decimal (if calculating by hand) or all available decimals (if using calculators or computers).

Example 2–1. If a measured distance d is 6051.78 m with an uncertainty of ± 30 m, you should report the distance as $d = 6050 \pm 30$ m. Similarly, if a time t is measured to be 3 seconds and the uncertainty is given as 1 part in 100, you

should report the time as $t = 3.00 \pm 0.03$ s.

2.2 Fractional Uncertainty

While the statement $x = x_0 \pm \delta x$ indicates the *precision* of the measurement, it is clear that such a statement will have different meanings depending on the value of x_0 relative to δx. Clearly, with $\delta x = 1$ m, we imply a different precision for $x_0 = 3$ m than for $x_0 = 1000$ km. Thus, we should consider the *fractional uncertainty*, written as

$$\frac{\delta x}{|x_0|}. \tag{2.2}$$

Note that the fractional uncertainty is a *nondimensional* quantity.

Example 2–2. We wish to report the uncertainty in the length of a 3m crocodile, which (due to a shortage of available arms) we only were able to measure with a precision of 6 cm. Using fractional uncertainty, we report the uncertainty $\delta l = \pm 100 \frac{0.06}{3}\% = \pm 2\%$.

2.3 Uncertainty in Derived Quantities

It is common to obtain measurements (and estimate their uncertainties) and then use these values in expressions that lead to derived quantities. In such situations we must be careful to *propagate* the uncertainties of the initial measurements so that the derived quantities can be reported with their combined uncertainties. The answers will differ depending on whether or not the individual measurements are *independent* or *dependent* on each other.

2.3.1 Uncertainty in sums and differences

Consider the two values $x \pm \delta x$ and $y \pm \delta y$. We would like to determine the correct expressions for the uncertainties in the four derived quantities

$$s = x + y, \tag{2.3}$$

$$d = x - y, \tag{2.4}$$

$$p = x \cdot y, \tag{2.5}$$

$$q = x/y. \tag{2.6}$$

For the sum, common sense would suggest that the maximum value of s must be $s = x + y + \delta x + \delta y$, while the minimum value is $s = x + y - \delta x - \delta y$. Thus

$$s \approx (x + y) \pm (\delta x + \delta y) = s_0 \pm \delta s \tag{2.7}$$

and similarly for differences we reason that

$$d \approx (x - y) \pm (\delta x + \delta y) = d_0 \pm \delta d. \tag{2.8}$$

Here, the subscript 0 indicates the nominal or "best" value of the results. We use the approximate sign $\approx$ since we anticipate that the stated uncertainties δs and δd probably overestimate the true uncertainties in the sum and difference, respectively. Note that the uncertainties in both the sum and the difference of x and y are identical.

2.3.2 Uncertainty in products and quotients

For the product, we first use fractional uncertainties and rewrite x and y as

$$x = x_0\left(1 \pm \frac{\delta x}{|x_0|}\right),$$

$$y = y_0\left(1 \pm \frac{\delta y}{|y_0|}\right).$$

Then, the maximum value of p is

$$p_{high} = x_0\left(1 + \frac{\delta x}{|x_0|}\right) \cdot y_0\left(1 + \frac{\delta y}{|y_0|}\right), \tag{2.9}$$

which becomes

$$p_{high} = p_0\left(1 + \frac{\delta x}{|x_0|} + \frac{\delta y}{|y_0|} + \ldots\right), \tag{2.10}$$

where the higher order term proportional to $\delta x \cdot \delta y$ has been ignored. We note that the best value is

$$p_0 = x_0 \cdot y_0. \tag{2.11}$$

The minimum value is found by reversing the signs of δx and δy. Both expressions yield the uncertainty in p as

$$\frac{\delta p}{|p_0|} \approx \frac{\delta x}{|x_0|} + \frac{\delta y}{|y_0|}. \tag{2.12}$$

For quotients, the maximum value will be

$$q_{high} = \frac{x_0\left(1 + \frac{\delta x}{|x_0|}\right)}{y_0\left(1 - \frac{\delta y}{|y_0|}\right)} = q_0\frac{1+a}{1-b}. \tag{2.13}$$

Using the binomial theorem, we expand $(1-b)^{-1}$ as $1 + b + b^2 + b^3 \ldots$, hence

$$\frac{1+a}{1-b} \approx (1+a)(1+b) = 1 + a + b + ab \approx 1 + a + b, \tag{2.14}$$

where we again ignore higher-order terms in b by assuming that $(\delta x/|x_0|) \ll 1$ and $(\delta y/|y_0|) \ll 1$. Similarly, for the minimum value, find

$$\frac{1-a}{1+b} \approx (1-a)(1-b) = 1 - a - b + ab \approx 1 - (a+b). \tag{2.15}$$

It therefore follows that

$$q = q_0\left[1 \pm \left(\frac{\delta x}{|x_0|} + \frac{\delta y}{|y_0|}\right)\right] = q_0\left[1 \pm \frac{\delta q}{q_0}\right] \tag{2.16}$$

and that the fractional uncertainty for both products and quotients are the same.

2.3.3 Uncertainty for general expressions

The stated uncertainties are the maximum values possible, but we suspect these are likely to be exaggerated. Later, we will show that if we assume our errors to be *normally* distributed (i.e., we have "Gaussian" errors) and our measurements are *independent*, then a better estimate of the uncertainty in sums and differences is

$$\delta s = \delta d = \sqrt{(\delta x)^2 + (\delta y)^2}, \tag{2.17}$$

and for products and quotients it becomes

$$\frac{\delta p}{p_0} = \frac{\delta q}{q_0} = \sqrt{\left(\frac{\delta x}{x_0}\right)^2 + \left(\frac{\delta y}{y_0}\right)^2}. \tag{2.18}$$

Note that in the case $s = nx$, where n is a constant, we must use $\delta s = n\delta x$ since all the x values are the same and obviously *not independent* of each other. Similarly, the product $p = x^n$ will have the fractional uncertainty

$$\frac{\delta p}{p_0} = n\frac{\delta x}{x_0}, \tag{2.19}$$

since the (repeated) measurements are *dependent*. In conclusion, if $s = x + y + nz - u - v - mw$, then

$$\delta s = \sqrt{(\delta x)^2 + (\delta y)^2 + (n\delta z)^2 + (\delta u)^2 + (\delta v)^2 + (m\delta w)^2}. \tag{2.20}$$

Even if our assumption of independent measurements is incorrect, δs cannot exceed the ordinary sum

$$\delta s = \delta x + \delta y + n\delta z + \delta u + \delta v + m\delta w. \tag{2.21}$$

Similarly, if

$$q = \frac{x \cdot y \cdot z^n}{u \cdot v \cdot w^m}, \tag{2.22}$$

then we find the fractional uncertainty in q to be

$$\frac{\delta q}{q_0} = \sqrt{\left(\frac{\delta x}{x_0}\right)^2 + \left(\frac{\delta y}{y_0}\right)^2 + \left(n\frac{\delta z}{z_0}\right)^2 + \left(\frac{\delta u}{u_0}\right)^2 + \left(\frac{\delta v}{v_0}\right)^2 + \left(m\frac{\delta w}{w_0}\right)^2}. \tag{2.23}$$

While this is the likely error for independent data, we can confidently say that the fractional uncertainty will be less than the linear sum

$$\frac{\delta q}{q_0} = \frac{\delta x}{x_0} + \frac{\delta y}{y_0} + n\frac{\delta z}{z_0} + \frac{\delta u}{u_0} + \frac{\delta v}{v_0} + m\frac{\delta w}{w_0}. \tag{2.24}$$

Example 2–3. As most students who have taken an introductory physics lab in mechanics will know, measuring the period T of a pendulum of length ℓ allows one to estimate the acceleration of gravity as

$$g = \frac{4\pi^2 \ell}{T^2}. \tag{2.25}$$

We wish to determine the uncertainty in this estimate given measurements of ℓ and T and their uncertainties. Being independent measurements we use (2.23) and find

$$\frac{\delta g}{|g_0|} = \sqrt{\left(\frac{\delta \ell}{\ell}\right)^2 + \left(2\frac{\delta T}{T}\right)^2}, \tag{2.26}$$

since the constant 4π has no uncertainty. Given the measurements $\ell = 92.95 \pm 0.10$ cm and $T = 1.936 \pm 0.004$ s, we obtain

$$g_0 = \frac{4\pi^2 0.9295 \text{ m}}{1.936^2 \text{ s}^2} = 9.79035 \text{ m s}^{-2}. \tag{2.27}$$

We can now evaluate the fractional uncertainty as

$$\frac{\delta g}{|g_0|} = \sqrt{\left(\frac{0.1}{92.95}\right)^2 + \left(2\frac{0.004}{1.936}\right)^2} \approx 0.4\%. \tag{2.28}$$

The answer, therefore, is

$$g = 9.790 \pm 0.042 \text{ m s}^{-2}, \tag{2.29}$$

where we have only used three significant decimals.

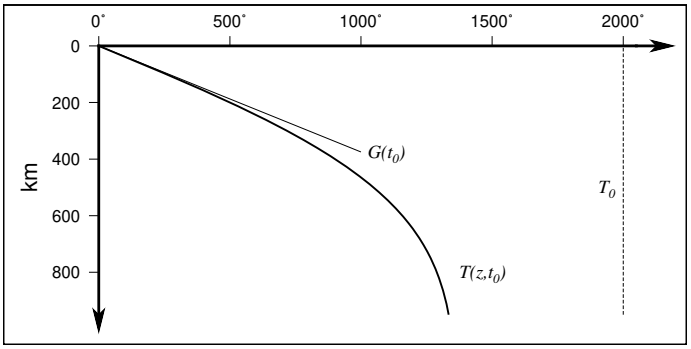


Figure 2.1: Lord Kelvin's model for the vertical temperature profile of the Earth (the *geotherm*) at a time t_0 since its initial formation at a constant temperature T_0 (dashed vertical geotherm). The tangent to the curve at the surface represents the vertical temperature gradient (G), which could be estimated from temperature measurements in mines.

Another case study revisits the debate that raged in the 19th century regarding the age of the Earth. Observing the slow process of erosion, Charles Darwin had implied that perhaps the Earth might be as old as 300 million years. Lord Kelvin, the preeminent physicist of his times, strongly objected and set out to calculate the age using the conductive cooling of the Earth (this and many other fascinating stories from the development of the geological sciences are portrayed in the classic book, *Great Geological Controversies* by A. Hallam [Oxford University Press]).

Example 2–4. Lord Kelvin assumed the whole Earth was once at a uniform temperature T_0 and had since cooled at the surface to $\sim 0°$C (as indicated in Figure 2.1). Then, the physics of heat conduction in solids dictates that

$$t_0 = \frac{T_0^2}{\pi \kappa G^2}, \tag{2.30}$$

with initial temperature $T_0 \approx 2000 \pm 200$ °K, thermal diffusivity $\kappa = 1 \pm 0.25$ mm^{2}s^{-1}, and observed near-surface temperature gradient $G = 25 \pm 5$ °K km^{-1}. Given the procedures established earlier, we first determine

$$\frac{\Delta t_0}{t_0} = \left[\left(2\frac{\Delta T_0}{T_0} \right)^2 + \left(\frac{\Delta \kappa}{\kappa} \right)^2 + \left(2\frac{\Delta G}{G} \right)^2 \right]^{\frac{1}{2}}. \tag{2.31}$$

Inserting the estimated parameters, we find

$$\frac{\Delta t_0}{t_0} = \left[\left(2\frac{200}{2000} \right)^2 + \left(\frac{0.25}{1} \right)^2 + \left(2\frac{5}{25} \right)^2 \right]^{\frac{1}{2}} \approx 51\%. \tag{2.32}$$

We evaluate Kelvin's estimate of the age of the Earth to be

$$t_0 = \frac{(2000\ °\text{K})^2}{\pi \cdot 10^{-6} \text{m}^2\text{s}^{-1} (25\ °\text{K} \cdot 10^{-3} \text{m}^{-1})^2} \approx 2 \cdot 10^{15}\ \text{s} = 65\ \text{Myr}. \tag{2.33}$$

Given the fractional uncertainty, we obtain $t_0 = 65 \pm 33$ Myr. As the debate raged on, positions hardened and Lord Kelvin continued to revise his estimates downwards, finally settling on 25 Myr. Modern science estimates that the Earth is closer to 4.6 *billion* years old. Where did Kelvin go wrong?

As a final case, let us imagine we are measuring the length of the coastline segment in Figure 2.2 using two different methods: (1) Set a compass to a fixed aperture $\Delta x = 1 \pm 0.025$ cm and march along the line counting the steps, and (2) use a digitizing tablet and sample the line approximately every $\Delta x = 1 \pm 0.1$ cm. Let us assume that it took $N = 50$ clicks or steps so the measured line length in both cases is 50 cm. What is the uncertainty in the length for the two methods? First, let us state that there will be an uncertainty for both methods that has to do with the undersampling of short-wavelength coastline "wiggles" (also, as you learn about the *fractal* nature of coastlines

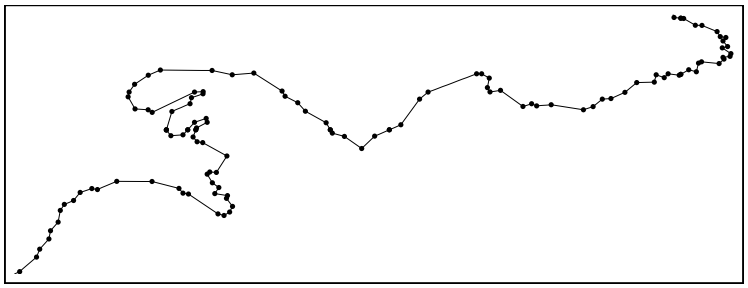


Figure 2.2: Example of a coastline segment whose length we attempt to estimate using both a compass and via digitizing.

then perhaps our simple approach here will seem a bit naive). That aside, we can see that the errors accumulate very differently. For the compass length-sum the errors are all dependent (since the aperture is fixed) and we must use the summation rule to find the uncertainty $\delta l = N \cdot 0.025$ cm $= 1.25$ cm. For the digitizing operations all the uncertainties associated with points 2 through 49 largely cancel and we are left with the uncertainty of the endpoints. Those are clearly independent and hence the uncertainty is $\delta l = (0.1^2 + 0.1^2)^{1/2}$ cm $= 0.14$ cm. The systematic errors using the compass accumulate while the errors in digitizing only affect the end-points. This discussion of digitizing errors is a bit oversimplified, but it does illustrate the difference between the two types of errors and how they accumulate.

2.3.4 Uncertainty in a function

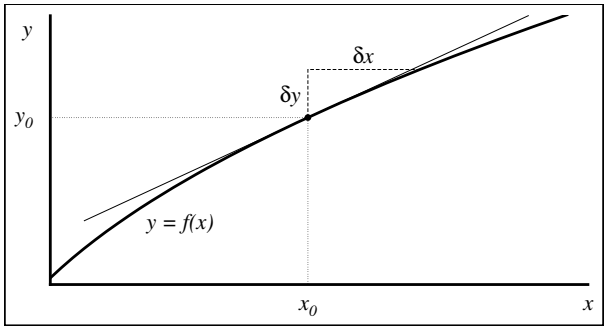


Figure 2.3: As δx becomes very small, the derivative of any well-behaved function can be approximated by a *tangent* at the point $(x_0, y_0 = y(x_0))$; this is the second term in Taylor's expansion.

Many solutions to scientific or engineering problems require the evaluation of functions with our uncertain measurements as arguments. If x is measured with uncertainty δx and is used to evaluate the function $y = f(x)$, then the uncertainty δy is related to the *derivative* of the function at x, i.e.,

$$\delta y = \left| \frac{df}{dx} \right|_{x_0} \cdot \delta x, \tag{2.34}$$

where the derivative is evaluated at x_0 (e.g., Figure 2.3).

Example 2–5. Let $y(x) = \cos x$ and $x = 20 \pm 3^\circ$. Following (2.34),

$$\delta y = \left| \frac{d\cos(x)}{dx} \right|_{x=20^\circ} \left(\frac{\pi}{180^\circ} \right) 3^\circ = |-\sin 20^\circ| \, \frac{3\pi}{180} = 0.342 \cdot 0.0524 \approx 0.02, \tag{2.35}$$

where we have converted the angle from degrees to radians (why?). The final answer then becomes

$$y = \cos(x) = 0.94 \pm 0.02. \tag{2.36}$$

Finally, for a function of multiple variables, $f(x,...,z)$, we extend our analysis to find

$$\delta f = \sqrt{\left(\frac{\partial f}{\partial x}\delta x\right)^2 + \cdots + \left(\frac{\partial f}{\partial z}\delta z\right)^2} \tag{2.37}$$

when $x,...,z$ are all random and independent. As before, δf cannot exceed the ordinary sum

$$\delta f \leq \left|\frac{\partial f}{\partial x}\right|\delta x + \cdots + \left|\frac{\partial f}{\partial z}\right|\delta z, \tag{2.38}$$

which is suitable for dependent measurements.

Example 2–6. Consider the spherical function

$$f(r,\theta,\phi) = \frac{1}{2}r^2\cos^2\theta\sin\phi. \tag{2.39}$$

We measured the parameters and found $r = 10 \pm 0.1, \theta = 60 \pm 1^\circ$, and $\phi = 10 \pm 1^\circ$. Using (2.37), the uncertainty in the evaluated expression in (2.39) is found as

$$\delta f = \sqrt{(r\cos^2\theta\sin\phi\delta r)^2 + (-r^2\cos\theta\sin\theta\sin\phi\delta\theta)^2 + \left(\frac{1}{2}r^2\cos^2\theta\cos\phi\delta\phi\right)^2}, \tag{2.40}$$

which means our final estimate of $f(r,\theta,\phi)$ evaluates to

$$f(r,\theta,\phi) = 2.17 \pm 0.26. \tag{2.41}$$

While the simple expressions for uncertainty in a derived quantity (2.21 or 2.23) generally apply, one must be careful with expressions where one or more of the measurements appear in different *subgroups* within the expression. In such cases one must treat the entire expression as a function of several variables and apply the general expression given in (2.37), as our next example illustrates.

Example 2–7. Given the multivariate function

$$f(a,b) = 2ab^2 + \pi b + 1, \tag{2.42}$$

we wish to find its value and uncertainty for $a = 0.3 \pm 0.02$ and $b = 1 \pm 0.01$. Because we have more than one term that depends on b we cannot easily employ the rules in (2.21) and (2.23) but must use (2.37) instead. We first evaluate $f(0.3,1)$ to be 5.7416... and next evaluate the uncertainty using (2.37):

$$\delta f = \sqrt{(2b^2\delta a)^2 + [(4ab+\pi)\delta b]^2} = 0.05903.... \tag{2.43}$$

Consequently, this yields a final estimate of

$$f = 5.74 \pm 0.06, \tag{2.44}$$

where we have rounded the answer to two decimals only.

2.4 Problems for Chapter 2

Problem 2.1. The distance to a building is estimated from a map to be 2550 ± 25 m. With a theodolite, a student determines the angle between the horizontal plane and the building roof to be $1°21' \pm 1'$. What is the height of the building and the uncertainty in this estimate?

Problem 2.2. The area of a cornfield is being estimated from aerial photographs. Because of an unknown stretching factor the uncertainty in linear distances has been set to 1% and hence uncertainties are *dependent*. What is the area and uncertainty of a field with measured dimensions 235.5 m by 115.6 m?

Problem 2.3. Some neighborhood kid is driving his scooter way too fast while passing your house. Annoyed, you set out to apply basic high-school physics in order to compute his speed: You measure a fixed length section in front of your house and, while hiding in the bushes, use a stop-watch to time how long he takes to cover that distance. Your old measuring tape reports 18.20 ± 0.05 m and you clock him covering that distance in 0.82 ± 0.10 s. In your passive-aggressive letter to the kid's parents, what is the speed and uncertainty that you report?

Problem 2.4. From compositional data you infer that two volcanic rock samples represent the initial and final pulse of activity, respectively. Your two samples have been radiometrically dated at 25.53 ± 0.1 Ma and 29.66 ± 0.2 Ma, respectively. What is the likely duration of volcanic activity (including the error)?

Problem 2.5. You have carefully measured the densities of an exposed ore body and the surrounding bed rock and determined $\rho_o = 3.15 \pm 0.05$ g cm^{-3} and $\rho_r = 2.67 \pm 0.05$ g cm^{-3}, respectively. A gravimetric survey over the region will be sensitive to variations in lateral density only, in other words the density contrast between the ore and the host rock. Assuming your samples are independent, what is the density contrast and its uncertainty?

Problem 2.6. Assigned to a crummy summer internship on Mars, you are tasked with the low-tech job of measuring the perimeters of circular craters using a $69.99 measuring wheel from Home Depot. You obtain the circumferences of two craters before you realize you are in fact the subject of a psychological experiment. Although upset, you nevertheless decide to complete the estimates for the two craters. The measuring wheel reports 12,311 and 9,045 clicks, respectively, and the manufacturer says each click represents a distance increment of 25 cm (i.e., the circumference of the measuring wheel). Previous studies suggest that this method is accurate to 1% of the circumference. What is the area (and its uncertainty) of each crater in m^2?

Problem 2.7. The Bouguer equation for gravity due to a constant thickness slab is given by $g = 2\pi\rho\gamma h$, where $\gamma = 6.6738 \cdot 10^{-11}$ m^3 kg^{-1} s^{-2} is the universal gravitational constant with fractional uncertainty $1.2 \cdot 10^{-4}$, $\rho = 2850$ g cm^{-3} is the slab density, and $h = 230 \pm 1$ m is the thickness of the slab.

a) To achieve a precision of 1% in the Bouguer calculation, how well do you need to know the density value (i.e., what is the maximum uncertainty you can tolerate)?

b) Report the Bouguer value for the slab, including its uncertainty, in mGal ($= 10^{-5}$m s^{-2}).

Problem 2.8. The subsidence of young (< 80 Myr) oceanic crust due to lithospheric cooling has been shown to follow approximately a linear $\sqrt{\text{age}}$ relationship, given by

$$z = z_r + \frac{2\rho_m \alpha_v T_m}{(\rho_m - \rho_w)} \sqrt{\frac{\kappa t}{\pi}}.$$

Given estimates of thermal diffusivity $\kappa = 1.00 \pm 0.04$ mm$^2 s^{-1}$, water density $\rho_w = 1.027 \pm 0.001$ g cm^{-3}, mantle density $\rho_m = 3.30 \pm 0.01$ g cm^{-3}, volumetric thermal expansion coefficient $\alpha_v = (3.00 \pm 0.02) \cdot 10^{-5}$ °K^{-1}, average ridge depth $z_r = 2500 \pm 200$ m, and mantle temperature $T_m = 1300 \pm 25$ °K, determine the predicted depth and its uncertainty for a location where rocks of age $t = 29.7 \pm 0.5$ Myr were recovered. Which term dominates the final uncertainty?

Problem 2.9. The flexural parameter α encountered in studies of elastic plate flexure is given by

$$\alpha = \left[\frac{4D}{(\rho_m - \rho_w)g}\right]^{\frac{1}{4}},$$

where D is the flexural rigidity and is related to the *elastic plate thickness*, h, via

$$D = \frac{Eh^3}{12(1-\nu^2)},$$

where E and ν are Young's modulus and Poisson's ratio, respectively. It can be shown that the distance from a 2-D line load (approximating an island chain, perhaps) to the maximum peripheral uplift (the "bulge") is given by $x_b = \pi\alpha$. Given the measurements $\rho_m = 3300 \pm 50$ kg m^{-3}, $E = 70 \pm 7$ GPa, $\nu = 0.25 \pm 0.01$, assuming no uncertainty in $g = 9.81$ m s^{-2} and $\rho_w = 1027$ kg m^{-3}, and observing $x_b = 252 \pm 10$ km, what is the corresponding elastic plate thickness (and its uncertainty)?

Problem 2.10. When collecting gravity measurements on-board a moving platform, such as a ship or aircraft, one must account for the *Eötvös effect*, which will reduce or increase the apparent gravity depending on the moving platform's heading (α), latitude (θ) and speed (v). This effect (in mGal) is

$$E = 14.585 v \cos\theta \sin\alpha + 0.015696 v^2,$$

where speed v is given in m/s.

a) If the ship is at the equator and moving at 10.0 ± 0.5 knots with heading $\alpha = 35 \pm 2^\circ$, what is the Eötvös effect and its uncertainty ? (You can assume α and v have independent errors and that the uncertainty in latitude is negligible). Hint: Consider E to be the function $E(v,\alpha)$ and use the rule for uncertainty in a function (i.e., 2.37).

b) Consider an aero-gravity survey using a plane capable of flying at 200 knots. We wish to run parallel lines in a certain direction α such that the *changes* in E with small changes in v are minimized. Find the optimal survey line orientation. You may again assume that the uncertainty in latitude is negligible and that we are flying at a latitude of 45 degrees north.

Chapter 3

BASIC STATISTICAL CONCEPTS

"The most important questions of life are, for the most part, really only problems of probability."

Pierre Simon de Laplace, Mathematician

Probability is mostly organized common sense. However, being able to be specific about what probability is enables us to more accurately calculate probabilities and to employ theoretical statistical distributions to address confidence limits on data-derived quantities.

3.1 Probability Basics

In data analysis and hypothesis testing we are concerned with separating the probable from the possible. First, let us have a look at possibilities. In many situations we can either list all the possibilities or say how many such outcomes there are. In evaluating possibilities, we are often concerned with finding all the possible choices that are offered. Studying these choices leads us to the "multiplication of choices" rule:

> If a choice consists of k steps, of which the first can be made in n_1 ways and the k^{th} in n_k ways, the total number of choices is $\Pi n_i, i = 1,k$.

This can often be seen most clearly with a tree diagram (Figure 3.1). The number of choices here are $3 \times 4 = 12$.

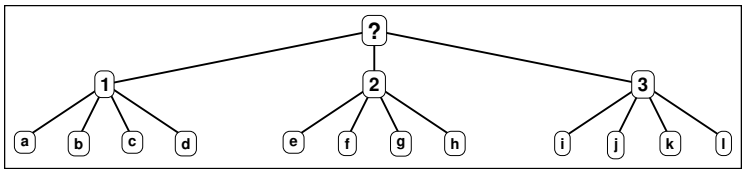


Figure 3.1: Tree diagram for illustrating all possible choices.

3.1.1 Permutations

How many ways can we arrange r objects selected from a set of n distinct objects? This question applies to numerous statistical and probabilistic situations. We will first consider a simple example.

Example 3–1. We have a tray with 20 water samples. How many ways can you select three samples from the 20? The first sample can be any of 20, the second will be any of the remaining 19, while the third is one of the remaining 18. The total ways must therefore be $20 \times 19 \times 18 = 6840$.

We can write the number of choices as $20 \times (20-1) \times (20-2)$, and by induction we find

$$\text{ways} = n(n-1)(n-2)\dots(n-r+1) = {}_nP_r. \tag{3.1}$$

It is convenient to introduce the factorial $n!$, defined as

$$n! = \prod_{i=1}^{n} i. \tag{3.2}$$

For convenience, we also define 0! to equal 1. We can then rewrite (3.1) as

$${}_nP_r = \frac{n(n-1)(n-2)\dots(n-r+1)(n-r)(n-r-1)\dots 1}{(n-r)(n-r-1)\dots 1} = \frac{n!}{(n-r)!}. \tag{3.3}$$

This quantity is called the number of *permutations* of r objects selected from a set of n distinct objects.

Example 3–2. We wish to determine how many different hands one can be dealt in a game of poker. With $n = 52$ (total number of cards in the deck) and $r = 5$ (number of cards in a hand), we find

$${}_{52}P_5 = \frac{52!}{(52-5)!} = \frac{52!}{47!} = 48 \cdot 49 \cdot 50 \cdot 51 \cdot 52 = 3 \cdot 10^8. \tag{3.4}$$

However, this calculation assumes that the *order* in which you receive the cards is important.

3.1.2 Combinations

In many situations we do not care about the exact ordering of the r objects, i.e., *abc* is the same choice as *acb* for our purpose. In general, r objects can be arranged in $r!$ different ways (${}_rP_r = r!$). Since we are only concerned about *which* r objects have been selected and not their order, we can use ${}_nP_r$ but must now normalize the result by $r!$, i.e.,

$${}_nC_r = \frac{{}_nP_r}{r!} = \frac{n!}{r!(n-r)!} = \binom{n}{r}. \tag{3.5}$$

The quantity ${}_nC_r$ is called the number of *combinations*, and the factors $\binom{n}{r}$ are called the *binomial coefficients*. After picking the r objects, $n-r$ objects are left, so consequently there are as many ways of selecting $n-r$ objects from n as there are of selecting r objects, i.e.,

$$\binom{n}{r} = \binom{n}{n-r}. \tag{3.6}$$

Example 3–3. How many ways can you select three tide gauge records from 10 available stations? This is a question of combinations:

$${}_{10}C_3 = \binom{10}{3} = \frac{10!}{3!7!} = \frac{8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3} = 8 \cdot 3 \cdot 5 = 120. \tag{3.7}$$

Likewise, per (3.6), there are also 120 ways to select 7 tide gauge records from the same 10 stations.

3.1.3 Probability

So far we have studied only what is *possible* in a given situation. We have listed all possibilities or determined how many possibilities there are. However, to be of use to us we need to be able to judge which of the possibilities are *probable* and which are *improbable*. The basic concept of probability can be stated thus: If there are n possible

outcomes or results, and s of those are regarded as favorable (or as "successes"), then the probability of success is given by

$$P = s/n. \tag{3.8}$$

This classical definition applies only when all possible outcomes are *equally likely*.

Example 3–4. What is the probability of drawing an ace from a deck of cards? *Answer*: $P = 4/52 = 1/13 = 7.7\%$. How about getting a 3 *or* a 4 with a balanced die? *Answer*: $s = 2$ and $n = 6$, so $P = 2/6 = 33\%$

While equally likely possibilities are found mostly in games of chance, the classical probability concept also applies to random selections, such as making selections to reduce a large set of data down to a manageable quantity without introducing sampling bias.

Example 3–5. If three of 20 water samples have been contaminated and you select four random samples, what is the probability of picking one of the bad samples?

Answer: We have $\binom{20}{4} = 3 \cdot 5 \cdot 17 \cdot 19 = 4845$ ways of making the selection of our four samples. The number of "favorable" outcomes is $\binom{17}{3}$ [we pick three good samples of the 17 good ones] times $\binom{3}{1}$ [we pick one of the three bad samples] = 2040. It then follows that the probability is $P = s/n = 2040/4845 = 42\%$. Here we used the rule of multiplicative choices.

Obviously, the classical probability concept will not be useful when some outcomes are more likely than others. A better definition would then be

> *The probability of an event is the proportion of the time that events of the same kind will occur in the long run.*

So, when the National Weather Service says that the chance of rain on any day in June is 0.2, it is based on past experiences that on average June had 6 days of rain. Another important probability theorem is the *law of large numbers*, which states

> *If a situation, trial, or experiment is repeated again and again, the proportion of successes will tend to approach the probability that any one outcome will be a success.*

which is basically our probability concept in reverse.

Coin tosses illustrate the law of large numbers nicely. We toss the coin and keep track of how many times we get "heads" versus the total number of tosses. For a nice symmetric coin we expect the proportion of heads to total tosses to approach 0.5 over the long haul, but initially we are not surprised that there can be large departures from this expectation. Figure 3.2 shows how the proportion may oscillate for a small number of tosses but eventually it will approach the expected value.

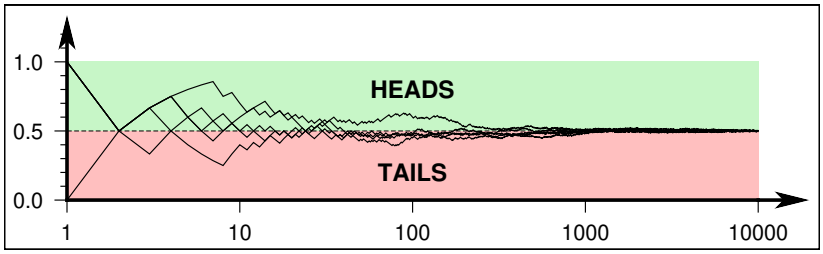


Figure 3.2: Proportion of heads in a series of coin tosses. The more tosses we complete, the closer the ratio of heads to total tosses will approach 0.5. Shown are five separate sequences. They differ considerably for small numbers but all converge on the expected proportion.

3.1.4 Some rules of probability

In statistics, the set of all possible outcomes of an experiment is called the *sample space*, usually denoted by the letter S. Any subset of S is called an *event*. An event may contain more than one item. Sample spaces may be finite or infinite. Two events that have no elements in common are said to be *mutually exclusive*, meaning they cannot both occur at the same time.

There are only positive (or zero) probabilities, symbolically written

$$P(A) \geq 0 \tag{3.9}$$

for any event A. Every sample space has probability 1, so that

$$P(S) = 1, \tag{3.10}$$

where $P = 1$ means absolute certainty. If two events are mutually exclusive, the probability that one *or* the other will occur equals the sum of their probabilities

$$P(A \cup B) = P(A) + P(B). \tag{3.11}$$

Regarding the notation, $\cup$ means *union* (which we read as "OR"), $\cap$ means *intersection* ("AND"), and $'$ (the prime symbol) means *complement* ("NOT"). We can furthermore state that

$$P(A) \leq 1, \tag{3.12}$$

since absolute certainty is the most we can ask for. Also,

$$P(A) + P(A') = 1, \tag{3.13}$$

since it is certain that an event either will or will not occur.

3.1.5 Probabilities and odds

Bookmakers in London use a slightly different system of reporting probabilities. If the probability of an event is p, then the *odds* for its occurrence are

$$a : b = \frac{p}{1-p}. \tag{3.14}$$

The inverse relation gives

$$p = \frac{a}{a+b}. \tag{3.15}$$

If you are still reading this book then odds are you will pass this course!

3.1.6 Addition rules

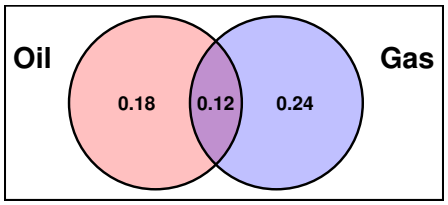


Figure 3.3: A Venn diagram illustrating the probabilities of finding hydrocarbons. The overlapping magenta wedge graphically represents the probabilities of finding *both* oil and gas.

The addition rules demonstrated above only holds for *mutually exclusive events*. Let us now consider a more general case. The sketch in Figure 3.3 is a *Venn diagram*, a handy graphical way of illustrating the various combinations

of possibilities and probabilities. The diagram illustrates the probabilities associated with finding hydrocarbons during a hypothetical exploration campaign. We see from the diagram that

$$\begin{aligned} P(\text{oil}) &= 0.18+0.12=0.3, \\ P(\text{gas}) &= 0.24+0.12=0.36, \\ P(\text{oil}\cup\text{gas}) &= 0.18+0.12+0.24=0.54. \end{aligned} \tag{3.16}$$

Now, if we used the simple addition rule (3.11), we would find

$$P(\text{oil}\cup\text{gas}) = P(\text{oil})+P(\text{gas}) = 0.3+0.36 = 0.66. \tag{3.17}$$

This value overestimates the probability, because finding oil and finding gas are *not* mutually exclusive since we might find both. We can correct the equation by writing

$$P(\text{oil}\cup\text{gas}) = P(\text{oil})+P(\text{gas})-P(\text{oil}\cap\text{gas}) = 0.3+0.36-0.12 = 0.54. \tag{3.18}$$

The general addition rule for probabilities thus becomes

$$P(A\cup B) = P(A)+P(B)-P(A\cap B). \tag{3.19}$$

Note that if the events *are* mutually exclusive then $P(A\cap B)=0$ and we recover the original rule.

3.1.7 Conditional probability and Bayes basic theorem

We must sometimes evaluate the probability of an event *given that another event already has occurred*. We write the probability that A will occur given that B already has occurred as

$$P(A|B) = \frac{P(A\cap B)}{P(B)}. \tag{3.20}$$

In our exploration example, we can find the probability of finding oil given that gas already has been found as

$$P(\text{oil}|\text{gas}) = \frac{P(\text{oil}\cap\text{gas})}{P(\text{gas})} = \frac{0.12}{0.36} = \frac{1}{3}. \tag{3.21}$$

We can now derive a general multiplication rule from (3.20) by multiplying it by $P(B)$ and exchange A and B, which gives

$$\begin{aligned} P(A\cap B) &= P(B)P(A|B) \\ P(A\cap B) &= P(A)P(B|A) \end{aligned} \tag{3.22}$$

and implies that the probability of both events A and B occurring is given by the probability of one event occurring multiplied by the probability that the other event will occur given that the first one already has occurred (occurs, or will occur). This rule is called the *joint probability* or *Bayes basic theorem*. Now, if the events A and B are independent events, then the probability that A will take place is not influenced by whether B has taken place or not, i.e.

$$P(A|B) = P(A). \tag{3.23}$$

Substituting this expression into (3.22) we obtain

$$P(A\cap B) = P(A)\cdot P(B). \tag{3.24}$$

That is, the probability that two independent events A and B both will occur equals the product of their probabilities. In general, for n independent events with individual probability p_i, the probability that all n events occur is

$$P = \prod_{i=1}^{n} p_i. \tag{3.25}$$

Example 3–6. What is the probability of rolling three ones in a row with a balanced die?

Answer: With $n = 3$ and $p = 1/6$,

$$P = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \approx 0.005. \tag{3.26}$$

While $P(A|B)$ and $P(B|A)$ may look similar, they can be vastly different. For example, let A be the event of a death on the Bay Bridge connecting San Francisco and Oakland, and B the event of a magnitude 8 earthquake in the area. Then, $P(A|B)$ is the probability of a fatality on the Bay Bridge *given* that a large earthquake has taken place nearby, while $P(B|A)$ is the probability that we will have a magnitude 8 quake *given* that a death has been reported on the bridge. Clearly $P(A|B)$ seems more likely than $P(B|A)$ since we know the former to have happened in the past. On the other hand, we can list many causes of fatalities on the freeway other than earthquakes (e.g., traffic accidents, heart attacks, old age, road rage, talk radio rants, and so on).

We can arrive at a relation between $P(B|A)$ and $P(A|B)$ by equating the two expressions for $P(A \cap B)$ in (3.22). We obtain $P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$, or

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{P(A)}. \tag{3.27}$$

This is a useful relation since we may sometimes know one conditional probability but are interested in the inverse relationship. For example, we may know that salt domes (known as potential traps for hydrocarbons) often are associated with large curvatures in the gravity field. However, we may be more interested in the converse: Given that large curvatures in the gravity field exist, what is the probability that salt domes are the cause of such anomalies?

3.1.8 Bayes general theorem

If there are more than one event B_i (all mutually exclusive) that are conditionally related to an event A, then $P(A)$ is simply the sum of the conditional probabilities of the events B_i times their individual probabilities, i.e.

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i). \tag{3.28}$$

Substituting (3.28) into (3.27) gives, for any of the n events B_i,

$$P(B_i|A) = \frac{P(B_i) \cdot P(A|B_i)}{\sum_{j=1}^{n} P(A|B_j) \cdot P(B_j)}. \tag{3.29}$$

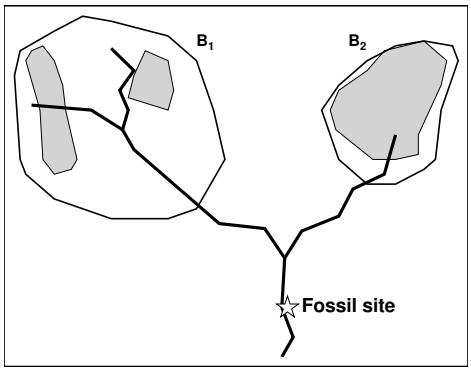


Figure 3.4: Location of a fossil discovery with respect to the two drainage basins from which it must have originated. Bayes theorem provides a formal way to assign likelihood to the possible origins.

This is the general *Bayes theorem*.

Example 3–7. Let us assume that an unknown marine fossil fragment was found in a dry stream bed in northern

Sahara. Excited, a paleontologist would like to send out an expendable graduate student field party to search for a more complete specimen of the unknown species. Unfortunately, the source of the fragment cannot be identified uniquely since it was found several kilometers below the junction of two dry stream tributaries (Figure 3.4). The drainage basin B_1 of the larger stream covers 407.5 km^2, while the other basin (B_2) covers only 207.5 km^2. Based on this difference in basin size alone we might expect the probabilities that the fragment came from one of the basins are

$$\begin{aligned} P(B_1) &= \tfrac{407.5}{615} = 0.66, \\ P(B_2) &= \tfrac{207.5}{615} = 0.34, \end{aligned} \tag{3.30}$$

based solely on the proportion of each basin's area to the combined area. However, inspecting an ancient British-produced geological map reveals that only 31% of the outcropping rocks in the larger basin B_1 are marine, whereas almost 85% of the outcrops in basin B_2 are marine. We can now state two conditional probabilities:

$P(A|B_1) = 0.31$ (Probability of a marine fossil, given it was derived from basin B_1.)
$P(A|B_2) = 0.85$ (Probability of a marine fossil, given it was derived from basin B_2.)

With these probabilities and Bayes general theorem (3.29) we can find the conditional probability that the fossil came from basin B_1 given that the fossil is marine:

$$P(B_1|A) = \frac{P(A|B_1)\cdot P(B_1)}{P(A|B_1)\cdot P(B_1) + P(A|B_2)\cdot P(B_2)} = \frac{0.31\cdot 0.66}{0.31\cdot 0.66 + 0.85\cdot 0.34} = 0.41. \tag{3.31}$$

Consequently, the probability of the fossil coming from the smaller basin B_2 is the complimentary probability

$$P(B_2|A) = 0.59. \tag{3.32}$$

It therefore seems somewhat more likely that the smaller basin was the source of the fossil and that this area should be the initial target for the student-led expedition. However, $P(B_1|A)$ and $P(B_2|A)$ are not dramatically different and depends to some extent on the assumptions used to select $P(B_i)$ and $P(A|B_i)$ in the first place. Bayes general theorem is extensively used in such search and find scenarios and the probabilities that go into the procedure are constantly being revised as more is learned during the search.

3.2 The M&M's of Statistics

When discussing exploratory data analysis we mentioned that it is useful to be able to present large data sets using just a few parameters. We saw the box-and-whisker diagram graphically summarized a data distribution. However, it is often desirable to represent a data set by a *single* number which, in its way, is descriptive of the entire data set. We will see there are several ways to select this "representative" value. We will mostly be concerned with measures that somehow describe the center or middle of the data set. These are called estimates of *central location*.

3.2.1 Population and samples

If a data set consists of all conceivably possible (or hypothetically possible) observations of a certain phenomenon then we call it a *population*. A population can be finite or infinite. Any subset of the population is called a *sample*. Thus, a series of 12 coin-tosses is a sample of the potentially unlimited number of tosses in the population. We will most often find that we are analyzing samples taken from a much larger population, and our aim will be to learn something about the population by studying the smaller sample set (Figure 3.5).

3.2.2 Measures of central location (mean, median, mode)

The best known estimate of central location is called the *arithmetic mean*, defined as

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{3.33}$$

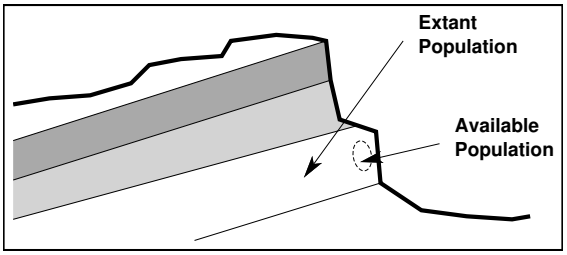


Figure 3.5: We must always try to select an unbiased sample from the population. In this example we are sampling the weathered outcrop of a sedimentary layer, which most likely is not representative of the entire formation.

The mean is also loosely called the "average." Resist being that sloppy! When reporting the mean value, always say "mean" and not "average" so that the reader knows exactly what you have done. We call $\bar{x}$ the *sample mean* to distinguish it from the true mean of the population, denoted

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i, \tag{3.34}$$

which likely will remain unknown to us. The mean has many useful properties, which explains its common use:

- It can always be calculated for any numerical data, i.e., it always exists.
- It is unique and straightforward to calculate.
- It is relatively stable and does not fluctuate much from sample to sample taken from the same population.
- It lends itself to further statistical treatment: several $\bar{x}$ estimates from subgroups can later be combined into an overall grand mean.
- It takes into account every data value.

However, the last property can sometimes be a liability. Should a few points deviate excessively from the bulk of the data then it does not make sense to include them in the sample. A better estimate for the central location may then be the *sample median*:

$$\text{median } x_i = \tilde{x} = \begin{cases} x_{n/2+1}, & n \text{ is odd} \\ \dfrac{1}{2}(x_{n/2+1} + x_{n/2}), & n \text{ is even} \end{cases} \tag{3.35}$$

Here, the data first must be sorted into ascending (or descending) order. We then choose the middle value (or mean of the two middle values for even n) as our median estimate.

Consider this sample of sandstone densities: $\{2.30, 2.20, 2.35, 2.25, 2.30, 23.0, 2.25\}$, $n = 7$. The median density can be found to be $\tilde{x} = 2.30$, a reasonable value, while the mean density $\bar{x} = 5.24$, which is a rather useless estimate since it is clearly far outside the bulk of the data *and* outside the range of known sandstone densities anywhere. For this reason we say that the median is a *robust* estimate of central location. Here it is rather obvious that the value 23.0, which probably is a clerical error, threw off the mean and we could correct for that by excluding it from the calculation and find $\bar{x} = 2.28$ instead. However, in many cases our data set will be very large and we must anticipate that some values may be erroneous.

The disadvantage of the median is the need to sort the data, which can be slow. (Do you think this is a valid reason not to use it?). However, like the mean, the median always exists and is unique.

Our final traditional estimate for central location is the *mode*. The mode is defined as the observation that occurs the most frequently. For defining the central location the mode is at a disadvantage since it may not exist (perhaps no two values are the same) or it may not be unique (our densities actually have two modes). Of course, if our data set is expected to have more than one "peak," modal estimates are important, and we will return to that later. The mode will be denoted as $\hat{x}$. The mean, median and mode of a distribution typically are related as indicated in Figure 3.6.

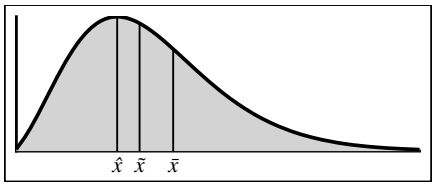


Figure 3.6: The relationship between the mean, median, and mode estimates of central location for a skewed data distribution. These estimates will all coincide for a perfectly symmetric and unimodal distribution.

Returning to the mean, it is occasionally the case that some measurements are considered more important than others. It could be that some observations were made with a more precise instrument, or simply that some values are not as well documented as others. These are examples of situations where we should use a *weighted mean*

$$\bar{x} = \sum_{i=1}^{n} w_i x_i \Bigg/ \sum_{i=1}^{n} w_i \ , \tag{3.36}$$

where w_i is the weight of the i'th data value. If all $w_i = 1$ then we recover the original definition for the mean (3.33). This general equation is also convenient when we need to compute the overall, or *grand mean* based on the individual means from several data sets. The grand mean based on m data sets may be written as

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{m} n_i \bar{x}_i}{\sum_{i=1}^{m} n_i}, \tag{3.37}$$

where the sample sizes n_i take the place of the weights in (3.36).

3.2.3 Measures of variation

While a measure of central location is an important attribute of our data, it says little about how the data are distributed. We need some way of representing the *variation* of our observations about the central location. In the EDA section, we used the *range* and *hinges* to indicate data variability. Another way to define the variability would be to compute the deviations from the mean,

$$\Delta x_i = x_i - \bar{x}, \tag{3.38}$$

and take the average of the sum of deviations, $\frac{1}{n}\sum_{i=1}^{n} \Delta x_i$. Sadly, it turns out that this sum is always zero, which makes it rather useless for our purposes. A more useful quantity might be the mean of the absolute value (AD) of the deviations:

$$AD = \frac{1}{n}\sum_{i=1}^{n} |\Delta x_i|. \tag{3.39}$$

Because of the absolute value sign this function is nonanalytic and often completely ignored by statisticians. You will find very superficial treatment of medians and absolute deviations in most elementary statistics books. However, when dealing with real data that include occasional bad values, the AD is useful, just as the median can be more useful than the mean. However, the most common way to describe variation of a population is to define it as the average *squared* deviation. Hence, the population *variance* is

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2, \tag{3.40}$$

and the population *standard deviation* is therefore

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2}. \tag{3.41}$$

Most often we will be working with samples rather than entire populations, and we hope (and will later test) that the sample is representative of the population. The sample standard deviation s is given by

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{3.42}$$

Note that we are dividing by $n-1$ rather than by n. This is done because $\bar{x}$ must first be *estimated* from the sample rather than being a *given* parameter of the population, such as μ and N. This reduces the degrees of freedom by one; hence we divide by $n-1$ (we will have more to say about degrees of freedom in Section 4.2.1). We can now show one property of the mean: It is clear that s^2 depends on the choice for $\bar{x}$. Let us find the value for $\bar{x}$ in (3.42) that gives the smallest value for s^2. Consider

$$f(\bar{x}) = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2. \tag{3.43}$$

The function f has a minimum where $df/d\bar{x} = 0$ and $d^2f/d\bar{x}^2 > 0$, so we find

$$\frac{df}{d\bar{x}} = \frac{\sum_{i=1}^{n} -2(x_i - \bar{x})}{n-1} = \frac{-2}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}) = 0, \tag{3.44}$$

which gives

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0. \tag{3.45}$$

We can solve this equation and find

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{3.46}$$

Since

$$\frac{d^2f}{dx^2} = \frac{2n}{n-1} > 0, \tag{3.47}$$

we know that f has a minimum for this value of $\bar{x}$. Thus, we have shown that the value $\bar{x}$ that minimizes the standard deviation equals the mean we defined earlier in (3.33). This is a very useful and important property of the mean. Because $\bar{x}$ minimizes the squared "misfit", it is also called the *least-squares estimate* of central location (or L_2 estimate for short). When computing the mean and standard deviation on a computer we do not normally use (3.42) since it requires two passes through the data: One to compute the $\bar{x}$ and another to solve (3.42). Rather, we rearrange (3.42) to give

$$\begin{aligned} s &= \sqrt{\sum_{i=1}^{n}\frac{(x_i - \bar{x})^2}{n-1}} = \sqrt{\sum_{i=1}^{n}\frac{x_i^2 - 2x_i\bar{x} + \bar{x}^2}{n-1}} \\ &= \sqrt{\frac{n\sum x_i^2 - 2n\bar{x}\sum x_i + n\sum \bar{x}^2}{n(n-1)}} = \sqrt{\frac{n\sum x_i^2 - (\sum x_i)^2}{n(n-1)}}. \end{aligned} \tag{3.48}$$

3.2.4 Robust estimation

We found that the arithmetic mean is the value that minimizes the sum of the squared deviations from the central value. Can we apply the same argument to the mean absolute deviation and find what the best value for $\tilde{x}$ may be? In other words, let

$$\frac{d}{d\tilde{x}}\left(\frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}|\right) = -\frac{1}{n}\sum_{i=1}^{n}\frac{x_i - \tilde{x}}{|x_i - \tilde{x}|} = 0. \tag{3.49}$$

The term inside the summation can only take on the values -1, 0, or $+1$. Thus, the only $\tilde{x}$ that can satisfy (3.49) is a value chosen such that half the x_i are smaller (giving -1) and half the x_i are larger (giving $+1$), and for odd sample

sizes we also get one or more exact zeros. Thus, we have proven that the median is the location estimate that minimizes the mean absolute deviation. The median is also called the L_1-estimate of central location.

The discussion of mean and median brings up the general issue of *robust estimation*: How to calculate a stable and reasonable estimate of central location in the presence of contaminated data? As an indicator of how robust a method is, we will introduce the concept of "breakdown point." It is the *smallest fraction* of the observations that must be replaced by outliers in order to throw the estimator outside reasonable bounds.

We have already seen that even a single bad value is enough to throw the mean way off. For our densities of sandstone, we had $\rho = \{2.2, 2.25, 2.25, 2.3, 2.3, 2.35, 23.0\}$, with $n = 7$. If we realized that 23.0 should be 2.3, we find $\bar{\rho} = 2.28 \pm 0.05$, while if we included $\rho_7 = 23.0$ we would find $\bar{\rho} = 5.24 \pm 7.8$. The second estimate is obviously far outside the 2.20–2.35 range we first determined. We can therefore say that the least squares estimate (i.e., the mean) has a breakdown value of $1/n$; it only takes one outlier to ruin our day. On the other hand, note that the median is ~ 2.3 in both cases, well inside the acceptable interval. It is found that the breakdown point of the median is 50%: We would have to replace half the data with bad outliers to move the estimate of the median outside the range of the original (good) data values.

Apart from the central location estimator, we also want a robust estimate of the spread of the data. Clearly, the classical standard deviation is problematic since only one bad value will make it biased due to the x^2 effect. From the success of taking the median of a set of numbers rather than summing them up, could we do something similar with the deviations? Consider what value of $\tilde{x}$ would minimize the median of $\{|x_i - \tilde{x}|\}$. You can probably see for yourselves that the $\tilde{x}$ must equal our old friend the median. Because of the robustness of the median operator, we will often use the quantity called the *median absolute deviation* (*MAD*) as our robust estimate of "spread" or variation. Note: Many textbooks and software packages (such as MATLAB) use *MAD* to indicate *mean absolute deviation* instead, as defined in (3.39) and called *AD* in these notes. Thus, we define

$$MAD = 1.4826 \text{ median } |x_i - \tilde{x}|, \tag{3.50}$$

where the factor 1.4826 is a correction term that makes the *MAD* equal to the standard deviation of normally distributed data[1]. Like the median, the *MAD* has a breakdown point of 50%. The *MAD* for our example was 0.07 and it remained unchanged by using the contaminated value. Having robust estimates of central location and scale, we can attempt to identify *outliers*. We may compute the robust *standard units*

$$z_i = \frac{x_i - \tilde{x}}{MAD} \tag{3.51}$$

and compare them to a cutoff value: If $|z_i| > z_{cut}$ we say we have detected an outlier. The choice for z_{cut} is to a certain extent arbitrary. It is, however, quite standard to choose $z_{cut} = 2.5$. Chances that any z_i will exceed z_{cut} is very small if the z_i's came from a normal distribution. Our normalized densities (including the contaminated value) using $\bar{x}$ and s to compute z_i gives

$$z_{L_2} = \{-0.39, -0.38, -0.38, -0.377, -0.377, -0.37, 2.28\}, \tag{3.52}$$

where none of the values qualify as an outlier. Using the median and *MAD* instead, we find

$$z_{L_1} = \{-1.35, -0.68, -0.68, 0.0, 0.0, 0.68, 280.0\}, \tag{3.53}$$

and we see that the bad observation gives a huge z-value two orders of magnitude larger than any other. Clearly, the least-squares technique alone is not trustworthy when it comes to detecting bad points. The outlier-detecting scheme presents us with an elegant two-step technique: First find and remove the outliers from the data, then use classical *least-squares* techniques on the remaining data points. The resulting statistics are called the *least trimmed squares* estimates (LTS). We will return to the concept of robustness when discussing regression in Chapter 6.

3.2.5 Central limit theorem

How well does our sample mean, $\bar{x}$, compare to the true population mean, μ? An important theorem, called the *central limit theorem*, states

> *If n (the sample size) is large, the theoretical sampling distribution of the mean can be approximated closely with a normal distribution.*

[1] This factor equals $1/P_c^{-1}(0.75)$, where $z = P_c^{-1}(p)$ is the inverse cumulative normal distribution; see Table A.1.

This is rather important since it justifies the use of the normal distribution in a wide range of situations. It simply states that the sample mean $\bar{x}$ is an *unbiased estimate* of the population mean and that the scatter about μ is *normally distributed*. It can be shown that the standard deviation of the sampling mean, $s_{\bar{x}}$, is related to the population deviation, σ, by

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \tag{3.54}$$

or

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \tag{3.55}$$

depending on whether the population is infinite (3.54) or finite of size N (3.55). Thus, as n grows large, $s_{\bar{x}} \to 0$. Furthermore, the sample variance s^2 has the mean value σ^2 with standard deviation

$$\sigma_s^2 = \frac{2\sigma^4}{n-1}, \tag{3.56}$$

which also $\to 0$ for large n. For our analysis we will substitute the sample standard deviation s *in lieu* of the unknown population standard deviation σ, since s is an *unbiased estimator* of σ.

3.2.6 Covariance and correlation

We found earlier that the sample variance was defined as

$$s_x^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})}{n-1}. \tag{3.57}$$

It is often the case that our data set consists of pairs of properties, such as sets of (depth, pressure), (time, temperature), concentrations of two elements, and more. Denoting the paired properties by x and y, we can compute the variance of each quantity separately. For instance, for y we find

$$s_y^2 = \frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1} = \frac{\sum_{i=1}^{n}(y_i-\bar{y})(y_i-\bar{y})}{n-1}. \tag{3.58}$$

We can now define the *covariance* between x and y in a similar way as

$$s_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n-1}. \tag{3.59}$$

While s_x and s_y tell us how the x and y values are distributed *individually*, s_{xy} tells us how the x and y values vary *together*.

Because the value of the covariance clearly depends on the units of x and y, it is difficult to state what covariance values are meaningful. This difficulty is overcome by defining the Pearson *correlation coefficient* r, which normalizes the covariance to yield correlations in the ± 1 range, i.e.,

$$r = \frac{s_{xy}}{s_x s_y}. \tag{3.60}$$

If $|r|$ is close to 1, then the variables are strongly correlated or anti-correlated. Values of r close to 0 mean that there is little significant correlation between the data pairs. Figure 3.7 shows some examples of data pairs and their correlations. We see that in general, r will tell us how well the data are "clustered" in some direction. Note in particular example (f), which presents data that are clearly correlated (i.e., all pairs lie on a circle), yet $r = 0$. This occurs because r is a measure of a *linear* relationship between values; a nonlinear relationship may not register a significant correlation. Thus, we must be careful with how we use r to draw conclusions about the interdependency of paired values. For example, if our (x, y) data are governed by a $y = \sqrt{x}$ law then we may find a fairly good correlation between x and y, but we would be wrong to conclude that x and y have a *linear* relationship (plotting y versus $\sqrt{x}$ *would give* a linear relationship and a much higher value of r). We will return to correlation under the rubrics of curve fitting and multiple regression in Chapter 6.

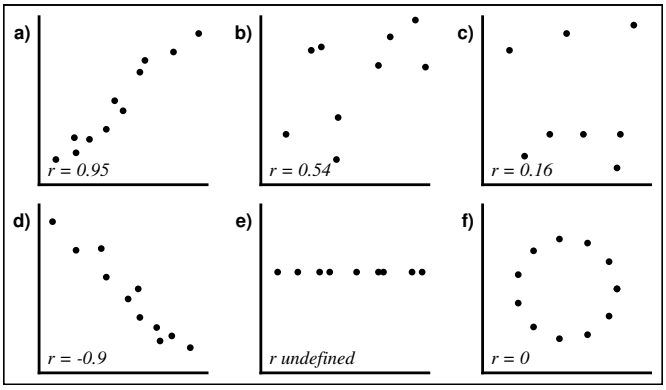


Figure 3.7: Some examples of data sets and their correlation coefficients. Note that the perfect circular correlation in (f) gives a zero linear correlation coefficient. While clearly x and y are correlated, their relationship is not *linear*.

3.2.7 Moments

Returning to the L_2 estimates, we will briefly introduce the concept of *moments*. In general, the r'th moment is defined as

$$m_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^r, \tag{3.61}$$

except for $r = 1$ where it is customary to use the "raw moment" about zero instead. From this definition it can be seen that the mean and variance are the first (raw) and second (central) moments, respectively. We will look at two higher order (central) moments that one may encounter in the literature. The first is called the *skewness* (SK) and it is the third central moment, given by

$$SK = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^3 = \frac{1}{n}\sum_{i=1}^{n} z_i^3, \tag{3.62}$$

where we normalize by s to get dimensionless values for SK. The skewness is used to investigate our data sets' *degree of symmetry* about the mean. A positive SK means we have a longer tail to the right of the mean than to the left, and vice versa for a negative SK (Figure 3.8).

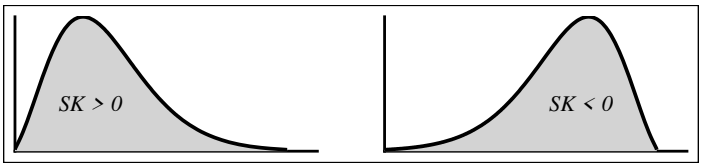


Figure 3.8: Examples of data distributions with positive and negative skewness. The sign of the skewness indicates which side of the distribution is long-tailed.

Unfortunately, if the data contain outliers then the SK will be very sensitive to these values and consequently be of little use to us. A more robust estimate of skewness is the *Pearson coefficient of skewness*,

$$SK_p = \frac{3(\bar{x} - \tilde{x})}{s}, \tag{3.63}$$

where we basically compare the mean and the median. An even higher-order central moment is the *kurtosis*,

$$K = \left\{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^4\right\} - 3 = \left\{\frac{1}{n}\sum_{i=1}^{n} z_i^4\right\} - 3. \tag{3.64}$$

The correction term -3 makes $K = 0$ for a normal distribution, which we will discuss shortly. The kurtosis K attempts to quantify a data distribution's "sharpness" ($K > 0$) or "flatness" ($K < 0$; Figure 3.9). However, for most real data K can be almost infinite and should be used only with "well-behaved" data.

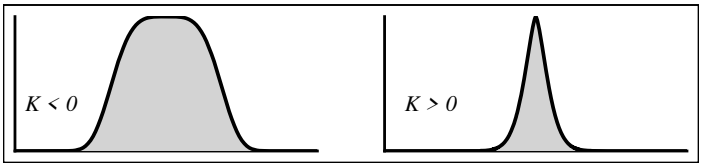


Figure 3.9: Examples of distributions with different kurtosis. Distributions with negative K are called *platykurtic*, while a positive K is called *leptokurtic*. You will of course be immensely pleased to learn that an intermediate case is called *mesokurtic*.

3.3 Discrete Probability Distributions

An important concept in statistics and probability is the notion of a *probability distribution*. It is a function $P(x)$, which indicates the probability that the event x will take place. $P(x)$ can be a discrete or continuous function. As an example of a discrete function, consider the function $P(x), x = 1,2,..,6$, that gives the probability of throwing an x with a balanced die:

$$P(x) = 1/6, \quad x = 1,2,\ldots,6, \tag{3.65}$$

or for flipping a coin:

$$P(x) = 1/2, \quad x = \{H,T\}. \tag{3.66}$$

Staying with the throws of the die, we can relate $P(x)$ to the area under the curve in Figure 3.10.

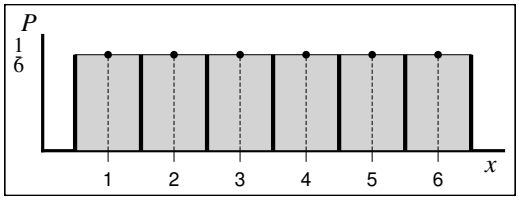


Figure 3.10: Probability of throwing any number on a die is a constant $1/6$, unless the die is "loaded".

Two important properties shared by all discrete probability distributions are

$$0 \leq P(x_i) \leq 1, \text{ for all } x_i, \tag{3.67}$$

$$\sum_{i=1}^{n} P(x_i) = 1. \tag{3.68}$$

3.3.1 Binomial probability distribution

Often we are more interested in knowing the probability of a certain outcome after n repeated tries, such as "what is the probability of receiving junk mail three days in one week?" To derive such a function, we will assume that each event is independent and has the same probability, p. Then, the probability that an event *does not* occur is the complement, $q = 1 - p$. Consequently, the probability of getting x successes in n tries (and thus $n - x$ failures) is

$$P_1(x) = p^x q^{n-x}. \tag{3.69}$$

However, this probability applies to a *specific order* of all possible outcomes. Since we may not care about the order in which the successful x events occurred, we must scale $P_1(x)$ by the number of possible combinations of x successes in n tries. We already know this amount to be given by $\binom{n}{x}$, so our discrete probability function becomes

$$P_{n,p}(x) = \binom{n}{x} p^x q^{n-x} = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0,1,\ldots,n. \tag{3.70}$$

This expression is known as the binomial probability distribution or simply the *binomial distribution* (Figure 3.11) and it is used to predict the probability that x events out of n tries will be successful, given that each independent x has the

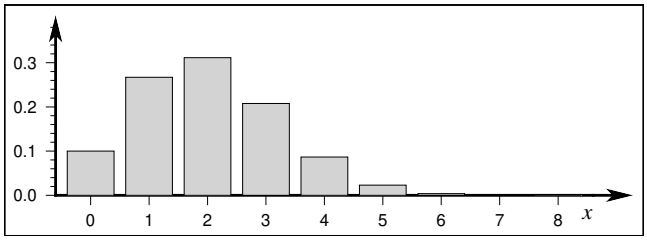


Figure 3.11: Binomial probability distribution $P_{n,p}(x)$, which shows the probability of having x successful outcomes out of a total of n tries, when each try has the probability p of success (and $q = 1 - p$ of failure). Here, $p = 0.25$ and $n = 8$.

probability p of success.

Example 3–8. What are the chances of drawing three red cards in six tries from a deck (assuming we place the card back into the deck after each try)? Here $p = 1/2$, so

$$P_{6,0.5}(3) = \frac{6!}{3!3!}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{6-3} = 0.31. \tag{3.71}$$

One might have thought that getting half red and half black cards would have a higher probability, but remember that we require *exactly* 3 reds. If we compute the probability of getting 1, 2, or 3 reds separately and used the summation rule to compute the probability that we would draw 1, 2, or 3 red cards then P would be much higher.

The binomial probability distribution can also be used to assess the likelihood of more serious scenarios, such as the next example presents.

Example 3–9. A silver-tonged con artist approaches you on a street in New York City with a simple proposition: He has 10 beads — 9 black and one white. You get to pick one bead from his bag. You are given six opportunities to draw a bead (the bead is returned to the bag after each try), and if anytime during the six tries you pick the white bead then you have won and he will give you $20. However, if you have not picked the white bead after six tries then you owe him $20 instead. Is this a good deal? Answer: Clearly, the probability of picking the white bead is fixed at $p = 0.1$. To lose the bet you will have to come up empty-handed six times in a row. For $n = 6$ and $r = 0$ the chances of that is simply

$$P_{6,0.1}(0) = \binom{6}{0}0.1^0(1-0.1)^6 = 0.53. \tag{3.72}$$

So while it is close to 50–50 the con-artist will most likely win, at least in the long run. You probably should also be concerned that there might be something else going on as well, such as sleight-of-hand removal of the white bead before each try...

3.3.2 The Poisson distribution

In some situations, the binomial distribution can be approximated by simpler expressions. One such case arises when the probability p for one event is very small and n is large. Such events are called *rare*, and the discrete distribution may then be approximated by

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots, n \tag{3.73}$$

where $\lambda = np$ is the *rate of occurrence*. The Poisson distribution can be used to evaluate the probabilities for the occurrence of rare events such as large earthquakes, volcanic eruptions, and reversals of the geomagnetic field. For instance, the number of floods occurring in a 50-year period has been shown to follow a Poisson distribution with $\lambda = 2.2$. What is the probability that we will have at least one flood in the next 50 year period? Here, $P = 1 - P_0$, the

probability of having no flood. Plugging in for $x = 0$ and $\lambda = 2.2$ we find $P_0 = 0.1108$, so $P = 0.8892$.

Example 3–10. A student is monitoring the radioactive decay of a certain sample that is expected to undergo three decays per minute. The student observes the number of decays over 100 individual one-minute periods and constructs the summary shown in Table 3.1. Does the data support the expected decay rate? We make a histogram of the data by

Decays	0	1	2	3	4	5	6	7	8	9
Observed	5	19	23	21	14	12	3	2	1	0

Table 3.1: Number of decays observed in one-minute interval.

normalizing the observed frequencies by the total count and superimposing the Poisson distribution for the expected rate. The result (Figure 3.12) shows a very good fit.

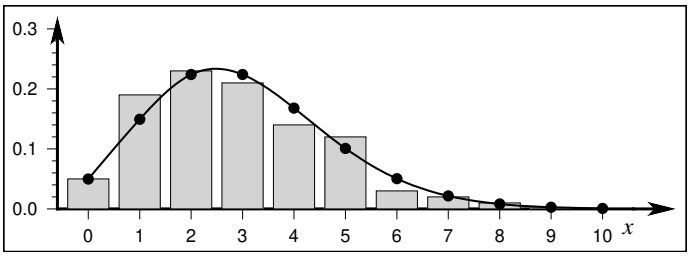


Figure 3.12: Histogram of observed decay rate frequencies (bars) and the theoretical Poisson distribution (circles) for the expected rate $\lambda = 3$.

3.4 Continuous Probability Distributions

While many populations are of a discrete nature (e.g., outcomes of coin tosses, numbers of microfossils in a core, etc.), we are very often dealing with observations of a phenomenon that can take on any of a continuous spectrum of values. We may sample the phenomenon at certain points in space-time and thus have discrete observations. Nevertheless, the underlying probability distribution is continuous (e.g., Figure 3.13).

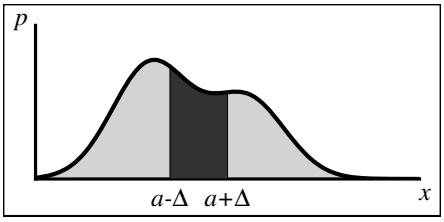


Figure 3.13: Example of a continuous probability density function (pdf). The area under any pdf must equal 1. The finite probability identified in (3.75) is indicated in dark gray.

Continuous distributions can be thought of as the limit for discrete distributions when the "spacing" between events shrinks to zero. Hence, we must replace the summation in (3.68) with the integral

$$\int_{-\infty}^{\infty} p(x)dx = 1. \tag{3.74}$$

Because of their continuous nature, functions such as $p(x)$ in (3.74) are called *probability density functions* (pdf). The probability of an event is still defined by the area under the curve, but now we must integrate to find the area and hence

the probability. E.g., the probability that a random variable will take on a value between $a-\Delta$ and $a+\Delta$ is

$$P(a\pm\Delta)=\int_{a-\Delta}^{a+\Delta}p(x)dx. \tag{3.75}$$

As $\Delta\to 0$ we find that the probability goes to zero. Thus, the probability of getting exactly $x=a$ is nil.

The *cumulative distribution function* (cdf) gives the probability that an observation less than or equal to a will occur. We obtain the integral expression for this distribution by replacing the lower limit by $-\infty$ and the upper limit by a, finding

$$P_c(a)=\int_{-\infty}^{a}p(x)dx. \tag{3.76}$$

Obviously, as $a\to\infty, P_c(a)\to 1$. Given the cumulative distribution function we can revisit (3.75) and instead state

$$P(a\pm\Delta)=P_c(a+\Delta)-P_c(a-\Delta). \tag{3.77}$$

3.4.1 The normal distribution

So far the function $p(x)$ has been arbitrary. Any continuous function with unit area under the curve (i.e., 3.74) would qualify. We will now turn our attention to the best known and most frequently used pdf: the *normal distribution*. Its study dates back to 18th century investigations into the nature of experimental error. It was found that repeat measurements of the same quantity displayed a surprising degree of regularity. In particular, the German scientist K. F. Gauss played a major role in developing the theoretical foundations for the normal distribution, hence its other name: the *Gaussian* distribution. It is given by

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{3.78}$$

where μ and σ have been defined previously. The constant term before the exponential normalizes the area under the curve to unity (Figure 3.14). As discussed in Section 3.2.4, it is often convenient to transform your data into so-called *standard scores*:

$$z_i=\frac{x_i-\mu}{\sigma}, \tag{3.79}$$

in which case (3.78) reduces to

$$p(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}, \tag{3.80}$$

which has zero mean and unit standard deviation.

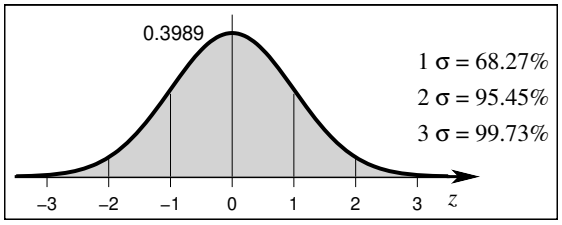


Figure 3.14: A normally distributed data set will have almost all of its values within $\pm 3\sigma$ of the mean (this corresponds to 99.73% of the data; see legend for percentages corresponding to other multiples of $\pm\sigma$).

Given the functional form of $p(z)$ we can evaluate the probability that an observation z will be $\leq a$:

$$P_c(a)=\int_{-\infty}^{a}p(z)=\int_{-\infty}^{0}p(z)+\int_{0}^{a}p(z)=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\int_{0}^{a}e^{-\frac{z^2}{2}}dz. \tag{3.81}$$

Let

$$u^2=\frac{z^2}{2},\ \text{hence}\ dz=\sqrt{2}du, \tag{3.82}$$

then

$$P_c(a) = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{\frac{a}{\sqrt{2}}} e^{-u^2} du = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} \operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{a}{\sqrt{2}}\right)\right]. \tag{3.83}$$

It follows that, for any value z, the cumulative distribution function is

$$P_c(z) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)\right]. \tag{3.84}$$

Here, erf represents the *error function* and it is defined by the definite integral in (3.83) and tabulated in Table A.1. Furthermore, the probability that z falls between a and b must necessarily be

$$P_c(a \le z \le b) = P_c(b) - P_c(a) = \frac{1}{2}\left[\operatorname{erf}\left(\frac{b}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{a}{\sqrt{2}}\right)\right]. \tag{3.85}$$

Example 3–11. Investigations into the strength of olivine have provided estimates of Young's modulus (E) that follow a normal distribution given by $\mu = 1.0 \cdot 10^{11}$ Pa and $\sigma = 1.0 \cdot 10^{10}$ Pa. What is the probability that a single estimate E will lie in the interval $9.8 \cdot 10^{10}$ Pa $< E < 1.1 \cdot 10^{11}$ Pa? We convert the limits to normal scores and find they correspond to the interval $-0.2 \le z \le 1.0$. Using these values for a and b in (3.85) (or using Table A.1) we find the probability to be 0.4206.

Approximate binomial distribution

Like the Poisson distribution, the normal distribution may also serve as an approximation to the binomial distribution when n is large. More specifically, this approximation holds when both np and $(1-p)n$ exceed 5. Under those circumstances, the mean and standard deviation of the approximate normal distribution become

$$\mu = np, \quad \sigma = \sqrt{np(1-p)}, \tag{3.86}$$

leading to the simplified distribution

$$P_b(x) = \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{\frac{-(x-np)^2}{2np(1-p)}\right\}. \tag{3.87}$$

Example 3–12. What is the probability that at least 70 of 100 sand grains will be larger than 0.5 mm if the probability that any single grain is that large is $p = 0.75$? Using the approximation (3.86) we find $\mu = np = 75$ and $s = \sqrt{np(1-p)} = 4.33$. Converting 69.5 (halfway between 69 and 70) to a z score gives -1.27, and we find via Table A.1 that the probability becomes 0.898 or about 90%.

3.4.2 The exponential distribution

Another important probability density distribution is the *exponential* distribution. It is given by

$$p_e(x) = \lambda e^{-\lambda x} \tag{3.88}$$

for some constant λ. However, most of the time we will see it used as a cumulative distribution function:

$$P_c(x) = 1 - e^{-\lambda x}. \tag{3.89}$$

Eq. (3.89) gives the probability that the observation a will be in the range $0 \le a \le x$.

Example 3–13. It has been reported that the heights (z) of Pacific seamounts follow an exponential distribution defined as

$$P_c(z \leq h) = 1 - e^{-h/340}, \tag{3.90}$$

which gives the probability that a seamount is shorter than h meters. Equation (3.90) then predicts that we might expect that

$$P_c(1000) = 1 - e^{-1000/340} \approx 95\% \tag{3.91}$$

of them are less than one km tall.

3.4.3 Log-normal distribution

Many data sets, such as grain-sizes of sediments and geochemical concentrations, have very skewed and long-tailed distributions (e.g., Figure 3.15). In general, such distributions arise when the observed quantities have errors that depend on *products* rather than *sums*. It therefore follows that the *logarithm* of the data may be normally distributed. Hence, taking the logarithm of your data may make the transformed distribution look normal. If this is the case, you can apply standard statistical techniques applicable to normal distributions to the logarithm of your data and convert the results (e.g., mean, standard deviation) back to get the proper units. The log-normal probability density distribution is therefore given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2}. \tag{3.92}$$

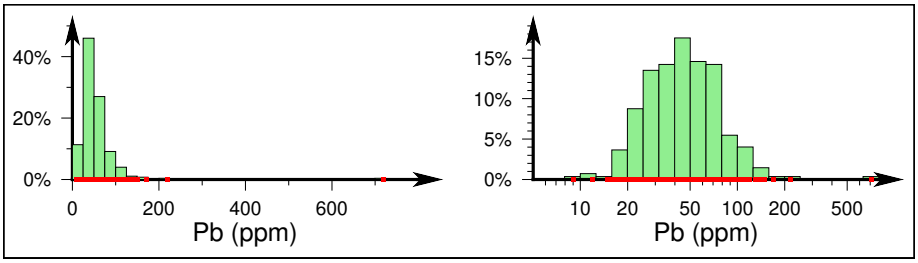


Figure 3.15: (left) The concentration of Pb in soil is very long-tailed and clearly not normally distributed. The red squares indicate individual sample values. (right) The same distribution after taking the logarithm of the data values. The resulting distribution is approximately normal, hence a log-normal distribution might be suitable to describe the data.

3.5 Inferences about Means

The central limits theorem states that the mean of a large sample taken from any distribution will be normally distributed even if the data themselves are not normally distributed, and furthermore it says that the sample mean is an unbiased estimator of the population mean. We can then use our knowledge of the normal distribution to quantify our faith in the precision of our sample mean. We already know that $s_{\bar{x}} = \sigma/\sqrt{n}$, so we can state with probability $1 - \alpha$ that $\bar{x}$ will differ from μ by at most E, which is given by

$$E = z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}, \tag{3.93}$$

where s is our estimate of σ. In other words, the chance that $\bar{x}$ exceeds the $\pm z_{\alpha/2}$ confidence interval is α. These error estimates apply to large samples ($n \geq 30$) and infinite populations. In those cases we can use our sample standard

deviation s in place of σ, which we usually do not know. Here, (3.93) can be inverted to yield the sample size necessary to be confident that the error in our sample mean is no larger than E, and we find

$$n = \left(\frac{z_{\alpha/2} \cdot s}{E}\right)^2. \tag{3.94}$$

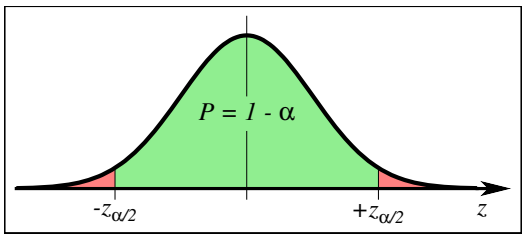


Figure 3.16: Probability is α that a value will fall in one of the two tails of the normal distribution, and $\alpha/2$ that it will fall in a specific tail.

The normal score for our sample mean is

$$z = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{\bar{x} - \mu}{s/\sqrt{n}}. \tag{3.95}$$

Since this statistic is normally distributed we know that the probability is $1-\alpha$ that z will take on a value in the interval $-z_{\alpha/2} < z < +z_{\alpha/2}$. Plugging in for the limits on z,

$$-z_{\alpha/2} < \frac{\bar{x} - \mu}{s/\sqrt{n}} < +z_{\alpha/2} \tag{3.96}$$

or

$$\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}. \tag{3.97}$$

Rearranging, we find

$$\mu = \bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}. \tag{3.98}$$

Eq. (3.98) shows the *confidence interval* on μ at the $1-\alpha$ confidence level. Very often, our confidence levels will be 95% ($\sim 2\sigma$) or 99% ($\sim 3\sigma$).

3.5.1 Small samples

The previous section dealt with large ($n \geq 30$) samples, where we could assume that $\bar{x}$ would be normally distributed as dictated by the central limits theorem. For smaller samples we must assume instead that the *population we are sampling* is normally distributed. We can then base our inferences on the statistic

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{\bar{x} - \mu}{s/\sqrt{n}}, \tag{3.99}$$

whose distribution is called the *Student's* t-distribution (Figure 3.17). It is similar to the normal distribution but its shape depends on the degrees of freedom, $\nu = n-1$. For large n (and hence ν) the t statistics approach the z statistics. As for z statistics, one can find tables with t values for various combinations of confidence levels and degrees of freedom (see Table A.2). For insight into what the t-distribution and others really are, get *Numerical Recipes* by Press et al. This excellent book gives both theory and computer code (in C++, C, FORTRAN, Java and a host of legacy languages).

Example 3–14. Given our sandstone density estimates from earlier, i.e., $\{2.2, 2.25, 2.25, 2.3, 2.3, 2.3, 2.35\}$, what is the 95% confidence interval on the population mean?

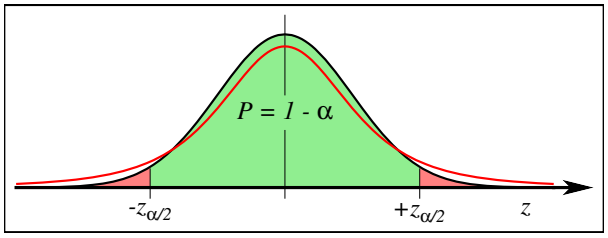


Figure 3.17: The same normal distribution and critical tails as in Figure 3.16, overlain by the Student's t-distribution for $\nu = 3$ degrees of freedom (red line). For small samples the probability distribution becomes wider.

Answer: We have $\bar{x} = 2.28$ with $s = 0.05$, and $\alpha = 1 - 95\% = 0.05$. The degrees of freedom $\nu = n - 1 = 6$. Table A.2 gives $t_{\alpha/2,\nu} = t_{0.025,6} = 2.447$. Using (3.97), we find our sample mean brackets the population mean, thus (with $t_{\alpha/2}$ instead of $z_{\alpha/2}$ and s instead of σ)

$$2.28 - 2.447 \cdot \frac{0.05}{\sqrt{7}} < \mu < 2.28 + 2.447 \cdot \frac{0.05}{\sqrt{7}} \tag{3.100}$$

or (since the bounds are symmetrical)

$$2.234 < \mu < 2.326 \tag{3.101}$$

or

$$\mu = 2.280 \pm 0.046. \tag{3.102}$$

3.6 Problems for Chapter 3

Problem 3.1. A student prepares for an exam in a data analysis class by studying a list of 10 specific topics. She is confident that she can answer any question related to six of these topics, but is ill prepared to handle the remaining four. For the exam, the instructor selects five topics at random from the same list of 10 topics. What is the probability that the student can solve all five problems on the exam?

Problem 3.2. During an expedition to Antarctica a research team collects 22 oriented rock cores to be used to determine the paleo-magnetic field. Among the 22 samples there are 7 of utmost importance: three are from an exciting new basaltic outcrop and four were recovered from another area with no prior samples. During the flight back to Punta Arenas the Principal Investigator samples too much Chilean wine and proceeds to trip over the box with the rock samples, causing 8 of the rock cores to fall out and break.

a) What is the probability of total disaster (i.e., all 7 important samples are destroyed)?

b) What is the probability that the 7 samples are all intact?

c) What is the probability that *at least* two samples from each of the two exciting areas have been ruined?

Problem 3.3. Returning from extensive fieldwork in the Congo, a biologist calmly discovers that the glue behind the labels on her glass specimen jars with spiders has dissolved and all the labels have separated from their corresponding jars. Of a total of 28 specimens, 8 of the specimens are previously unknown spiders; the remaining 20 are known to be harmless to humans. Given past experience that half of newly discovered spiders are venomous, what is the probability that, in randomly selecting four specimens:

a) One of the specimens is a venomous spider?

b) All four are nonvenomous?

Note: Unlike the biologist, the technician making the selection has no knowledge of spider species.

Problem 3.4. Tuco owns a tire manufacturing plant in South America that makes automobile tires for American-produced cars. For each batch of 100 tires, Tuco's quality control team goes to work: They randomly select four tires, mount them on a vehicle, and give each tire a solid kick. If any one of the four tires fall apart then the whole batch is sent back for reprocessing. How many defective tires can one batch have and still have at least a 50/50 chance of passing the test? [Hint: Plot the probability of passing the test as a function of the number of defective tires and graphically find the answer.]

Problem 3.5. A manufacturer of compasses used in geological field mapping has three physical plants that make the compasses. Plant A produces 55%, plant B 30%, and plant C 15% of the total production. If 0.4 % of the compasses from plant A are defective, and the corresponding numbers for plants B and C are 0.6 % and 1.2%, what is the probability that a defective compass purchased by mail-order was produced by plant A?

Problem 3.6. In the poor kingdom of Parador the levels of pollutants in water wells are measured using older, imprecise instruments. Here, 25% of all wells have excessive amounts of pollutants. When tested, 99% of all wells that have excessive amounts of pollutants will fail the test, but 17% of the wells that do *not* have excessive amounts of pollutants will also fail due to the poor quality of the instruments. What is the probability that a well failing the test actually has excessive amounts of pollutants?

Problem 3.7. A sample of 64 specimens of a particular fossil gives a mean length of 52.8 mm, with a standard deviation of 4.5 mm. Find the 99% confidence intervals for the mean length.

Problem 3.8. An expedition measuring the heat flux, q, out of the seafloor near a mid-ocean ridge returned with 13 independent measurements (in mW m^{-2}), given by

$$q = \{45.2, 47.4, 55.1, 39.2, 51.2, 46.3, 49.9, 42.9, 75.3, 53.1, 48.8, 58.8, 42.2\}.$$

a) Estimate the sample mean and sample standard deviation, as well as the median and the median absolute deviation (MAD).

b) Using robust scores, are there any outliers? If so, identify them, remove them from the data, and redo the answers from (a).

Problem 3.9. Walter, a youngster growing up in rural North Dakota, spends two winter weeks in bed with chicken pox. To stave off boredom he decides to measure the lowest temperatures (in °C) reached during each night. He obtains the following series of 14 measurements:

-51.32	-49.34	-42.41	-56.72	-45.92	-50.33	-47.09
-53.39	-24.23	-27.41	-44.21	-48.08	-39.08	-54.02

a) Estimate the sample mean and sample standard deviation, as well as the median and the median absolute deviation (MAD).

b) Using robust scores, are there any outlying data that suggest Walter's fever may have affected his measurements? If so, identify them, remove them from the data, and redo your analysis.

Problem 3.10. The data set *depths.txt* contains the bathymetric depths (in meters) for part of an ocean basin.

a) Find the mean depth, the standard deviation, and the 95% confidence interval on the mean depth.

b) What is the probability that a random depth measurement will be shallower than -4000m?

c) Determine the median and median absolute deviation (MAD).

d) Given the criteria for outliers ($|z_i| > 2.5$) using the median and MAD, find the robust confidence limits in meters. How many measurements are considered outliers?

Problem 3.11. Previous sampling of the salinity in tap water from a local water company has revealed that the salinity content is well described as following a normal distribution, with $\mu = 10$ ppm and $\sigma = 5$ ppm. When randomly testing households in this neighborhood for the salinity of the drinking water, what is the probability that the salinity will be:

a) Less than 4 ppm?

b) Between 8 and 16 ppm?

Make sure to illustrate the two cases graphically.

Problem 3.12. In Texas, past experience has shown that, on average, only one in 10 exploratory wells drilled discovers oil. Let n be the number of holes drilled until the first success (i.e., oil is struck). Assume that the exploratory wells represent independent events.

a) Find $P(1)$, $P(2)$, and $P(3)$.

b) Derive a formula for $P(n)$.

c) Plot $P(n)$ for $n = 1$ to 30.

d) Find $P_c(n)$, the cumulative probability that we will find oil in *n or less* tries, and plot it.

e) How many holes should we expect to drill in order to have a 90% probability of finding oil?

Chapter 4

TESTING OF HYPOTHESES

"The great tragedy of science – the slaying of a beautiful hypothesis by an ugly fact."

Thomas Huxley, Biologist

Much of statistics is concerned with testing hypotheses for certain properties of entire populations based on sample data taken from these populations. There are numerous standard techniques used to perform these tests, and we will broadly group them into two sets:

1. Parametric tests
2. Nonparametric tests

This chapter will discuss the assumptions behind both kinds of tests and how they are performed.

4.1 The Null Hypothesis

At the core of all tests lies the concept of the "null hypothesis" (also known informally as the "boring hypothesis"). The null hypothesis, denoted H_0, is stated and we will use our tests to see if we can reject it. For instance, if we want to test whether two rock types (i.e., two separate populations) have different densities, we obtain samples from each population and form the null hypothesis that they have equal densities (a boring result); we then test if we can reject H_0. We will illustrate this approach with an example:

Example 4–1. It is claimed that the density of a particular sandstone unit is 2.35 g cm^{-3}. We are handed a sample of 50 specimens from an outcrop in the same area and decide to set the criteria that the sample comes from another lithological unit if the sample mean is less than 2.25 or larger than 2.45. In other words, our null hypothesis H_0 is $2.25 < \mu < 2.45$. This statement is a clear-cut criterion for accepting or rejecting the claim that the sample originates from the same unit, but it is not infallible. Since our decision will be based on a sample, there is the possibility that the sample mean $\bar{x}$ may indeed satisfy $\bar{x} < 2.25$ or $\bar{x} > 2.45$ *even though* the population mean μ *is* 2.35. We will therefore want to know what the chances are that we could make a wrong decision and reject H_0. Clearly, we must investigate what is the probability that $\bar{x} < 2.25$ or $\bar{x} > 2.45$ when μ in fact is 2.35. Here, $s = \sigma = 0.42$. This probability is given by the area under the two tails in Figure 4.1.

Since $n = 50 \gg 30$, we will treat our sample as of infinite size. We find the uncertainty in the sample mean to be

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{0.42}{\sqrt{50}} = 0.06. \tag{4.1}$$

We can now evaluate the normal scores of the two limits as

$$z_0 = \frac{2.25 - 2.35}{0.06} = -1.67, \tag{4.2}$$

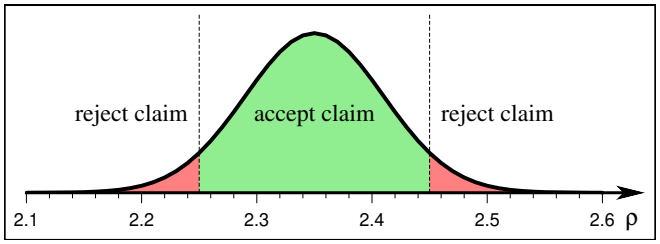


Figure 4.1: Type I error: Possibility of erroneously rejecting a correct hypothesis.

$$z_1 = \frac{2.45 - 2.35}{0.06} = +1.67. \tag{4.3}$$

We find the area under each tail to be $\frac{\alpha}{2} = \frac{1}{2}\left[1 + \text{ erf}\left(-1.67/\sqrt{2}\right)\right] = 0.0475$. Consequently, the probability of getting a sample mean that falls in the distal tails of the distribution is

$$\alpha = 2 \cdot 0.0475 = 0.095 \text{ or } 9.5\%. \tag{4.4}$$

This result means there is a 9.5% chance we will erroneously reject the hypothesis that $\mu = 2.35$ when it is in fact *true*. In statistics, we say we have committed a *type I error*[1].

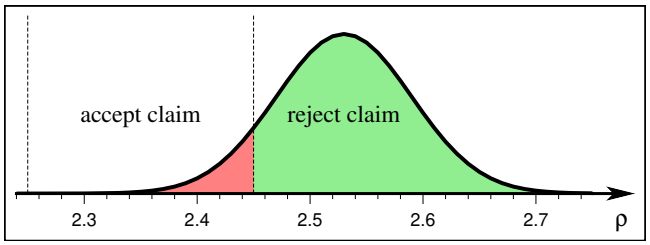


Figure 4.2: Type II error: Possibility of erroneously accepting an incorrect hypothesis.

Let us look at another possibility, where our test will fail to detect that μ is *not* equal to 2.35.

Example 4–2. Suppose, for the sake of argument, that the true mean $\mu = 2.53$. Then, the probability of getting a sample mean in the range $2.25 - 2.45$, and hence erroneously accept the claim that $\mu = 2.35$, is now given by the tail area in Figure 4.2. As before, $s_{\bar{x}} = 0.06$ so the normal scores become

$$z_0 = \frac{2.25 - 2.53}{0.06} = -4.67, \qquad z_1 = \frac{2.45 - 2.53}{0.06} = -1.333. \tag{4.5}$$

It follows that the probability $\beta = \frac{1}{2}\left[\text{erf}\left(-1.333/\sqrt{2}\right) - \text{ erf}\left(-4.67/\sqrt{2}\right)\right] = 0.092$ or 9.2%. This is the risk we run of accepting the incorrect hypothesis H_0. We call this committing a *type II error*.

Therefore, we recognize that there are several possible outcomes when testing a null hypothesis, as shown in Table 4.1. If the hypothesis is true, but is rejected, we have committed a Type I error, and the probability of doing so is designated α. In our example, α was 0.095. If our hypothesis is incorrect, but we still accept it, then we have committed a Type II error, and the probability of doing so is designated β. In our case, with $\mu = 2.53, \beta$ was 0.092.

We saw in our example that the type II error probability *depended on the value of* μ. Since μ is typically not known and we cannot evaluate β, it is common to simply either reject H_0 or *reserve judgment* (i.e., never accept H_0). This way we avoid committing a type II error altogether, at the expense of never accepting H_0. We call this a *significance*

[1] Its name suggests there are more ways to make mistakes...

	Accept H_0	**Reject** H_0
H_0 is true	Correct Decision	Type I Error
H_0 is false	Type II Error	Correct Decision

Table 4.1: The four possible decision scenarios when testing an hypothesis. Of these, we always seek to avoid making a Type II error.

test and say that the results are *statistically significant* if we can reject H_0. If not, the results are *not* statistically significant, and we attempt no further decisions. Of course, we may be wrong in rejecting H_0 but we can always state the likelihood of being wrong as α. Hence, in statistics we can only disprove hypotheses, but never prove them.

4.2 Parametric Tests

Parametric tests are used to make decisions based on parameters derived by assuming the data are approximately described by a known probability distribution. The parameters typically used are properties of the distribution, such as the mean and standard deviations. Note that while we obtain our statistical parameters from the *sample*, our hypothesis testing applies to the *parent* population.

4.2.1 What are the degrees of freedom?

When dealing with statistical tests we encounter the concept of "degrees of freedom" (usually denoted by the variable ν) and it is often confusing to know what it is for various cases. We need to understand what ν is in order to perform the tests. Typically, we will compute a *statistic* from our sample data. It may represent our best estimate of one *parameter* of the theoretical population distribution we are investigating via our limited sample. For instance, we may wish to compute a statistical quantity that requires us to use the mean of our data set of n points. While we will initially use $\nu = n$ in evaluating the mean (since all the points are independent), once we use the mean in *subsequent* calculations we must reduce ν by one, since any individual data point can now be obtained from that mean and any $n-1$ data points. In general, we say we lose one degree of freedom for each parameter we have estimated from the data. We will also see that certain transformations of the data, say, grouping the data into a set of k bins, will also affect ν. While you may have had n data points to start with, after binning you now only have k items (the bin counts), so any testing involving these bins will have to consider a ν that may only be $k-1$ (you still lose one since the bin counts must sum to n). However, the degrees of freedom may be even less than $k-1$ if any other parameters had to be estimated in order to facilitate the binning in the first place. In this Chapter, we will illuminate these situations using relevant examples.

4.2.2 Differences between sample means (equal variance)

We will often want to know if an observed difference between two sample means can be attributed to chance. We will again rely on the Student's t distribution. Here, it is assumed that the two populations have the *same variance* but possibly *different means*.

We are interested in the distribution of $\bar{x}_1 - \bar{x}_2$, the difference in sample means. Of course, we only have one such estimate, but if the two samples are independent and random then the hypothetical difference distribution will be approximately normal, with mean $\mu_1 - \mu_2$ and standard deviation

$$\sigma_e = \sigma_p \sqrt{(1/n_1 + 1/n_2)}. \tag{4.6}$$

Here, σ_p is the *pooled* standard deviation,

$$\sigma_p = \sqrt{\frac{(n_1-1)\sigma_1^2 + (n_2-1)\sigma_2^2}{n_1+n_2-2}}. \tag{4.7}$$

We find the t-statistic by evaluating

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{\bar{x}_1 - \bar{x}_2}{s_e} \tag{4.8}$$

and test the hypothesis H_0: $\mu_1 = \mu_2$ based on the t-distribution for $\nu = n_1 + n_2 - 2$ degrees of freedom. Here, we have lost one degree of freedom for each sample mean we computed. For large n_1, n_2, the t-distribution becomes very close to a normal distribution and we may instead use z-statistics based on

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}. \tag{4.9}$$

We will illustrate the two-sample t-test with an example.

Example 4–3. We have obtained random samples of magnetite-bearing rocks from two separate basalt outcrops. The measured magnetizations (in $\mathrm{Am^2kg^{-1}}$) are

Outcrop 1: $\{87.4, 93.4, 96.8, 86.1, 96.4\}$ $n_1 = 5$

Outcrop 2: $\{106.2, 102.2, 105.7, 93.4, 95.0, 97.0\}$ $n_2 = 6$

We state our null hypothesis H_0: $\mu_1 = \mu_2$; the alternative hypothesis is of course $\mathrm{H}_1 : \mu_1 \neq \mu_2$. We decide to use a 95% confidence level, so the level of significance $\alpha = 0.05$. In this case, $\nu = 5 + 6 - 2 = 9$, and Table A.2 shows that the critical t value is 2.262. We will reject H_0 if our calculated t exceeds this critical value. Using the data, we find

$$\bar{x}_1 = 92.0 \text{ with } s_1 = 5.0, \tag{4.10}$$

$$\bar{x}_2 = 99.9 \text{ with } s_2 = 5.5. \tag{4.11}$$

Using (4.8) we obtain

$$t = \frac{99.9 - 92.0}{\sqrt{\frac{4 \cdot 5^2 + 5 \cdot 5.5^2}{9}\left(\frac{1}{5} + \frac{1}{6}\right)}} = 2.5. \tag{4.12}$$

Since $t > 2.262$ we must reject H_0. We conclude that the magnetizations at the two outcrops are not the same.

4.2.3 Differences between sample means (unequal variance)

In situations where the two populations *do not* have equal variance (which we will test for in Section 4.2.6) one can use *Welch's* t-test instead. It is similar to the regular two-sample t-test but now the observed t value is obtained directly from (4.9). However, the estimation of the degrees of freedom differs and is given by

$$\nu \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2 \nu_1} + \frac{s_2^4}{n_2^2 \nu_2}}. \tag{4.13}$$

Rounding ν to the nearest integer allows you to look up critical t from Table A.2.

So far, we have put confidence limits on sample means and compared sample means to investigate whether two populations have different means. We will now turn our attention to inferences about the standard deviation.

4.2.4 Inferences about the standard deviation

The most popular way of estimating σ is to compute the sample standard deviation. When investigating properties of σ using s as a proxy we will rely on the "chi-square" statistic, given by

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}. \tag{4.14}$$

The χ^2 distribution depends on the degrees of freedom, $\nu = n - 1$, and it is restricted to positive values because of the squared terms. It portrays how the sample variances would be distributed if we selected random samples of n items from a normal distribution with a standard deviation of σ. Figure 4.3 shows a typical χ^2 distribution. In the same way

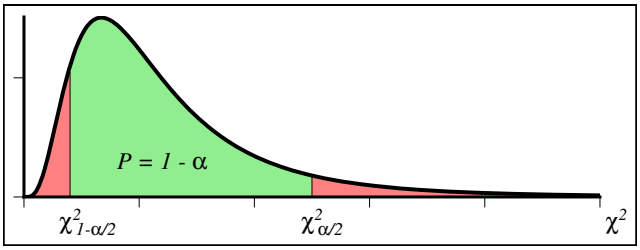


Figure 4.3: A typical chi-square probability density function, with mean value ν and mode at $\nu - 2$.

we used z_α and t_α, we now use χ^2_α as the value for which the area to the right of χ^2_α equals α. Because the distribution is not symmetrical, we must evaluate the critical values for $\alpha/2$ and $1-\alpha/2$ separately. Furthermore, in the same way we placed confidence intervals on μ, we now use (4.14) to place confidence intervals on the variance, i.e.,

$$\chi^2_{1-\frac{\alpha}{2}} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}} \tag{4.15}$$

or

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}, \tag{4.16}$$

which gives the α confidence interval on the variance. For large samples ($n > 30$) this can be simplified to

$$\frac{s}{1+z_{\frac{\alpha}{2}}/\sqrt{2n}} < \sigma < \frac{s}{1-z_{\frac{\alpha}{2}}/\sqrt{2n}}. \tag{4.17}$$

Note that the confidence interval is *not* symmetrical about the sample standard deviation.

4.2.5 Testing a sample standard deviation

We might want to test whether our sample standard deviation s is equal to or different from a given population standard deviation σ. In such a case the null hypothesis becomes $H_0 : s = \sigma$, with the alternative hypothesis $H_1 : s \neq \sigma$. As usual, we select our level of significance to be α = 0.05.

Example 4–4. We have 15 estimates of temperatures with $s = 1.3^\circ$ C and we want to know if s is significantly different from $\sigma = 1.5^\circ$ C based on past experience. From α = 0.05 and ν = 14 (since we lose one degree of freedom by first computing $\bar{x}$ to obtain s), we find the critical χ^2 values from Table A.3 to be $\chi^2_{0.975} = 5.63$ and $\chi^2_{0.025} = 26.119$. Based on our sample statistic, we compute

$$\chi^2 = \frac{14 \cdot 1.3^2}{1.5^2} = 10.5. \tag{4.18}$$

We see that we cannot reject H_0 at the 0.05 significance level, so we simply reserve judgment. This was a *two-sided* test since we had to check that χ^2 did not land in either of the two tails.

For large samples ($n \geq 30$), the χ^2 critical values do not vary much with ν and we may use the simpler statistic

$$z = \frac{s-\sigma}{\sigma/\sqrt{2n}} \tag{4.19}$$

and use the standard z-statistics, i.e., the $n = \infty$ entry in Table A.2.

4.2.6 Differences between sample standard deviations

In the t-test for differences between two means (Section 4.2.2) we *assumed* that the standard deviations of the two samples were the same. Often this is not the case and one should first test whether this assumption is valid. We want

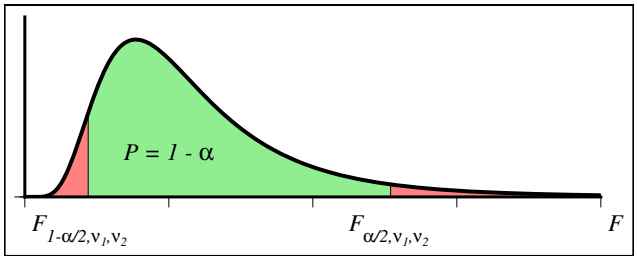


Figure 4.4: A typical F probability density function, for $\alpha = 0.10$, $\nu_1 = 20$ and $\nu_2 = 12$.

to know whether the two variances are different or not. The statistic that is most appropriate for such tests is called the F-statistic, defined as the ratio

$$F = \begin{cases} {s_1}^2/{s_2}^2 & , s_1 > s_2 \\ {s_2}^2/{s_1}^2 & , s_2 > s_1 \end{cases} . \tag{4.20}$$

An example of this distribution as shown in Figure 4.4. For normal distributions, this variance ratio is a continuous distribution called the F-distribution. It depends on the two degrees of freedom, $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$. As before, we will reject the null hypothesis $H_0 : \sigma_1 = \sigma_2$ at the α level of significance and (possibly) entertain the alternative H_1: $\sigma_1 \neq \sigma_2$ when our observed F-statistic exceeds the critical value $F_{\alpha/2,\nu_1,\nu_2}$. Note that because of the way (4.20) forces F to be equal to or larger than unity, any F-ratio that would have been located into the left tail is now instead mapped to the right tail. This is why we use $\alpha/2$ in determining the critical value as it is still a *two-sided* test.

Example 4–5. In our case of rock magnetizations we assumed that the two standard deviations were approximately the same. Let us now show that this was actually justified. We find our observed statistic to be

$$F = \frac{5.5^2}{5.0^2} = 1.21. \tag{4.21}$$

From Table A.6 we find $F_{0.025}(\nu_1 = 5, \nu_2 = 4) = 9.36$. Hence, we cannot reject H_0 and conclude instead that the difference in sample standard deviations is not statistically significant at the 95% level. In fact, s_1 would have to be over three times larger than s_2 ($\sqrt{9.36} = 3.06$ to be exact) before we would be able to reject H_0.

4.2.7 Testing distribution shape: The χ^2 test

The next parametric test we shall be concerned with is the chi-squared test. It is a sample-based statistic using normal scores that are squared and summed up:

$$\chi^2 = \sum_{i=1}^{n} z_i^2 = \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s} \right)^2 . \tag{4.22}$$

If we draw all possible samples of size n from a normal population and plotted a histogram of the resulting $\sum z^2$ they would approximate the χ^2 distribution mentioned earlier. The χ^2 test is used to compare the *shape* of our data distribution to a distribution of known shape (which is usually a normal distribution but need not be). The χ^2 test is most often used on data that have been categorized or *binned*. Assuming that our observations have been binned into k bins, the test statistic is found as

$$\chi^2 = \sum_{j=1}^{k} \frac{(O_j - E_j)^2}{E_j}, \tag{4.23}$$

where O_j and E_j are the number of observed and expected values in the j'th bin. At first glance you may think we have a typographical error in (4.23) — we do not. Note that this χ^2 still is nondimensional since we are using counts, even if the denominator is not squared. With counts, the probability that m out of n counts will fall in a given bin j is determined by the binomial distribution approximated by (3.3.1), with

$$\bar{x}_j = E_j = np_j \tag{4.24}$$

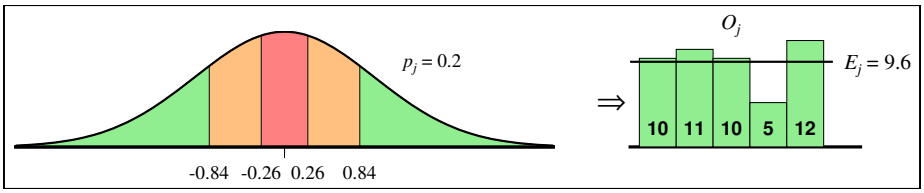


Figure 4.5: By desiring equal probabilities ($p = 0.2$) we obtain z-values for the bin boundaries that are not equidistant. This ensures that each bin has an acceptable number of observations and simplifies the calculation of the expected values.

and

$$s_j = \sqrt{np_j(1-p_j)} \approx \sqrt{np_j} = \sqrt{E_j}. \tag{4.25}$$

Here, p_j is the probability that any value will fall in the j'th bin. Plugging in for χ^2, we find

$$\chi^2 = \sum_{j=1}^{k} \left(\frac{x_j - \bar{x}_j}{s_j} \right)^2 = \sum_{j=1}^{k} \left(\frac{O_j - E_j}{\sqrt{E_j}} \right)^2 = \sum_{j=1}^{k} \frac{(O_j - E_j)^2}{E_j}. \tag{4.26}$$

as stated in (4.23). Since $\chi^2 = 0$ would mean a perfect match between observations and expectations we realize that this χ^2 test is *one-sided*, hence we will only check if our observed χ^2 statistic exceeds the critical $\chi^2_{\alpha,\nu}$ value.

Note that due to the normal approximation to the binomial distribution used to derive (4.23), we must ensure that the expected counts, E_j, for each bin are all 5 or greater. Depending on the size of your data set this may require some careful decisions on the number of bins you can use, as we shall see in our next example.

Example 4–6. Consider the 48 measurements of salinity from Whitewater Bay in Florida (Table 4.2). We would like to know if these observations came from a normal distribution or not. The answer might have implications for models of mixing salt and fresh water in the bay.

The first step is to transform the data into normal scores. We find $\bar{x} = 49.54$ and $s = 9.27$, and thus convert all values via

$$z_i = \frac{x_i - 49.54}{9.27}. \tag{4.27}$$

We choose to bin the data into five bins whose boundaries were chosen so that the area under the curve for each bin is the same, i.e., 0.2. This choice ensures that the expected value will be the same for all bins (Figure 4.5). Using Table A.1, we find that the corresponding z-values for the intervals are (-∞, -0.84), (-0.84, -0.26), (-0.26, +0.26), (0.26, 0.84), and (0.84, ∞). Counting the values in Table 4.2 we find the observed number of samples for each of the five bins to be 10, 11, 10, 5, and 12. These are the observed O_j's. The expected values E_j are all the same, i.e.,

$$E_j = \frac{n}{k} = \frac{48}{5} = 9.6, \tag{4.28}$$

which satisfies the binomial approximation criteria above. Using (4.23) we find the observed value $\chi^2 = 3.04$.

The χ^2 distribution depends on ν, the degrees of freedom, which normally would be $\nu = k - 1 = 4$ in our case (we lose one since the bin counts must sum to n). However, we also used our observations to compute $\bar{x}$, then s, in order to *determine the bin boundaries*. These estimations further reduce ν by two, leaving just two degrees of freedom. From the relevant Table A.3 we find critical χ^2 for $\nu = 2$ and $\alpha = 0.05$ to be 5.99. Since this is much larger than our computed value we conclude that we cannot, at the 0.05 level of confidence, reject the null hypothesis that the salinities were drawn from a normal distribution. We stress that while we used a normal distribution for comparison in this example, the E_j could have been derived from any other distribution we may wish to compare our data to.

Sample Number	Original Sample	Standardized Sample	Sample Number	Original Sample	Standardized Sample
1	46.00	-0.38	25	35.00	-1.57
2	37.00	-1.35	26	49.00	-0.06
3	62.00	1.34	27	48.00	-0.17
4	59.00	1.02	28	39.00	-1.14
5	40.00	-1.03	29	36.00	-1.46
6	53.00	0.37	30	47.00	-0.27
7	58.00	0.91	31	59.00	1.02
8	49.00	-0.06	32	42.00	-0.81
9	60.00	1.13	33	61.00	1.24
10	56.00	0.70	34	67.00	1.88
11	58.00	0.91	35	53.00	0.37
12	46.00	-0.38	36	48.00	-0.17
13	47.00	-0.27	37	50.00	0.05
14	52.00	0.27	38	43.00	-0.71
15	51.00	0.16	39	44.00	-0.60
16	60.00	1.13	40	49.00	-0.06
17	46.00	-0.38	41	46.00	-0.38
18	36.00	-1.46	42	63.00	1.45
19	34.00	-1.68	43	53.00	0.37
20	51.00	0.16	44	40.00	-1.03
21	60.00	1.13	45	50.00	0.05
22	47.00	-0.27	46	78.00	3.07
23	40.00	-1.03	47	48.00	-0.17
24	40.00	-1.03	48	42.00	-0.81

Table 4.2: Standardized scores of salinity measurements from Whitewater Bay.

4.2.8 Test for a correlation coefficient

We recall that the conventional correlation coefficient was defined as

$$r = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}. \tag{4.29}$$

Often, we need to test if an observed r is *significant*. In such tests, r is our sample-derived estimate of ρ, the actual correlation of the population pairs. The most useful null hypothesis is simply $H_0 : \rho = 0$. It can be shown that the sampling distribution of correlations for a population that has zero correlation ($\rho = 0$) is a normal distribution with mean $\mu = 0$ and $\sigma = \sqrt{(1-r^2)/(n-2)}$. Hence, a t statistic can be calculated as

$$t = \frac{r - \mu}{\sigma} = \frac{r}{\sqrt{(1-r^2)/(n-2)}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}. \tag{4.30}$$

The degrees of freedom, ν, is $n-2$ since we needed to compute both $\bar{x}$ and $\bar{y}$ first to determine r.

Example 4–7. Suppose we roll a pair of dice, one red and one green (Table 4.3). Using (4.29) we obtain $r = 0.66$ which seems quite high, especially since there is no reason to believe a correlation should exist at all. Let us run a test to determine if the correlation is significant. Choosing $\alpha = 0.05$, Table A.2 shows that critical $t_{\alpha/2,3} = 3.182$. Applying (4.30) gives the observed $t = 1.52$, hence the correlation of 0.66 is most likely caused by random fluctuations found in small samples and (gratefully) we cannot reject H_0.

How high would r have to be for us to find it significant and commit a type I error by rejecting the (true) null hypothesis? We can solve for r in

$$t_{\alpha/2,n-2} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \Rightarrow 3.182^2 = \frac{3r^2}{1-r^2} \Rightarrow r = \pm 0.88. \tag{4.31}$$

Red (x)	**Green** (y)
4	5
2	2
4	6
2	1
6	4

Table 4.3: Result of tossing a pair of dice five times.

So, if r happens to equal or exceed ± 0.88 we would find ourselves awkwardly concluding that red and green dice give correlated pairs of values, but we would only make that mistake once in twenty tries on average.

Going a bit further, we can use (4.30) to determine what the critical correlation, r_c, has to be for a given sample size (i.e., number of rolls) before we would suspect the dice have been tampered with. We simply solve the equation for r given a level of confidence and sample size. The results are displayed in Figure 4.6. High correlations are required when the number of rolls are relatively small.

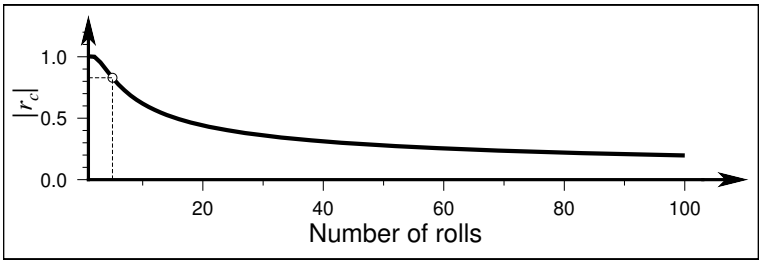


Figure 4.6: The critical values of correlation, i.e., how large an observed correlation would have to be before we must reject the null hypothesis. The circle and dashed lines reflect the situation in Example 4.7.

4.2.9 Analysis of variance

We found earlier that we could use the Student's t-test to decide if two samples had different means. However, very often we are faced with the task of deciding whether observed differences among *more* than two sample means can be attributed to chance, or whether there are real differences among the means of the populations sampled. Statistical scientists developed an exploratory procedure to analyze such observations known as *ANOVA*, which stands for "ANalysis Of VAriance". ANOVA usually comes in two flavors: One-way and two-way.

One-way ANOVA

In one-way ANOVA, the basic idea is to express the total variation of the data as a sum of two terms, each of which can be attributed to a specific *source*. The two sources of variation are:

1. Actual differences (because of different means) among the populations that the samples represent.
2. Chance or experimental error.

The measure of the total variation that we shall use here is the *total sum of squares*, *SST*, which is simply

$$SST = \sum_{j=1}^{k} \sum_{i=1}^{n} (x_{ij} - \bar{\bar{x}})^2, \tag{4.32}$$

where x_{ij} is the i'th observation from the j'th sample. Thus, we have k samples with n observations each and $\bar{\bar{x}}$ is the grand mean of all observations. Note that if we divided SST by $kn-1$ we would obtain the variance of the combined

data set. Let $\bar{x}_j$ be the mean of the j'th sample. By adding and subtracting $\bar{x}_j$ we can then rewrite (4.32) as

$$\begin{aligned} SST &= \sum_j^k \sum_i^n [(\bar{x}_j - \bar{\bar{x}}) + (x_{ij} - \bar{x}_j)]^2 = \sum_j^k \sum_i^n [(\bar{x}_j - \bar{\bar{x}})^2 + 2(\bar{x}_j - \bar{\bar{x}})(x_{ij} - \bar{x}_j) + (x_{ij} - \bar{x}_j)^2] \\ &= n\sum_j^k (\bar{x}_j - \bar{\bar{x}})^2 + \sum_j^k \sum_i^n (x_{ij} - \bar{x}_j)^2 + 2\sum_j^k \sum_i^n (\bar{x}_j x_{ij} - \bar{x}_j^2 - \bar{\bar{x}} x_{ij} + \bar{\bar{x}}\bar{x}_j) \\ &= n\sum_j^k (\bar{x}_j - \bar{\bar{x}})^2 + \sum_j^k \sum_i^n (x_{ij} - \bar{x}_j)^2 + 2\sum_j^k \bar{x}_j n \bar{x}_j - 2n\sum_j^k \bar{x}_j^2 - 2\bar{\bar{x}}\sum_j^k \sum_i^n x_{ij} + 2\bar{\bar{x}} n \sum_j^k \bar{x}_j \\ &= n\sum_j^k (\bar{x}_j - \bar{\bar{x}})^2 + \sum_j^k \sum_i^n (x_{ij} - \bar{x}_j)^2 + 2n\sum_j^k \bar{x}_j^2 - 2n\sum_j^k \bar{x}_j^2 - 2\bar{\bar{x}} k n \bar{\bar{x}} + 2\bar{\bar{x}} n k \bar{\bar{x}} \\ &= n\sum_j^k (\bar{x}_j - \bar{\bar{x}})^2 + \sum_j^k \sum_i^n (x_{ij} - \bar{x}_j)^2. \end{aligned} \tag{4.33}$$

Looking at the last two terms remaining we see that the first is a measure of the variation *among* the sample means. Similarly, the second term is a measure of the variation *within* the individual samples. It is customary to refer to the first term as the *treatment sum of squares*, $SS(Tr)$ and to the second term as the *error sum of squares*, SSE (Note: these names originated from agricultural experiments, i.e., "treatments" of different fertilizers.) Our overall goal with the ANOVA test is to investigate if the sample means are all statistically equal, hence the null hypothesis is

$$H_0 = \mu_1 = \mu_2 = \cdots = \mu_k \text{ versus } H_1 : \text{they are not all equal.}$$

We usually state that $\mu_i = \mu + \varepsilon_i$, so $\sum \varepsilon_i \equiv 0$. Thus, the null hypothesis can also be written,

$$H_0 : \varepsilon_1 = \varepsilon_2 = \cdots = \varepsilon_k = 0.$$

Let us introduce two assumptions that are critical to this test:

1. The populations we are sampling are approximately normal.
2. They all have the same variance, σ^2.

If true, then we can look upon the k samples as if they came from the same (normal) population and therefore consider the variance of their means $s_{\bar{\bar{x}}}^2$ as an estimate of $\sigma_{\bar{x}}$. Since $\sigma_{\bar{\bar{x}}} = \sigma/\sqrt{n}$ for infinite populations (recall the central limits theorem), we find that $ns_{\bar{\bar{x}}}^2 = \sigma^2$. This is, upon examination, the same as $SS(Tr)/(k-1)$.

Since σ^2 is not known it must be estimated from the data. Furthermore, since the populations are assumed to be normal then any one of the sample variances could be used. To improve on the estimate we chose to take the mean of the sample variances,

$$\sigma^2 \approx s^2 = \frac{s_1^2 + s_2^2 + \cdots s_k^2}{k}. \tag{4.34}$$

This turns out to equal the (normalized) second variance term, i.e., $SSE/(k(n-1))$. We now have two competing estimates of s^2 for the population variance σ^2. If the first estimate (based on variation among the sample means) is much larger than the second estimate (based on variations within the samples, i.e., variation due to chance) we should reject H_0. This is simply the F-test comparing two variances. Consequently, we shall use the statistic

$$F = \frac{SS(Tr)/_{k-1}}{SSE/_{k(n-1)}}. \tag{4.35}$$

where the numerator is our estimate of σ^2 from variation among the $\bar{x}_j$ while the denominator is our estimate of σ^2 from variation within the samples. The expected value of F is therefore unity. In practice, we construct an ANOVA table (Table 4.4).

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Treatments	$k-1$	$SS(Tr)$	$MS(Tr) = \frac{SS(Tr)}{k-1}$	$\frac{MS(Tr)}{MSE}$
Error	$k(n-1)$	SSE	$MSE = \frac{SSE}{k(n-1)}$	
Total	$kn-1$	SST		

Table 4.4: Table used for a one-way analysis of variance (ANOVA).

We will apply this procedure to a small data set of four samples with six observations each (Table 4.5).

Example 4–8. The porosity $\bar{\varphi}$ of a sandstone was measured at four different locales and the observed values are listed in Table 4.5. At the 0.05 level of significance, is the porosity the same at all four locations? The null hypothesis

Loc A	Loc B	Loc C	Loc D
23.1	21.7	21.9	19.8
22.8	23.0	21.3	20.4
23.2	22.4	21.6	19.3
23.4	21.1	20.2	18.5
23.6	21.9	21.6	19.1
21.7	23.4	23.8	21.9

Table 4.5: Four samples of sandstone porosity take from different locations.

becomes

$$H_0 : \bar{\varphi}_1 = \bar{\varphi}_2 = \bar{\varphi}_3 = \bar{\varphi}_4. \tag{4.36}$$

We first find

$$\bar{x}_1 = 22.97, \bar{x}_2 = 22.25, \bar{x}_3 = 21.73, \bar{x}_4 = 19.83, \bar{\bar{x}} = 21.70. \tag{4.37}$$

The resulting ANOVA table is given as Table 4.6. From Table A.5 we find the critical F value for 3 and 20 degrees

Source of Variation	ν	SS	MS	F
Treatments	3	32.35	10.78	10.78
Error	20	20.02	1.000	
Total	23	52.37		

Table 4.6: One-way ANOVA table based on the statistics from the porosity data given in Table 4.5.

of freedom at the 5% level to be $F_{0.05,3,20} = 3.10$ (This is a *one-sided* test since we only consider if $F > 1$ or not). Since our observed F is much larger than the critical value we must reject our null hypothesis: The porosity at the four locations are not all the same.

To simplify the calculations, we rewrite the expressions for the various sums of squares as

$$SST = \left\{ \sum_{j=1}^{k} \sum_{i=1}^{n} x_{ij}^2 \right\} - \frac{1}{kn} S^2, \tag{4.38}$$

$$SS(Tr) = \left\{ \frac{1}{n} \sum_{j=1}^{k} S_j^2 \right\} - \frac{1}{kn} S^2, \tag{4.39}$$

and

$$SSE = SST - SS(Tr). \tag{4.40}$$

Here, S_j is the sum of the values in the j'th sample, and S is the total sum of all observations.

So far we assumed that each sample has the same number of observations. Instead, if there are n_j observations in the j'th sample we get

$$SST = \left\{ \sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}^2 \right\} - \frac{1}{N} S^2, \tag{4.41}$$

$$SS(Tr) = \left\{ \sum_{j=1}^{k} \frac{S_j^2}{n_j} \right\} - \frac{1}{N} S^2, \tag{4.42}$$

where $S_j = \sum_{i=1}^{n_j} x_{ij}$ and $N = \sum_{j=1}^{k} n_j$. Also, $\nu_{\text{treat}} = k-1$ but $\nu_{\text{error}} = N-k$.

Two-way ANOVA

The previous ANOVA test was concerned only with the task of checking if the means of the populations were the same. The two-way ANOVA procedure extends the test to whether there also are variations *across* the populations. In other words, the population mean for j'th treatment and i'th (agricultural) *block* is expected to be

$$\mu_{ij} = \mu + \varepsilon_j + \gamma_i. \tag{4.43}$$

Thus, ε_j are the treatment effects that vary from sample to sample, and γ_i are called the *block effects* and vary within each sample. Examples might be porosity at four locations where the ε_j represent differences among the locations and γ_i may represent variations with depth across all samples. Again, we test

$$H_0 : \varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_k = 0,$$

but now we also consider the second null hypothesis

$$H_0 : \gamma_1 = \gamma_2 = \ldots = \gamma_n = 0.$$

To do this we must obtain a quantity, similar to the treatment sum of squares, which measures the variation among the block means. If we let S_i be the total of all values in the i'th block (e.g., depth) and substitute it for S_j, sum over i instead of j, and swap n and k, we find the *block sum of squares* via

$$SSB = \left\{ \frac{1}{k} \sum_{i=1}^{n} S_i^2 \right\} - \frac{1}{kn} S^2. \tag{4.44}$$

Hence, we compute SST and $SS(Tr)$ as before, SSB as just given, and SSE now becomes

$$SSE = SST - [SS(Tr) + SSB]. \tag{4.45}$$

Table 4.7 shows the extended ANOVA table for two-way analysis.

Source of Variation	**Degrees of Freedom**	**Sum of Squares**	**Mean Square**	F
Treatments	$k-1$	$SS(Tr)$	$MS(Tr) = \dfrac{SS(Tr)}{k-1}$	$\dfrac{MS(Tr)}{MSE}$
Blocks	$n-1$	SSB	$MSB = \dfrac{SSB}{n-1}$	$\dfrac{MSB}{MSE}$
Error	$(k-1)(n-1)$	SSE	$MSE = \dfrac{SSE}{(k-1)(n-1)}$	
Total	$kn-1$	SST		

Table 4.7: Two-way table for the analysis of variance (ANOVA).

Example 4–9. We have measured the nickel concentration in a shale at four locations where we have obtained three observations at different depths in the unit. Our data are given in parts per million (ppm), with depth increasing downward in Table 4.8. We want to do a two-way ANOVA to see if the variations among the locations and among observations at the same depth (our "blocks") are similar at the 95% level of confidence. We state

$$H_0 : \varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = 0,$$

$$\gamma_1 = \gamma_2 = \gamma_3 = 0.$$

	Loc 1	**Loc 2**	**Loc 3**	**Loc 4**	S_i
Depth 1	71	44	50	67	232
Depth 2	92	51	64	81	288
Depth 3	89	85	72	86	332
S_j	252	180	186	234	852

Table 4.8: Four samples of nickel concentrations from different locations, sorted by depth.

Following the procedure, we compute the total sum of squares to be

$$\sum\sum x^2 = 63,414. \tag{4.46}$$

Combined with the sums in Table 4.8 we find

$$SST = 63,414 - \frac{1}{12}(852)^2 = 63,414 - 60,492 = 2922, \tag{4.47}$$

$$SS(Tr) = \frac{1}{3}[252^2 + 180^2 + 186^2 + 234^2] - 60,494 = 1260, \tag{4.48}$$

$$SSB = \frac{1}{4}[232^2 + 288^2 + 332^2] - 60,492 = 1256, \tag{4.49}$$

$$SSE = 2922 - [1260 + 1256] = 406. \tag{4.50}$$

We construct a two-way ANOVA table presented as Table 4.9.

Source of Variation	ν	**SS**	**MS**	F
Treatments	3	1260	420	6.21
Blocks	2	1256	628	9.28
Error	6	406	67.67	
Total	11	2922		

Table 4.9: Two-way ANOVA table resulting from the statistics of the nickel concentrations given in Table 4.8.

With the help of Table A.5 we reach the following decisions:

Treatments (locations): Critical value $F_{0.05,3,6} = 4.76$, so we reject the hypothesis that $\varepsilon_i = 0$.

Blocks (depth): Critical value $F_{0.05,2,6} = 5.14$, so again we reject the hypothesis that all $\gamma_j = 0$.

In other words, we conclude that the average nickel concentration is not the same at the four locations, and that it is not the same at all depths.

4.3 Nonparametric Tests

The last section concluded the examination of standard parametric tests (i.e., the t-, F-, and χ^2-tests.) We justified using these tests by *either* having large samples and invoking the central limits theorem *or* simply assuming that the distribution we have sampled is approximately normal. Sometimes, however, none of these conditions are met. The two typical situations that can arise are:

1. You have a small sample ($n < 30$) and you *cannot* assume that the population it came from is normal.
2. You have *ordinal* data (which can be ranked, but not operated on numerically).

In those cases we must consider *nonparametric* methods, which make no assumptions about the shape of the data distribution. In particular, nonparametric tests *do not* involve the calculation of distribution parameters, such as the mean and standard deviation.

4.3.1 Sign test for the one-sample mean or median

The nonparametric sign test is a robust alternative to the standard one-sample t-test. It can be used when the distribution we have sampled has a continuous *symmetrical* population. This implies that the probability of getting a data value *less* than the mean is the same as getting one *larger* than the mean: both probabilities equal 0.5. However, if we cannot assume that the population is symmetrical, the test should instead apply to the median value rather than the mean. Since we will be testing whether or not our sample mean (or median) is statistically indistinguishable from a specified hypothetical mean (or median), the procedure relies on properties of the binomial distribution encountered in Chapter 3 and is reminiscent of the simple coin-toss analogy. Values may be less than or larger than the hypothetical mean (median), and the probability of finding x values out of n values to be less than the mean (median) follows directly from (3.70), with $p = 0.5$. To perform the sign test we need to evaluate the cumulative binomial distribution or consult pre-tabulated distributions.

Example 4–10. The following data constitute a random sample of 15 measurements of salinity content (in ppt):

97.5, 95.2, 97.3, 96.0, 96.8, 100.3, 97.4, 95.3, 93.2, 99.1, 96.1, 97.6, 98.2, 98.5, 94.9

We will use the one-sample sign test to consider the null hypothesis, H_0: $\tilde{\mu} \geq 98.5$ against the alternative hypothesis, H_1: $\tilde{\mu} < 98.5$ at the $\alpha = 0.01$ level of significance. Because of the inequality we have a *one-sided* test. We replace all values greater than 98.5 with a plus sign and all values less than 98.5 with a minus sign. Values that equal 98.5 exactly are discarded; in our case we lose one value, thus $n = 14$, resulting in the following series:

\- - - - - + - - - + - - - -

We find $x = 2$ values (represented by the two plus-signs) larger than the hypothetical median. The probability of finding $x \leq 2$ is given by the binomial distribution (3.70) by adding up the probabilities for $x = 0$, $x = 1$, and $x = 2$. We find

$$P = P_{14,0.5}(0) + P_{14,0.5}(1) + P_{14,0.5}(2) = C_{14,0.5}(2) = \sum_{x=0}^{2} P_{14,0.5}(x), \tag{4.51}$$

which evaluates to

$$P = \binom{14}{0}\frac{1}{2}^{14} + \binom{14}{1}\frac{1}{2}^{14} + \binom{14}{2}\frac{1}{2}^{14} = 0.00006 + 0.0009 + 0.0056 \approx 0.0065. \tag{4.52}$$

Since 0.0065 is less than 0.01, we must reject H_0. We conclude that the null hypothesis must be rejected as the data suggest that the median salinity from the sampled region is less than 98.5 ppt. Note that in this test we did not compute a critical value for x but compared the probability for the observed case with a specified probability α.

When both np and $n(1-p)$ are greater than 5 (here they are both equal to 7) we are allowed to use the normal approximation to the binomial distribution. Per (3.87), the sign test may then be based on the statistic

$$z = \frac{x - np}{\sqrt{np(1-p)}}, \tag{4.53}$$

which in our situation (with $p = 0.5$) simplifies to

$$z = \frac{2x - n}{\sqrt{n}}. \tag{4.54}$$

We may now simply compare the observed z statistic with the chosen $z_{\alpha/2}$ critical value as in the standard parametric case (or z_α for a one-sided test like the present case). Here, $z_\alpha = -2.326$ while observed $z = -2.676$ and we again must reject the null hypothesis.

4.3.2 Mann-Whitney test (U-test)

This test is a nonparametric alternative to the two-sample Student's t-test. It also goes by the names *Wilcoxon* test and the U-test. The Mann-Whitney test is performed by combining the two data sets we want to compare, sorting the combined set into ascending order, and assigning each point a *rank*: the smallest value is given rank = 1, while the largest observation is ranked $n_1 + n_2$. Should some of the observations be identical one assigns the mean rank to all tied values. E.g., if the 7th and 8th sorted values were identical, we would assign to each the mean rank of 7.5. The idea here is that if the samples consist of random drawings from the same population (i.e., when H_0 is true) then we would expect the ranks for both samples to be scattered more-or-less uniformly throughout the sequence. This would be true regardless of the distribution that characterizes the population.

After sorting the data we add up the ranks for each data set separately into *rank sums*, which we denote S_1 and S_2. The sum of $S_1 + S_2$ must obviously equal the sum of the first $(n_1 + n_2)$ integers, which is

$$\frac{1}{2}(n_1 + n_2)(n_1 + n_2 + 1). \tag{4.55}$$

Many early rank sum tests were based on S_1 or S_2 but now it is customary to use the statistic U, defined as $U = \min(U_1, U_2)$, i.e., the smallest of U_1 and U_2, with

$$U_1 = S_1 - \frac{1}{2}n_1(n_1 + 1) \tag{4.56}$$

and

$$U_2 = S_2 - \frac{1}{2}n_2(n_2 + 1). \tag{4.57}$$

This statistic can range from 0 to $n_1 \cdot n_2$ and its sampling distribution is symmetrical about $n_1 \cdot n_2/2$. The test, then, consists of these steps:

1. Compute the U-statistic using the smallest value of (4.56) and (4.57).
2. Evaluate critical U_{α,n_1,n_2}, given the sample sizes and the desired level of significance α.
3. Compare the calculated U statistic to the critical U_{α,n_1,n_2} and reject H_0 if U is *less than* the critical value.

Note that for the U-test, H_0 is rejected when our U-value is *less* than and not larger than the critical value, as is common for most other tests we have discussed.

Example 4–11. We want to compare the grain size of sand obtained from two different locations on the moon on the basis of measurements of grain diameters (in mm), as follows:

Location 1	0.37	0.70	0.75	0.30	0.45	0.16	0.62	0.73	0.33		$n_1 = 9$
Location 2	0.86	0.55	0.80	0.42	0.97	0.84	0.24	0.51	0.92	0.69	$n_2 = 10$

We do not know what type of distribution that grain sizes of sand on the moon might follow, so we choose the U-test to see if the mean grain size differ between the two samples. Computing the sample means gives $\bar{x}_1 = 0.49$ and

Data	Location	Rank	Data	Location	Rank
0.16	1	1	0.69	2	11
0.24	2	2	0.70	1	12
0.30	1	3	0.73	1	13
0.33	1	4	0.75	1	14
0.37	1	5	0.80	2	15
0.42	2	6	0.84	2	16
0.45	1	7	0.86	2	17
0.51	2	8	0.92	2	18
0.55	2	9	0.97	2	19
0.62	1	10			

Table 4.10: Sorted listing of the grain sizes of moon sand with auxiliary column containing the location of each specimen.

$\bar{x}_2 = 0.68$. If we wanted to use the t-test we would have to assume that the underlying distributions are normal, since the samples are small. The U-test requires no such assumptions. We start by arranging the data jointly into ascending order and keep track of which location each point originated from (Table 4.10).

We first evaluate the rank sum for location 1, giving $S_1 = 69$, from which it follows that

$$S_2 = \frac{19 \cdot 20}{2} - S_1 = 190 - 69 = 121. \tag{4.58}$$

We now form the null hypothesis $\mathrm{H}_0 : \mu_1 = \mu_2$, with $\mathrm{H}_1 : \mu_1 \neq \mu_2$, and state the level of significance $\alpha = 0.05$. Table A.10 has critical values for U and we find $U_{\alpha,9,10} = 20$. Thus, we will reject the null hypothesis if U is ≤ 20. From S_1 and S_2 we find

$$U_1 = 69 - \frac{9 \cdot 10}{2} = 24, \tag{4.59}$$

$$U_2 = 121 - \frac{10 \cdot 11}{2} = 66, \tag{4.60}$$

and hence $U = \min(24, 66) = 24$. This is larger than the critical value of 20, suggesting we *cannot reject* the null hypothesis. In other words, the observed difference in mean grain size at the two locations is not statistically significant at the 95% level of confidence.

For large samples $(n_1, n_2 > 30)$ the procedure again simplifies and it can be shown that the mean and standard deviation of the U sampling distribution approach

$$\mu_U = \frac{n_1 n_2}{2}, \quad \sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}, \tag{4.61}$$

provided there are *no tied ranks*. We could then evaluate standard z-scores as $z = (U - \mu_U)/\sigma_U$ and use the familiar critical values $\pm z_{\alpha/2}$ from Table A.2.

4.3.3 Comparing distributions: The Kolmogorov-Smirnov test

Another very useful nonparametric method is the Kolmogorov-Smirnov test (or K-S for short). It is a test for goodness of fit or *shape* and is often used instead of the χ^2-test. We may use it to test the null hypothesis that two distributions have the same probability density function (i.e., the same shape). A big advantage of the K-S test over the χ^2-test is that one does not have to bin the data, which is an arbitrary procedure anyway (how do you select bin size and why?). In the K-S test we convert the data distribution to a cumulative distribution $C(x)$. Clearly, $C(x)$ then gives the fraction of data points to the "left" of x. While different data sets will in general have different distributions, all cumulative distributions agree at the smallest x ($C(x) \equiv 0$) and at the largest x ($C(x) \equiv 1$). Thus, it is the behavior *between* these points that sets distributions apart (e.g., Figure 4.7). There is of course an infinite number of ways to measure the overall difference between two cumulative distributions: we could look at the absolute value of the area between the

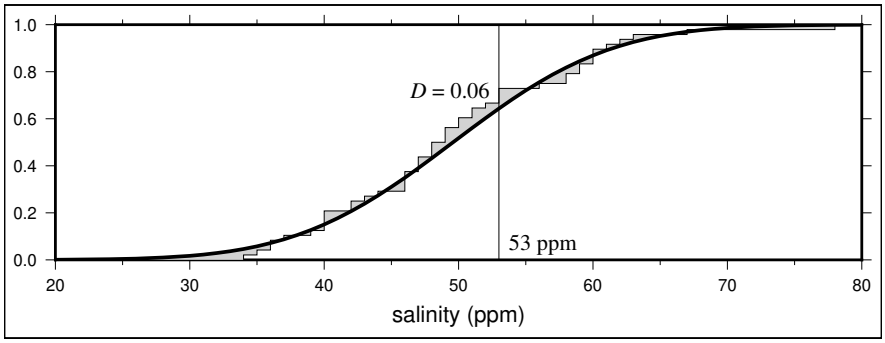


Figure 4.7: Solid line is a cumulative normal distribution for $\mu = 49.59$ ppm and $\sigma = 9.27$ ppm. The stair-case curve is the observed cumulative distribution which has its maximum difference, D, from the theoretical curve at the 53 ppm point.

curves, the mean square difference, etc. The K-S statistic chooses a simple measure: It determines the maximum absolute difference between the two cumulative curves. Thus, when comparing two cumulative distributions $C_1(x)$ and $C_2(x)$ our K-S statistic becomes

$$D = \max_{-\infty<x<\infty} |C_1(x) - C_2(x)|. \tag{4.62}$$

Note that C_2 may be another data-derived cumulative distribution (the test is then a *two-sample* test) or a theoretical cumulative probability function like the cumulative normal distribution (we call this case a *one-sample* test). The distribution of the K-S statistic itself can be calculated under the assumption that C_1 and C_2 are drawn from the same distribution (i.e., H_0), thus providing critical values for D. For the one-sample scenario there are two different cases to consider:

1. C_2 is a *known* cumulative distribution function, i.e., its parameters are prescribed by the null hypothesis.

2. C_2 is an *unknown* cumulative distribution function, i.e., its parameters must first be computed from C_1.

The former case is the problem studied by Kolmogorov and Smirnov. The latter case, however, clearly reduces the degrees of freedom and the standard K-S critical values are overestimated. A different set of critical values (called the *Lilliefors* critical values) has been developed for the normal distribution. Hence, when we need to compute the mean and standard deviation first we say we are performing a *Lilliefors test* rather than a Kolmogorov-Smirnov test.

We will use this test on the salinity measurements we looked at previously (Table 4.2). We sort the salinity measurements, convert them to a cumulative distribution (e.g., C_1), and plot the cumulative function on the same graph as that of a normal cumulative distribution with the same mean and standard deviation (e.g., C_2). Inspecting Figure 4.7 we find the maximum absolute difference to occur at the 53 ppt observation. The D estimate is $0.701 - 0.641 = 0.06$. Based on a significance level of $\alpha = 0.05$ and $n = 48$, the critical Lilliefors value for a two-sided test is found in Table A.13 to be ~ 0.128, which is much larger than observed. Hence we cannot reject the null hypothesis that the samples were collected from a normally distributed population. In this example, both the K-S and χ^2 tests reached the same conclusion.

At this point, we may wish to compare the usage of the χ^2-test and the K-S test. Both tests examine whether observed data match a theoretical (or another observed) distribution. However, only the χ^2-test imposes the criterion that we must *bin* the data first, a somewhat arbitrary process. Furthermore, it also restricts the number of bins we can use, especially for smaller data sets, since we need to maintain $E_j \geq 5$. For these reasons, we recommend you only use the χ^2-test where having or setting up bins come naturally (e.g., testing daily or monthly data), while using the K-S for true testing of distribution shapes.

4.3.4 Spearman's rank correlation

Finally, we will look at nonparametric correlation called Spearman's *rank correlation*, denoted by r_s. The rank correlation is carried out by ranking the x_i's and y_i's *separately*, then computing the standard correlation coefficient (i.e.,

4.29) using the ranks *in lieu* of the data values. Let u_i be the rank of the i'th pair's x-value and v_i be the rank of the i'th pair's y-value. Then, Spearman's rank correlation depends on the covariance and variances of the *ranks*:

$$r_s = \frac{s_{uv}}{s_u s_v} = \frac{n\sum_{i=1}^n u_i v_i - \left(\sum_{i=1}^n u_i\right)\left(\sum_{i=1}^n v_i\right)}{\sqrt{\left[n\sum_{i=1}^n u_i^2 - \left(\sum_{i=1}^n u_i\right)^2\right]\left[n\sum_{i=1}^n v_i^2 - \left(\sum_{i=1}^n v_i\right)^2\right]}}. \tag{4.63}$$

If there are runs of tied ranks then we assign those points their *average* rank. Fortunately, for situations where there are no ties (4.63) simplifies greatly to

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}, \tag{4.64}$$

where $d_i = u_i - v_i$ is the difference in rank for each (x_i, y_i) pair. In the case where the null hypothesis H_0: $\rho = 0$ is true, the sampling distribution of r_s is approximately normal and has zero mean ($\mu = 0$) and standard deviation $\sigma = 1/\sqrt{n-1}$. We could therefore base our statistics on

$$z = \frac{r_s - \mu}{\sigma} = \frac{r_s - 0}{1/\sqrt{n-1}} = r_s\sqrt{n-1} \tag{4.65}$$

and compare this observed z-value to critical $z_{\alpha/2}$ values. However, it turns out that a better approximation is the one given by (4.30), which we utilized when testing the standard correlation coefficient. Even so, for small data sets ($n < 20$) either approximation deviates from the true distribution and hence special tables are required (see Table A.15).

A comparison between the standard correlation and the Spearman's rank correlation reveals some interesting differences:

- Spearman's rank correlation is more tolerant of outliers since only their outlying *ranks* and not actual data values enter into the calculation.
- While the standard correlation measures the degree of *linear* correlation between x and y, Spearman's rank correlation measures the degree of *monotonicity* of the two rank series. Any data set whose ranks u and v vary monotonically will yield $r_s = \pm 1$, even if they do not form a linear trend.

In most other situations the two correlations will be similar.

Red (x)	**Rank** x	**Green** (y)	**Rank** y	**d**
4	3.5	5	4	0.5
2	1.5	2	2	0.5
4	3.5	6	5	1.5
2	1.5	1	1	-0.5
6	5	4	3	-2

Table 4.11: Evaluating the differences in ranks among $x - y$ pairs obtained by rolling red and green dice. Notice there are two groups of x-values with tied ranks but none among the y-values.

Ranking the dice data discussed earlier (Table 4.3) gives the values listed in Table 4.11. Using (4.64) we find $r_s = 0.65$ (surprisingly similar to the $r = 0.66$ we found using (4.29)), while the exact equation (4.63) yields $r_s = 0.6325$, which may be close enough for government work. For the simplified value the z-statistic from (4.65) becomes $z = 1.3$, which is well inside the 95% confidence limits ($z_{0.025} = \pm 1.96$) for a normal distribution. Likewise, using (4.30) we find $t = 1.41$ with critical $t_{0.025,3} = 3.18$. Hence, in either case we again arrive at the same conclusion that we cannot reject H_0. However, for such a small data set the approximations usually are quite poor; Table A.15 states the critical correlation is 1, meaning it would take a perfect nonparametric correlation to reject H_0.

In summary, there are numerous tests, both parametric and nonparametric, that can be applied to our data, and there are many others not covered in these notes. However, the ones presented here are the most common hypothesis tests that all scientists should be aware of. A simple guide to their use is given in Figure 4.8.

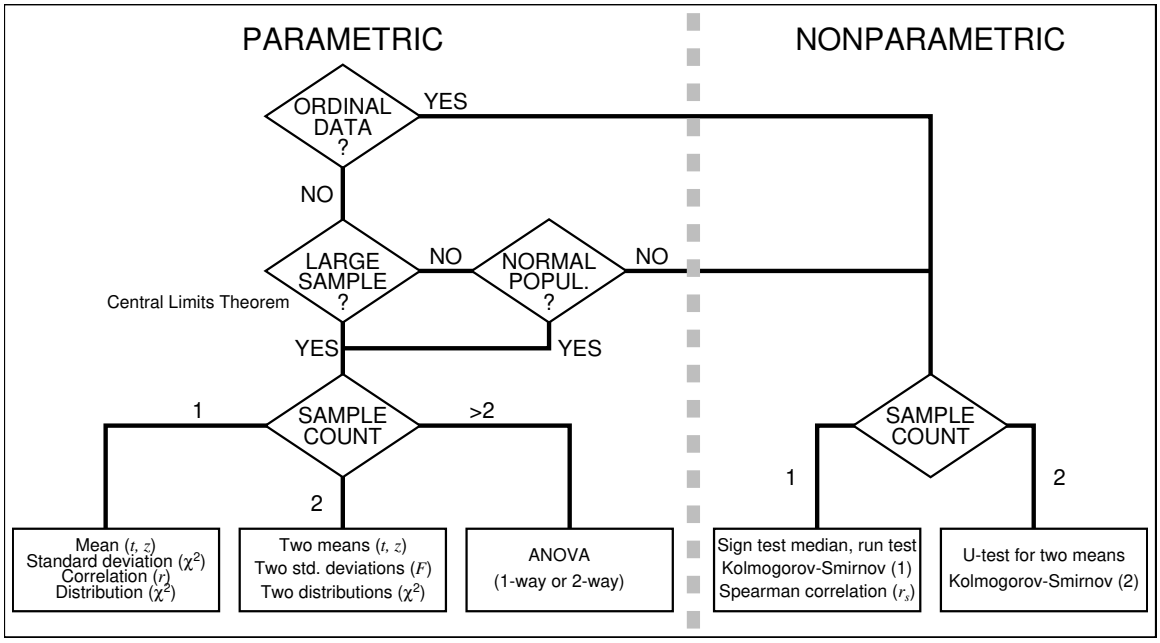


Figure 4.8: Simple decision chart for selecting standard parametric or nonparametric tests. The *run test* is a specific application of the sign test and will be discussed in Chapter 7.

4.4 Problems for Chapter 4

Problem 4.1. The following data were obtained in an experiment designed to check whether there is a systematic difference in the weights (in grams) obtained with two different scales. At the 0.01 level of significance, is there a systematic bias in the readings from the two scales? Hint: Given the hypothesis, do we have one or two samples? (Embrace the hint!).

Item	Scale 1	Scale 2
1	2.13	2.17
2	17.56	17.61
3	9.33	9.35
4	11.40	11.42
5	288.62	288.61
6	10.25	10.27
7	23.37	23.42
8	106.27	106.26
9	12.40	12.45
10	24.78	24.75

Problem 4.2. We have obtained two samples of magnesium concentrations (in ppm) from an igneous rock unit. The samples, obtained at two different locations, exhibit the concentrations listed below. At the 95% level of confidence, do the samples support the hypothesis that the average Mg concentration is the same at both locations? (Assume such data are normally distributed.).

Location 1	123	151	162	130	156	120	139	133	
Location 2	142	131	138	145	166	159	173	151	155

Problem 4.3. The following random samples are measurements of the heat-producing capacity (in millions of calories per ton) of specimens of coal from two mines:

Mine 1	8,400	8,230	8,380	7,860	7,930
Mine 2	7,510	7,690	7,720	8,070	7,660

At the 95% level of confidence, is the difference in mean value different? (You must first check if the standard deviations are similar).

Problem 4.4. If the seismic velocities v (in km/sec) in a limestone layer are normally distributed according to the probability density function

$$p(v) = \frac{1}{1.45\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{v - 5.15}{1.45} \right)^2 \right\},$$

then what is the probability that any single seismic refraction study will determine a velocity in the $5.5 < v < 6.0$ range?

Problem 4.5. After three months of strenuous field work a graduate student returns to the lab with new data to analyze. Water samples collected from two remote rivers are analyzed for mercury contamination. The results of the analysis (in ppm) are:

River A: 9.86, 12.02, 12.96, 10.40, 12.43, 9.61, 11.12, 10.64, 10.22.

River B: 11.36, 10.48, 11.06, 11.61, 13.28, 12.72, 13.91, 12.08, 12.38, 12.80.

a) At the 95% confidence level, do these samples support the alarming hypothesis that the average mercury contamination is higher in river B than in river A?

b) Given the statistical parameters you have obtained, plot on a single graph the two normal probability density distributions showing how the mercury samples are distributed, as well as the two *theoretical* probability density distributions of the sample means. Label your illustration and comment on what you see.

Problem 4.6. The water contents of soils (in volume %) were measured at two different sites A and B and are reproduced in *soilwater.txt*. At the 99% level of confidence, do the soils at the two sites have different water content?

Problem 4.7. A sensitive instrument measuring nickel content (in ppm) was tested on a set of reference specimens with known concentrations. The results (listed below in the "Before" columns) were unbiased and no calibration was necessary. After a thorough cleaning of the instrument the test sequence was repeated, with results listed in the "After" columns below.

Before	After	Before	After
211	198	172	166
180	173	155	154
171	172	185	181
214	209	168	164
183	179	203	201
194	192	180	175
160	161	245	240
181	182	146	142

At the 0.01 level of significance, is there a need to re-calibrate the instrument after the cleaning? Hint: Consider the differences in the measurements.

Problem 4.8. A soil scientist worrying about toxic waste wants to know if the standard deviation of a certain pollutant coming from several nearby factories is less than 10 ppm. Based on a sample of 15 values she finds $s = 11.3$ ppm. What can she conclude at the 95% level of confidence?

Problem 4.9. A designer of seismometers must know whether or not the standard deviation of the time it takes the instrument to start recording an event after being triggered by an earthquake is less than 0.010 s. Use the 0.05 level of significance to test the null hypothesis $H_0 : \sigma \leq 0.010$ s against the alternative hypothesis $H_1 : \sigma > 0.010$ on the basis of a random sample of size $n = 16$ for which the sample standard deviation was found to be $s = 0.012$ s.

Problem 4.10. An examination designed to measure basic knowledge in geology was given to random samples of freshmen at two major universities, and their scores were:

University A : $77, 72, 58, 92, 87, 93, 97, 91, 70, 98, 76, 90, 62, 69, 90, 78, 96, 84, 73, 80$

University B : $89, 74, 45, 56, 71, 74, 94, 88, 66, 62, 88, 63, 88, 37, 63, 75, 78, 34, 75, 68$

Apply the U-test at the 0.05 level of significance to test the null hypothesis that there is no significant difference in average knowledge of geology at the two universities. (MATLAB hint: Fill in vectors g and source (A = 1, B = 2), use [gsorted, gkey] = sort (g) to sort g, and use ranks1 = find (source(gkey) == 1) to get the ranks of entries from University A, etc.)

Problem 4.11. Given 10 data pairs you find $r = -0.85$. Is this correlation significant at the 99% level?

Problem 4.12. Given three data pairs (x_i, y_i), how high correlation coefficient would you need to determine for it to be significant at the 95% level of confidence?

Problem 4.13. A fringe seismologist goes on TV and claims there are more local earthquakes during certain months

of the year when planetary constellations are considered "favorable". A wee bit skeptical, you decide to examine the earthquake catalog for a whole year (see table *quakedays.txt*) to test his claim. The table gives the day number (1–365) when a local earthquake occurred. Assume this is not a leap year.

a) State the null hypothesis and the alternative hypothesis.

b) What are the expected number of events for each month?

c) Using the χ^2-test and a 95% level of confidence, do the data suggest that some months have more earthquakes than others?

Problem 4.14. An environmental scientist measures the sulfur dioxide emissions from an industrial plant over an 80 day period. The amounts (in tons per day) are given in the file *sulfur.txt*.

a) Bin the data using the categories less than 10, 10–15, 15–20, 20–25, 25 and above. Plot the histogram and indicate the counts.

b) The scientist wonders if the emissions are well described by the expected normal distribution. What are the mean and standard deviation for the raw data?

c) The scientist decides to use a χ^2 test. What are the expected counts in each bin?

d) Test whether or not the binned data are indistinguishable from a normal distribution at the 95% level of confidence. Should she reject H_0?

Problem 4.15. From past experience, we know there is a weak correlation between the the proportion of manganese nodules on the seafloor and the age of the underlying crust. How many data pairs, n, do we need to sample before an inferred positive correlation of $r = 0.32$ is significantly larger than zero at the 95% level of confidence? (MATLAB hint: Use `icdf` to get critical t-values – try `help icdf` to see how to use it.)

Problem 4.16. A senator from a "blue state" believes oil imported to the US from Norway has been infused with a nefarious toxic chemical. He theorizes that when drivers in the US use fuel from the imported oil they become debilitated from the fumes and the chance of accidents at railroad crossings increases significantly. He obtains two data sets from 1999–2009 showing the annual import of oil (in millions of barrels) and the number of drivers killed in collisions with railway trains, reproduced in the file *conspiracy.txt*. Plot the two time series on the same graph (with different vertical scales), then calculate the correlation between the annual data pairs. Is the correlation significant at the 95% level? Should the Senate punish the Norwegians by authorizing US forces to invade their little kingdom and confiscate their oil rigs?

Problem 4.17. A republican congressman is upset with the amount of funding the US National Science Foundation allocates to support research in the social sciences. Looking for arguments to reduce this wasteful spending, he comes across data that seem to suggest that awarding too many doctorates in sociology may cause an increase in deaths due to *anticoagulants*. He theorizes that the lack of rigor in the social sciences causes blood to thin, leading to these unnecessary deaths. As his intern, you are tasked with examining the data (see file *PhD_deaths.txt*) and determine if the correlation between the number of Ph.D.'s and deaths might be significant at the 99% level. What will your report to the Congressman conclude?

Problem 4.18. We have made determinations of chromium content in four shale units. For each unit we obtained six measurements; the data are summarized below (values are given in ppm). With a one-way ANOVA at the 0.05 level of significance, can the differences among the four sample means be attributed solely to measurement uncertainties?

Unit A	Unit B	Unit C	Unit D
181.3	160.6	163.2	132.3
176.8	179.9	154.1	140.8
182.7	170.9	158.6	124.5
186.0	151.3	137.8	123.2
188.4	163.4	158.7	121.1
160.2	185.8	191.8	163.2

Problem 4.19. Four different temperature gauges recorded the temperature in a storage room. The data are summarized below. At the 0.05 level of significance, can the differences among the four sample means be attributed to chance?

Gauge A	Gauge B	Gauge C	Gauge D
23.1	21.7	21.9	19.8
22.8	23.0	21.3	20.4
23.2	22.4	21.6	19.3
23.4	21.1	20.2	18.5
23.6	21.9	21.6	19.1
21.7	23.4	23.8	21.9

Problem 4.20. Water samples were obtained from four different locations along a river to determine whether the quantity of dissolved oxygen, a proxy for water pollution, varies between the locations. Locations A and B were selected above an industrial plant: one near the shore and another in midstream. Location C was adjacent to the industrial water discharge from the plant, and location D was slightly downriver in midstream. Water specimens were randomly selected at each location, but one specimen, from location D, was lost in the laboratory. The data are shown in the table below, with lower values corresponding to higher levels of pollution.

a) Do the data provide sufficient evidence to indicate a significant difference in mean dissolved oxygen content for the four locations?

b) Compare the mean dissolved oxygen content in midstream above the plant with the mean content adjacent to the plant (i.e., locations B versus C). Use a 95% confidence interval. What do you conclude?

Location	Mean Oxygen Content				
A	5.9	6.1	6.3	6.1	6.0
B	6.3	6.6	6.4	6.4	6.5
C	4.8	4.3	5.0	4.7	5.1
D	6.0	6.2	6.1	5.8	

Problem 4.21. A laboratory technician measures the breaking (tensile) strength of each of five samples of granite (G_1 – G_5) using four different measuring instruments (I_1 – I_4), and obtains the following results (in MPa):

	I_1	I_2	I_3	I_4
G_1	18.6	18.1	17.6	19.6
G_2	22.7	23.9	24.7	22.5
G_3	23.2	20.8	19.2	22.1
G_4	22.5	18.9	21.2	23.4
G_5	17.3	18.9	19.8	18.8

a) Using a two-way ANOVA test, do the data provide sufficient evidence to indicate a difference in strength for the five samples? (Use $\alpha = 0.1$)

b) Is there any statistical evidence of instrument bias? (Use $\alpha = 0.1$)

Problem 4.22. A study was conducted to compare automobile gasoline mileage for three brands of gasoline, A, B, C. Four cars, all of the same make and model, were employed in the experiment and each gasoline brand was tested in each car. The data, in miles per gallon, are as follows:

	CAR 1	**CAR 2**	**CAR 3**	**CAR 4**
GAS A	15.7	17.0	17.3	16.1
GAS B	17.2	18.1	17.9	17.7
GAS C	16.1	17.5	16.8	17.8

a) Using a two-way ANOVA, do the data provide sufficient evidence to indicate a difference in the mean mileage per gallon for the three gasoline brands? (Use $\alpha = 0.1$)

b) Is there any statistical evidence of differences mean mileage for the four automobiles? (Use $\alpha = 0.1$)

Problem 4.23. Densities ρ_i were obtained from a sample of 40 specimens from the same lithological unit (see *rho.txt*). Use the sign test at the 0.01 level of significance to test the null hypothesis $\tilde{\rho} \leq 2.42$ against the alternative hypothesis $\tilde{\rho} > 2.42$.

Problem 4.24. The science team on an oceanographic research vessel has measured the heat flux through the ocean floor at two sites with similar crustal ages. Site I was formed by seafloor spreading with a spreading rate of 15 cm/yr while site II was formed at half the spreading rate. The heat flux values (in mWm^{-2}) recorded were:

Site I	57.15	63.00	67.64	62.09	73.84	65.54	61.88	67.95
Site II	56.82	58.88	65.93	56.34	59.92	50.52	65.37	53.62

a) Use the U-test at the 0.05 level of significance to test the hypothesis that the heat flux from site II is different from that of site I. (MATLAB hint: Fill in vectors q and source, use [qsorted, key] = sort (q) to sort q, and use ranks1 = find (source(key) == 1) to get the ranks of entries from site 1, etc.)

b) Upon reviewing the literature for the region, it is discovered that a previous cruise recorded one heat flux measurement from Site I, yielding a value of 74.19 mWm^{-2}. Does this change the results obtained above?

Problem 4.25. The Earth's surface temperature (important in many agricultural as well as hydrological problems) can be tediously measured on the ground or conveniently recorded by remote infrared sensors mounted on airplanes or satellites. However, the remotely sensed data appear to have a bias. At the 95% level of confidence, is there such a bias in the data below?

Location	**Ground (°C)**	**Remote (°C)**
1	36.9	37.3
2	35.4	38.1
3	26.3	27.9
4	21.0	22.7
5	14.7	16.2

Problem 4.26. The textural properties of sandstones are believed to reflect environmental conditions when they first formed. Textural maturity is defined as the degree to which a sand is both well *sorted* and well *rounded*. These two characteristics are expected to be correlated. However, roundness and degree of sorting are not quantified on a ratio scale but assigned ordinal values such as "moderately sorted" or "well sorted". Hence, we can rank the assessments

but not calculate ordinary statistics from them. Based on the table below, is there a significant (95% level) correlation between roundness and degree of sorting? (P = poor, M = moderate, W = well sorted, and A = angular, SA = subangular, SR = subrounded.)

Sorting	Rank	Roundness	Rank
P	4	SR	11
W	10	SA	9
P	2	A	1
M	8	SA	4
M	6	SA	6
W	9	SR	12
P	3	SA	8
M	7	SA	3
W	11	SA	7
W	12	SR	10
M	5	A	2
P	1	SA	5

Problem 4.27. A homeowner decides to test whether or not the level of salt in his drinking water at home is significantly higher than that of the drinking water at his office. Measuring the salinity (in ppm) gives:

Home	76.19	84.00	79.89	82.78	98.45	87.38	82.50	90.60
Office	75.76	78.51	87.91	75.12	90.19	67.36	87.16	

a) Use the U-test at the 0.01 level of significance to test his hypothesis.

b) He decides to get an 8th measurement from the office and finds salinity to be 72.31. Does this change the results obtained above?

Problem 4.28. Permeabilities have been estimated in the lab for samples taken from two different sandstone outcrops. We obtained these values (in Darcy; 1 Darcy = $9.8697 \times 10^{-13}\text{m}^2$):

A	18.76	13.24	3.83	10.12	8.40	8.60	18.18	15.04	9.62	13.22
B	15.90	5.80	1.31	17.50	9.22	6.20	1.92	13.46	2.61	8.01

Using the Kolmogorov-Smirnov test at the 95% level of confidence, do these samples appear to come from the same lithological unit (i.e., population)? Use the MATLAB script `kolsmir.m` and the two-sample critical values provided in Table A.14.

Problem 4.29. We have obtained two samples of residual magnetization from a basaltic sill at two different sites. The values are:

Site 1	68.9	41.1	-6.1	25.6	17.0	18.0	65.9	50.0	23.1	41.1
Site 2	54.5	4.0	-18.5	62.5	21.0	6.0	-15.5	42.2	-13.0	15.0

Using the Kolmogorov-Smirnov test at the 95% level of confidence, do these samples come from the same population?

Problem 4.30. As discussed in Chapter 1, the Earth's magnetic field reverses direction. Table *GK2007.txt* contains data from the Gee and Kent (2007) geomagnetic time scale. It lists all normal and reversely magnetized chrons and gives the duration of each interval in Myr. Note: Examine the last letter in the chron: "n" means normalized and "r"

stands for reversed polarity. Using the Kolmogorov-Smirnov test at the 95% level of confidence, are the distributions of intervals for the normal and reverse polarities different?

Problem 4.31. We will revisit Problem 4.13, but this time we will utilize the Kolmogorov-Smirnov test and completely avoid any discussion about binning artifacts.

a) What is the cumulative distribution function you will compare your observed data to?

b) Can you reject the null hypothesis at the 99% level of confidence?

Problem 4.32. From a Carboniferous shale we obtain the following concentrations of chromium and nickel (in ppm):

Cr (ppm)	Ni (ppm)
122	604
340	311
522	173
61	503
133	495
235	444
498	272
371	362
239	384

a) What is the correlation coefficient, r. Is the correlation significant at $\alpha = 0.01$?

b) What is Spearman's rank correlation coefficient, r_s? Is it significant at $\alpha = 0.01$?

c) Suppose one additional measurement was added to the table, with Cr = 698, Ni = 597. What are the new correlations r and r_s? Are these correlations significant at $\alpha = 0.01$?

d) We suspect the last measurement to be an outlier. Given the recipe of detecting outliers discussed previously based on the statistical properties of the first 9 values, are the Cr and Ni values for the additional measurement outliers in their respective groups?

Problem 4.33. The abundance (in %) of four elements in seven samples of basalt from the Pacific has been recorded as

Sample #	Si	Al	Fe	Mg
1	22.5	9.6	6.6	3.4
2	22.1	8.4	7.8	3.6
3	25.9	8.7	4.8	4.0
4	23.5	8.1	5.0	5.2
5	21.7	10.0	8.2	4.9
6	21.9	8.2	9.3	4.9
7	23.7	7.2	9.5	3.3

Compute the correlations between all pairs of elements. Are any of the correlations significant at the 0.05 level?

Chapter 5

LINEAR (MATRIX) ALGEBRA

"Never send a human to do a machine's job."

Agent Smith in *The Matrix*

A large subset of data analysis techniques is simply the practical application of linear algebra. Undergraduate students in the natural sciences often lack any formal introduction to this material, or they suffered through an overly theoretical presentation in a course offered by a mathematics department. In this book we will not dwell too much on the finer theoretical points of linear algebra but instead present a simple overview of the aspects that are particularly pertinent in data analysis. There are of course an infinite number of books on matrix and linear algebra that the eager reader could consult beyond this brief introduction.

5.1 Matrix Algebra Terminology

A *matrix* is simply a rectangular array of elements arranged in a series of n rows and k columns. The *order* of a matrix is the specification of the number of rows by the number of columns. *Elements* of a matrix are denoted a_{ij}, where index i specifies the *row* position and index j specifies the *column* position; thus a_{ij} identifies the element at position i, j.

An element can be a number (real or complex), an algebraic expression, or (with some restrictions) another matrix or matrix expression. As an example of a real matrix, consider

$$\mathbf{A} = \begin{bmatrix} 12 & 4 & 10 \\ 8 & 1 & 11 \\ 15 & 3 & 11 \\ 14 & 1 & 11 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}. \tag{5.1}$$

This matrix, $\mathbf{A}$, is of order 4×3, with elements a_{11} = 12, a_{21} = 8, and so on. The notation used for matrices is not always consistent, but it is usually one of the following schemes:

Matrix: Designated by a bold, uppercase letter (the most common scheme), brackets, or hat (^), sometimes with one or (more commonly) two underscores. The order is also sometimes explicitly given. E.g., $\mathbf{A}_{(4,3)}$ means $\mathbf{A}$ is of order 4×3.

Order: Always given as rows x columns but uses letters n, k, m, p differently, e.g., n (rows) x k (columns) or k (rows) x n (columns).

Element: Most commonly a_{ij}, with $i =$ row; $j =$ column (sometimes other dummy indices like l, p, q are used).

The advantages of matrix algebra mainly lies in the fact that it provides a concise and simple method for manipulating large sets of numbers or computations, making it ideal for computers. Furthermore,

1. The compact form of matrices allows a convenient notation for describing large tables of data.

2. Matrix operations allow complex relationships to be seen, which otherwise would be obscured by the shear size of the data (i.e., it aids in clarification).

3. Most matrix manipulations involve just a few standard operations for which standard subroutines are readily available.

MATLAB, which stands for "Matrix Laboratory", is ideally suited to perform such manipulations, as is its Open Source clone, Octave. Python with numPy, R or Julia are also good choices.

As a convention with data matrices (i.e., when the elements are data values), the columns usually represent the different *variables* (e.g., one column contains temperatures, another salinity, etc.) while rows contain the *observations* (e.g., the values of the variables at different depths, times, or positions). Since there are usually more observations than variables, such data matrices are typically rectangular, having more rows (n) than columns (k), i.e., the matrix has an order $n \times k$ where $n > k$.

5.2 Matrix Definitions

Matrices whose smallest dimension equals one are called *vectors* and are typically designated by a bold, lowercase letter, but sometimes they may be typeset normally with either an arrow above them or a single underscore beneath. By having only one dimension, one of the two indices (row or column) is dropped. A *column vector* is a matrix containing only a single column of elements, such as

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}. \tag{5.2}$$

A *row vector* is a matrix containing only a single row of elements, e.g.,

$$\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}. \tag{5.3}$$

The size of a vector is simply the number of elements it contains (n, in both examples above). The *null matrix*, written as $\mathbf{0}$ or $\mathbf{0}_{(k,n)}$ has all its elements set equal to 0 — it plays the role of "zero" in matrix algebra. A *square matrix* has the same numbers of rows as columns, so its order is $n \times n$. A *diagonal matrix* is a square matrix with zeros in all positions except along the principal (or leading) diagonal:

$$\mathbf{D} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix} \tag{5.4}$$

or

$$d_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ \text{nonzero} & \text{for } i = j \end{cases} \tag{5.5}$$

This type of matrix is important for scaling the rows or columns of other matrices. The *identity matrix* is a diagonal matrix with all of its nonzero elements equal to one. Written as $\mathbf{I}$ or $\mathbf{I}_n$, it plays the role of "one" in matrix algebra. A *lower triangular matrix* is a square matrix with all elements equal to zero *above* the principal diagonal, e.g.,

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 7 & 0 \\ 8 & 2 & 6 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 3 & 7 & \\ 8 & 2 & 6 \end{bmatrix} \tag{5.6}$$

or

$$\ell_{ij} = \begin{cases} 0 & \text{for } i < j \\ \text{nonzero} & \text{for } i \geq j \end{cases} \tag{5.7}$$

An *upper triangular matrix* is thus a square matrix with all elements equal to zero *below* the principal diagonal

$$u_{ij} = \begin{cases} 0 & \text{for } i > j \\ \text{nonzero} & \text{for } i \leq j \end{cases} \tag{5.8}$$

If one multiplies two triangular matrices of the same form, the result is a third matrix of the same form.

We also have the *fully populated matrix* which is a matrix with all of its elements nonzero, the *sparse matrix* which is a matrix with only a small proportion of its elements nonzero, and the *scalar* which simply is a number (i.e., a matrix of order 1x1, representing a single element).

A *matrix transpose* (or the transpose of a matrix) is obtained by *interchanging* the rows and columns of the matrix. Thus, row i becomes column i and column j becomes row j. As a consequence, the order of the matrix is reversed:

$$\mathbf{A} = \begin{bmatrix} 1 & 14 \\ 6 & 7 \\ 8 & 2 \end{bmatrix} \Rightarrow \mathbf{A}^T = \begin{bmatrix} 1 & 6 & 8 \\ 14 & 7 & 2 \end{bmatrix} \tag{5.9}$$

As shown, taking the transpose is indicated by the superscript ^{T}. Repeated transposing yields the original matrix, i.e.,

$$(\mathbf{A}^T)^T = \mathbf{A}. \tag{5.10}$$

A diagonal matrix is its own transpose: $\mathbf{D}^T = \mathbf{D}$. In general, we find the transpose rule

$$a_{ij} \Leftrightarrow a_{ji}. \tag{5.11}$$

A *symmetric matrix* is a square matrix that is symmetric about its principal diagonal, so $a_{ij} = a_{ji}$. Therefore, a symmetric matrix is equal to its own transpose:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 6 & 3 \\ 5 & 3 & 4 \end{bmatrix} = \mathbf{A}^T. \tag{5.12}$$

A *skew symmetric matrix* is a matrix in which

$$a_{ij} = -a_{ji}. \tag{5.13}$$

Therefore, $\mathbf{A}^T = -\mathbf{A}$. Thus, a_{ii}, the principal diagonal elements, must all be zero. The following matrix is skew symmetric:

$$\mathbf{A} = \begin{bmatrix} 0 & 4 & -5 \\ -4 & 0 & 3 \\ 5 & -3 & 0 \end{bmatrix}. \tag{5.14}$$

Any square matrix can be decomposed into the *sum* of a symmetric and a skew-symmetric matrix:

$$\mathbf{A} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) + \frac{1}{2}(\mathbf{A} - \mathbf{A}^T). \tag{5.15}$$

The *trace* of a square matrix is simply the sum of the elements along the principal diagonal. It is symbolized as tr ($\mathbf{A}$). This property is useful in calculating various quantities from matrices. *Submatrices* are smaller matrix *partitions* of the larger *supermatrix*, i.e.,

$$\left[\begin{array}{c} \text{Supermatrix} \\ \hline \mathbf{F} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right]. \tag{5.16}$$

Such partitioning will frequently be useful.

5.3 Matrix Addition

Matrix addition and subtraction require matrices of the *same order* since each operation simply involves the addition or subtraction of corresponding elements. So, if $\mathbf{C} = \mathbf{A} + \mathbf{B}$ then

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}, \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}, \mathbf{C} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \\ a_{31}+b_{31} & a_{32}+b_{32} \end{bmatrix}, \tag{5.17}$$

and (with apologies to ABBA fans)

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}, \tag{5.18}$$

$$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C}), \tag{5.19}$$

where all matrices must be of the same order. *Scalar multiplication* of a matrix is achieved by multiplying all elements of a matrix by a constant (the scalar):

$$\beta\mathbf{A}=\beta\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{32} \\ a_{31} & a_{32} \end{bmatrix}=\begin{bmatrix} \beta a_{11} & \beta a_{12} \\ \beta a_{21} & \beta a_{22} \\ \beta a_{31} & \beta a_{32} \end{bmatrix}, \tag{5.20}$$

where β is a scalar. Thus, every element is multiplied by the scalar.

5.4 Dot Product

The *scalar product* (or *dot product* or *inner product*) is the product of two vectors of the same size, e.g.,

$$\mathbf{a}\cdot\mathbf{b}=\beta, \tag{5.21}$$

where $\mathbf{a}$ is a row vector (or the transpose of a column vector) of length n, $\mathbf{b}$ is a column vector (or the transpose of a row vector), also of length n, and β is the scalar product of $\mathbf{a}\cdot\mathbf{b}$. Given the two 3-D vectors

$$\mathbf{a}=[a_1\ a_2\ a_3],\quad \mathbf{b}=\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{5.22}$$

we sum the products of corresponding elements in the two vectors, obtaining

$$\beta=a_1b_1+a_2b_2+a_3b_3. \tag{5.23}$$

We may visualize this multiplication, as illustrated in Figure 5.1 for two 4-D vectors.

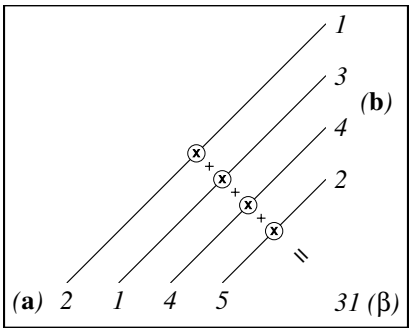


Figure 5.1: The dot product of the two 4-D vectors $\mathbf{a}=[2\ 1\ 4\ 5]$ and $\mathbf{b}=[1\ 3\ 4\ 2]$ is obtained by multiplying the component pairs and calculating the sum of these products.

Geometrically, this product can be thought of as multiplying the length of one vector by the component of the other vector that is parallel to the first, as shown in Figure 5.2:

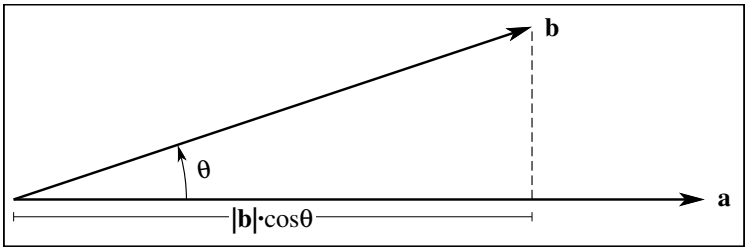


Figure 5.2: Geometrical meaning of the dot product of two vectors. Regardless of dimension, the dot product is proportional to the cosine of the angle between the two vectors.

As an example, think of $\mathbf{b}$ as a force and $|\mathbf{a}|$ as the magnitude of displacement, with their product equal to the work in the direction of $\mathbf{a}$. Thus:

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos(\theta), \tag{5.24}$$

where the *magnitude* of a vector $\mathbf{x}$ is given by

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}. \tag{5.25}$$

The maximum principle says that the unit vector $(\hat{\mathbf{n}})$ making $\mathbf{a}\cdot\hat{\mathbf{n}}$ a maximum is that unit vector pointing in the same direction as $\mathbf{a}$: If $\hat{\mathbf{n}} \parallel \mathbf{a}$ then $\cos(\theta) = \cos(0^\circ) = 1$ and $\mathbf{a}\cdot\mathbf{n} = |\mathbf{a}||\mathbf{n}|\cos(\theta) = |\mathbf{a}||\hat{\mathbf{n}}| = |\mathbf{a}|$. This is equally true where $\mathbf{d}$ is any vector of a given magnitude — that vector $\hat{\mathbf{n}}$ which parallels $\mathbf{d}$ will give the largest scalar product. Parallel vectors thus have $\cos(\theta) = 1$, then $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|$ and $\mathbf{a} = \beta\mathbf{b}$ (i.e., two vectors are parallel if one is simply a scalar multiple of the other — this property comes from equating direction cosines), where

$$\beta = |\mathbf{a}|/|\mathbf{b}|. \tag{5.26}$$

In contrast, *perpendicular* vectors have $\cos\theta = \cos 90^\circ = 0$, so that $\mathbf{a}\cdot\mathbf{b} = 0$, where $\mathbf{a}\perp\mathbf{b}$.

Squaring vectors is simply the dot product of a vector with its own transpose, i.e.,

$$\mathbf{a}^2 = \mathbf{a}\cdot\mathbf{a}^T \text{ for row vectors} \tag{5.27}$$

and

$$\mathbf{a}^2 = \mathbf{a}^T\cdot\mathbf{a} \text{ for column vectors.} \tag{5.28}$$

5.5 Matrix Multiplication

Matrix multiplication requires "conformable" matrices. Matrices are conformable when there are as many columns in the first matrix as there are rows in the second matrix. Consider

$$\mathbf{C}_{(n,k)} = \mathbf{A}_{(n,p)}\cdot\mathbf{B}_{(p,k)}. \tag{5.29}$$

The matrix product $\mathbf{C}$ is of order $n\times k$ and has elements c_{ij}, given by

$$c_{ij} = \sum_{q=1}^{p} a_{iq}b_{qj}. \tag{5.30}$$

This is an extension of the scalar product — in this case, each element of $\mathbf{C}$ is the scalar product of a row vector in $\mathbf{A}$ and a column vector in $\mathbf{B}$. For instance, if

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}, \tag{5.31}$$

then

$$c_{12} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32}. \tag{5.32}$$

We illustrate the situation in Figure 5.3. The order of multiplication is critical. Usually

$$\mathbf{A}\cdot\mathbf{B} \neq \mathbf{B}\cdot\mathbf{A}, \tag{5.33}$$

and unless $\mathbf{A}$ and $\mathbf{B}$ are square matrices or the order of $\mathbf{A}^T$ is the same as the order of $\mathbf{B}$ (or vice versa), one of the two products cannot even be formed. The multiplication order is specified by stating

$\mathbf{A}$ is *pre*-multiplied by $\mathbf{B}$ (yielding $\mathbf{B}\cdot\mathbf{A}$)

$\mathbf{A}$ is *post*-multiplied by $\mathbf{B}$ (yielding $\mathbf{A}\cdot\mathbf{B}$)

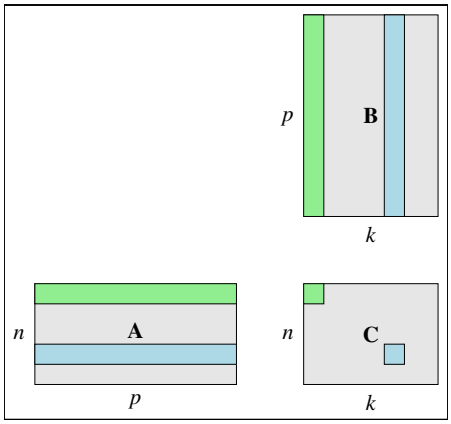


Figure 5.3: The matrix product of two matrices **A** and **B**. The light blue row in **A** is "dotted" with the light blue column in **B**, resulting in the single, light blue element in **C**. Similarly, the dot product of the light green vectors result in the single, light green element. This process is repeated by letting all rows in **A** be "dotted" with all the columns in **B**.

The order in which the pairs are multiplied is not important *mathematically*, i.e.,

$$\mathbf{D} = (\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C} = \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}), \tag{5.34}$$

but we will see later that the order matter *computationally*. The *transpose of a matrix product* is simply the multiplication of the transpose of each individual matrix in reverse order, i.e.,

$$\mathbf{D} = \mathbf{A}\cdot\mathbf{B}\cdot\mathbf{C}, \tag{5.35}$$

$$\mathbf{D}^T = \mathbf{C}^T\cdot\mathbf{B}^T\cdot\mathbf{A}^T. \tag{5.36}$$

Multiplication by **I** leaves the matrix unchanged, i.e.,

$$\mathbf{A}\cdot\mathbf{I} = \mathbf{I}\cdot\mathbf{A} = \mathbf{A}. \tag{5.37}$$

For example,

$$\begin{bmatrix} 3 & 6 & 9 \\ 2 & 8 & 7 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 2 & 8 & 7 \end{bmatrix}. \tag{5.38}$$

Premultiplication by a diagonal matrix is written $\mathbf{C} = \mathbf{D}\cdot\mathbf{A}$, where **D** is a diagonal matrix. Here, **C** is the **A** matrix with each *row* scaled by the corresponding diagonal element of **D**:

$$\mathbf{D} = \begin{bmatrix} d_{11} & & \\ & d_{22} & \\ & & d_{33} \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{5.39}$$

$$\mathbf{C} = \mathbf{D}\cdot\mathbf{A} = \begin{bmatrix} a_{11}d_{11} & a_{12}d_{11} & a_{13}d_{11} \\ a_{21}d_{22} & a_{22}d_{22} & a_{23}d_{22} \\ a_{31}d_{33} & a_{32}d_{33} & a_{33}d_{33} \end{bmatrix} \begin{array}{l} \leftarrow \text{ each element } \times d_{11} \\ \leftarrow \text{ each element } \times d_{22} \\ \leftarrow \text{ each element } \times d_{33} \end{array} \tag{5.40}$$

Postmultiplication by a diagonal matrix produces a matrix in which each *column* has been scaled by the corresponding diagonal element **D**. Hence,

$$\mathbf{C} = \mathbf{A}\cdot\mathbf{D} = \begin{bmatrix} a_{11}d_{11} & a_{12}d_{22} & a_{13}d_{33} \\ a_{21}d_{11} & a_{22}d_{22} & a_{23}d_{33} \\ a_{31}d_{11} & a_{32}d_{22} & a_{33}d_{33} \end{bmatrix}, \tag{5.41}$$

where each column in **A** has been scaled by the corresponding diagonal matrix elements, d_{ii}.

5.5.1 Computational considerations

The matrix product

$$\mathbf{C}_{(n,k)} = \mathbf{A}_{(n,p)} \cdot \mathbf{B}_{(p,k)} \tag{5.42}$$

involves $n \times k \times p$ multiplications and $n \times k \times (p-1)$ additions. hence,

$$\mathbf{E}_{(n,k)} = [\mathbf{A}_{(n,p)} \cdot \mathbf{B}_{(p,q)}] \cdot \mathbf{C}_{(q,k)} \tag{5.43}$$

gives $n \times p \times q$ multiplications, so

$$\mathbf{E} = [\mathbf{D}_{(n,q)}] \cdot \mathbf{C}_{(q,k)} \tag{5.44}$$

gives $n \times q \times k$ multiplications, and

$$\mathbf{E}_{(n,k)} = \mathbf{A}_{(n,p)} \cdot [\mathbf{B}_{(p,q)} \cdot \mathbf{C}_{(q,k)}] \tag{5.45}$$

gives $p \times q \times k$ multiplications, while

$$\mathbf{E}_{(n,k)} = \mathbf{A}_{(n,p)} \cdot [\mathbf{D}_{(p,k)}] \tag{5.46}$$

gives $n \times p \times k$ multiplications. Therefore, the total number of operations depend on the order of multiplications:

1. $(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} \Rightarrow nq(p+k)$ total multiplications
2. $\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) \Rightarrow pk(n+q)$ total multiplications

If both $\mathbf{A}$ and $\mathbf{B}$ are 100×100 matrices and $\mathbf{C}$ is 100×1, then $n = p = q = 100$, and $k = 1$. Multiplying using form (1) involves $\sim 1 \times 10^6$ multiplications, whereas form (2) involves 2×10^4; so computing $\mathbf{B} \cdot \mathbf{C}$ first, then premultiplying by $\mathbf{A}$ saves almost a million multiplications and almost an equal number of additions. Therefore, the order of operations is extremely important computationally for both speed and accuracy, as more operations lead to a greater accumulation of *round-off errors*.

5.6 Matrix Determinant

The *determinant of a matrix* is a single number representing a property of a square matrix (and is dependent upon what the matrix represents). The main use here is for finding the inverse of a matrix or for solving simultaneous linear equations. Symbolically, the determinant is usually written as det $(\mathbf{A})$, $|\mathbf{A}|$ or $||\mathbf{A}||$ (to differentiate from magnitude). The calculation of a 2×2 determinant is carried out using the definition

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \tag{5.47}$$

which is the difference of the cross products. The calculation of an $n \times n$ determinant is given by

$$|\mathbf{A}| = a_{11}m_{11} - a_{12}m_{12} + a_{13}m_{13} - \cdots - (-1)^n a_{1n}m_{1n}, \tag{5.48}$$

where m_{11} is the determinant of $\mathbf{A}$ with the first row and column missing; m_{12} is the determinant with the first row and second column missing, etc. For larger matrices the procedure is used recursively. The determinant of a 1×1 matrix is just the particular element. An example of a 3×3 determinant follows:

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{5.49}$$

$$m_{11} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32} \tag{5.50}$$

$$m_{12} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{21}a_{33} - a_{23}a_{31} \tag{5.51}$$

$$m_{13} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{21}a_{32} - a_{22}a_{31} \tag{5.52}$$

So

$$\begin{aligned} |\mathbf{A}| &= a_{11}m_{11} - a_{12}m_{12} + a_{13}m_{13} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}). \end{aligned} \tag{5.53}$$

For a 4×4 determinant, each m_{1i} would be an entire expansion given above for the 3×3 determinant — one quickly needs a computer.

A *singular matrix* is a square matrix whose determinant is zero. A determinant is zero if:

1. Any row or column is zero.
2. Any row or column is equal to a linear combination of other rows or columns.

As an example of a singular matrix, consider

$$|\mathbf{A}| = \begin{vmatrix} 1 & 6 & 4 \\ 2 & 1 & 0 \\ 5 & -3 & -4 \end{vmatrix}, \tag{5.54}$$

where row $1 = 3\cdot$(row 2) $-$ row 3. Then, the determinant becomes

$$\begin{aligned} |\mathbf{A}| &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= 1[1(-4) - 0(-3)] - 6[2(-4) - 0(5)] + 4[2(-3) - 1(5)] = -4 + 48 - 44 = 0. \end{aligned} \tag{5.55}$$

The *degree of clustering* symmetrically about the principal diagonal is another (of many) properties of a determinant. The more the clustering, the higher the value of the determinant. The *rank* of a matrix is the number of linearly independent vectors that it contains (either row or column vectors). Consider

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 0 & 2 \\ 1 & 0 & 1 & -1 \\ -3 & -4 & -2 & 0 \end{bmatrix}. \tag{5.56}$$

Since row $3 = -$(row 1)$-2\cdot$(row 2), or col $3 =$ col $1 - 1/4\cdot$(col 2) and col $4 = -$(col 1)$+3/4\cdot$(col 2), the matrix $\mathbf{A}$ has rank 2 (i.e., it has only two linearly independent vectors, independent of whether viewed by rows or columns).

The *rank* of a *matrix product* must be less than or equal to the smallest rank of the matrices being multiplied:

$$\mathbf{A}_{(\text{rank } 2)} \cdot \mathbf{B}_{(\text{rank } 1)} = \mathbf{C}_{(\text{rank } 1)\cdot} \tag{5.57}$$

Therefore (and seen from another angle), if a matrix has rank r then any matrix factor of it must have rank of at least r. Since the rank cannot be greater than the smallest of k or n in a $k \times n$ matrix, this definition also limits the size (order) of factor matrices. (That is, one cannot factor a matrix of rank 2, into two matrices of which either is of less than rank 2, so k and n of each factor must also be ≥ 2).

5.7 Matrix Division (Matrix Inverse)

Matrix division can be thought of as multiplying by the *inverse*. Consider the scalar division

$$\frac{x}{b} = x\frac{1}{b} = xb^{-1}, \tag{5.58}$$

where we can write

$$bb^{-1} = 1. \tag{5.59}$$

Likewise, matrices can be effectively divided by multiplying by an inverse matrix. *Nonsingular square matrices* may have an inverse symbolized as $\mathbf{A}^{-1}$ and satisfying $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. The calculation of a matrix inverse is usually done using elimination methods on the computer. For a simple 2 x 2 matrix, its inverse is given by

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}. \tag{5.60}$$

As an example, let

$$\mathbf{A} = \begin{bmatrix} 7 & 2 \\ 10 & 3 \end{bmatrix}. \tag{5.61}$$

We solve for the inverse as

$$\mathbf{A}^{-1} = \frac{1}{21-20}\begin{bmatrix} 3 & -2 \\ -10 & 7 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -10 & 7 \end{bmatrix}, \tag{5.62}$$

and as a check we note that

$$\mathbf{A}\mathbf{A}^{-1} = \begin{bmatrix} 7 & 2 \\ 10 & 3 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -10 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}. \tag{5.63}$$

Given the concept of a matrix inverse we may summarize a few useful matrix properties:

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}, \tag{5.64}$$

$$(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1} = \mathbf{A}^{-T}, \tag{5.65}$$

$$\mathbf{D} = \mathbf{A}\mathbf{B}\mathbf{C} \text{ then } \mathbf{D}^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{5.66}$$

This "reversal rule" for inverse products may be useful for eliminating or minimizing the number of matrix inverses requiring calculation.

5.8 Matrix Manipulation and Normal Scores

We will look at a few examples of matrix manipulations. Consider the data matrix

$$\mathbf{A} = \begin{bmatrix} 1\,2\,3 \\ 4\,5\,6 \\ 7\,8\,9 \end{bmatrix} \tag{5.67}$$

and unit row vector

$$\mathbf{j}_3^T = [1\ 1\ 1]. \tag{5.68}$$

To compute the mean of each column vector in $\mathbf{A}$ (here, each column has length $n = 3$), we note that

$$\bar{\mathbf{x}}_c = \frac{1}{3}\mathbf{j}_3^T\mathbf{A}, \text{ and in general } \bar{\mathbf{x}}_c = \frac{1}{n}\mathbf{j}_n^T\mathbf{A}. \tag{5.69}$$

For our example, we find

$$\bar{\mathbf{x}}_c = \frac{1}{3}[1\ 1\ 1]\cdot\begin{bmatrix} 1\,2\,3 \\ 4\,5\,6 \\ 7\,8\,9 \end{bmatrix} = \frac{1}{3}[12\ 15\ 18] = [4\ 5\ 6]. \tag{5.70}$$

To compute the mean of each row vector in $\mathbf{A}$ (here, each row has length $k = 3$), let

$$\bar{\mathbf{x}}_r = \frac{1}{3}\mathbf{A}\mathbf{j}_3, \text{ and in general } \bar{\mathbf{x}}_r = \frac{1}{k}\mathbf{A}\mathbf{j}_k. \tag{5.71}$$

Again, for our example, we find

$$\bar{\mathbf{x}}_r = \frac{1}{3}\begin{bmatrix} 1\,2\,3 \\ 4\,5\,6 \\ 7\,8\,9 \end{bmatrix}\cdot\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 6 \\ 15 \\ 24 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}. \tag{5.72}$$

Given these terms, how can we compute normal scores for a data table (or matrix)? What we want in each cell are the elements

$$z_{ij} = \frac{a_{ij} - \bar{a}_j}{s_j}. \tag{5.73}$$

In matrix terminology we would first need to form the difference matrix, **D**, given as

$$\mathbf{D} = \mathbf{A} - \frac{1}{n}\mathbf{JA}, \tag{5.74}$$

where **J** is the $n \times n$ unit matrix (all entries equal 1). Given the diagonal matrix **S** containing the standard deviation of each column defined as

$$\mathbf{S} = \begin{bmatrix} s_1 & 0 & \dots & 0 \\ 0 & s_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & s_n \end{bmatrix} \tag{5.75}$$

we get the normal scores (here, **I** is of size $n \times n$) as

$$\mathbf{Z} = \mathbf{DS}^{-1} = \left(\mathbf{A} - \frac{1}{n}\mathbf{JA}\right)\mathbf{S}^{-1} = \left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{AS}^{-1}. \tag{5.76}$$

5.9 Solution of Simultaneous Linear Equations

A system of four simultaneous linear equations in four unknowns x_1, x_2, x_3, x_4 can be written

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 &= b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 &= b_4 \end{aligned} \tag{5.77}$$

or, in matrix form,

$$\mathbf{Ax} = \mathbf{b}, \tag{5.78}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \tag{5.79}$$

is called the coefficient matrix,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{5.80}$$

is the unknown vector, and

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{5.81}$$

is the right hand side (i.e., the observations). Premultiplying both sides by $\mathbf{A}^{-1}$ yields

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b}, \tag{5.82}$$

hence

$$\mathbf{Ix} = \mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \tag{5.83}$$

gives the solution for values of x_1, x_2, x_3, x_4 which solve the system. For simplicity, the following example solves for two simultaneous equations only. Consider two equations in two unknowns (e.g., equations of lines in the $x_1 - x_2$ plane):

$$\begin{array}{c} 5x_1 + 7x_2 = 19 \\ 3x_1 - 2x_2 = -1 \end{array} . \tag{5.84}$$

In matrix form this system translates to

$$\left[\begin{array}{cc} 5 & 7 \\ 3 & -2 \end{array} \right] \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \left[\begin{array}{c} 19 \\ -1 \end{array} \right] \tag{5.85}$$

or

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}. \tag{5.86}$$

To solve this matrix equation we need the inverse of $\mathbf{A}$, which is simply

$$\mathbf{A}^{-1} = \frac{1}{-10-21} \quad \left[\begin{array}{cc} -2 & -7 \\ -3 & 5 \end{array} \right] = \left[\begin{array}{cc} \frac{2}{31} & \frac{7}{31} \\ \frac{3}{31} & \frac{-5}{31} \end{array} \right]. \tag{5.87}$$

Then, $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, where

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \left[\begin{array}{cc} \frac{2}{31} & \frac{7}{31} \\ \frac{3}{31} & \frac{-5}{31} \end{array} \right] \left[\begin{array}{c} 19 \\ -1 \end{array} \right] = \left[\begin{array}{c} \frac{38}{31} - \frac{7}{31} \\ \frac{57}{31} + \frac{5}{31} \end{array} \right] = \left[\begin{array}{c} 1 \\ 2 \end{array} \right]. \tag{5.88}$$

So, the values $x_1 = 1$ and $x_2 = 2$ solve the above system, or

$$\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \left[\begin{array}{c} 1 \\ 2 \end{array} \right]. \tag{5.89}$$

While this approach may seem burdensome, it is good because it is extremely general and allows for a straightforward solution to very large systems. However, it is true that direct (elimination) methods to the solution are in fact quicker for fully populated matrices:

1. A solution using the inverse matrix approach involves n^3 multiplications for the inversion and n^2k more multiplications to finish the solution, where n is the number of equations per set, and k is the number of sets of equations (each of the same form but different $\mathbf{b}$ vector). The total number of multiplications is $n^3 + n^2k$.

2. A solution by directly solving the linear equations involves $n^3/3 + n^2k$ multiplications.

Hence, while the matrix form is easy to handle, one should not necessarily always use it blindly. We will consider many situations for which matrix solutions are ideal. For sparse or symmetrical matrices, the above relationships may not hold.

5.9.1 Simple regression and curve fitting

Whereas an interpolant fits each data point exactly, it is frequently advantageous to produce a smoothed fit to the data — not exactly fitting each point, but producing a "best" fit. A popular (and convenient) method for producing such fits is known as the *method of least squares*.

The method of least squares produces a fit of a specified (usually continuous) basis to a set of data points which minimizes the sum of the squared misfit (error) between the fitted curve and the data. The misfit can be measured vertically, as in Figure 5.4. This *regression* of y on x is the most commonly used method. Less common methods (i.e., more work involved) is the regression of x on y and even orthogonal regression (which we will return to later; see Figure 5.5).

Consider fitting a single "best" linear curve to n data points. This can be a scatter plot of $x(t)$, $d(t)$ plotted at similar values of t, or a simple $d = f(x)$ relationship. At any rate, d (our data) are considered a function of x (which may be a spatial coordinate or time). We wish to fit a line of the form

$$d(x) = m_1 + m_2(x - x_0) \tag{5.90}$$

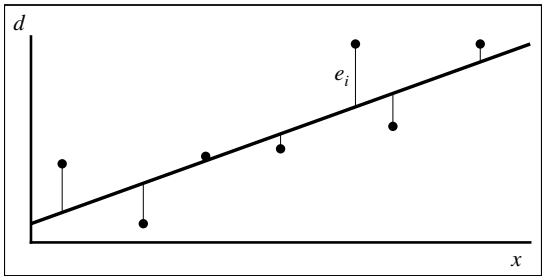


Figure 5.4: Graphical representation of the regression errors used in least-squares procedures. We measure misfit vertically in the d-direction from data point to regression curve.

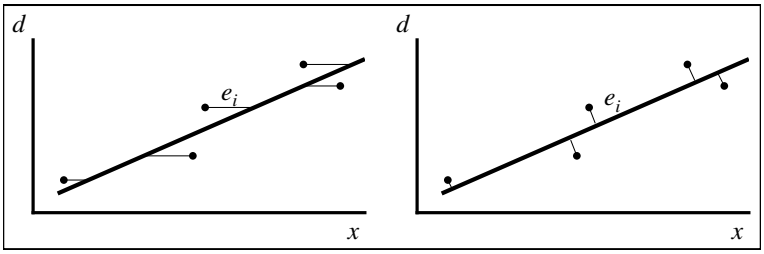


Figure 5.5: Two other regression methods: regressing x on d and orthogonal regression. Here we measure misfits horizontally from data point to regression line or orthogonally onto the regression line, respectively.

and must therefore determine values for the model coefficients m_1 and m_2 that produce a line that minimizes the sum of the squared misfits (here, x_0 is a constant specified beforehand). In other words,

$$\text{minimize} \sum_{i=1}^{n} \left[(d_{\text{computed}}(x_i) - d_{\text{observed}}(x_i) \right]^2 . \tag{5.91}$$

Ideally, for each observation d_i at location x_i we should have

$$\begin{aligned} m_1 + m_2(x_1 - x_0) &= d_1 \\ m_1 + m_2(x_2 - x_0) &= d_2 \\ m_1 + m_2(x_3 - x_0) &= d_3 \\ &\vdots \\ m_1 + m_2(x_n - x_0) &= d_n \end{aligned} \tag{5.92}$$

There are many more equations (n — one for each observed value of d) than unknowns (2 — m_1 and m_2). Such a system is *overdetermined* and there exists no unique solution (unless all the d_i's do lie exactly on a single line, in which case any two equations will uniquely determine m_1 and m_2). In matrix form,

$$\begin{bmatrix} 1 & (x_1 - x_0) \\ 1 & (x_2 - x_0) \\ \vdots & \vdots \\ 1 & (x_n - x_0) \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}, \tag{5.93}$$

i.e., $\mathbf{G} \cdot \mathbf{m} = \mathbf{d}$. Here, $\mathbf{G}$ represents how predictions of $\mathbf{d}$ are related to the model $\mathbf{m}$ and is often called the *design matrix*. However, since $\mathbf{G}$ is a not square it has no inverse, hence the equation cannot be inverted and solved as is. Consider instead the *misfit*, e_i, at each point, between prediction and observation:

$$\begin{aligned} m_1 + m_2(x_1 - x_0) - d_1 &= e_1 \\ m_1 + m_2(x_2 - x_0) - d_2 &= e_2 \\ &\vdots \\ m_1 + m_2(x_n - x_0) - d_n &= e_n \end{aligned} \tag{5.94}$$

We wish to determine the values for m_1 and m_2 that minimize

$$E(m_1, m_2) = \sum_{i=1}^{n} e_i^2 = \mathbf{e}^T \mathbf{e}, \tag{5.95}$$

where $\mathbf{e}^T = (e_1, e_2, ..., e_n)$ is the *misfit vector*. This condition will minimize the variance of the residuals about the regression line and give the desired least-squares fit. Thus, $E(m_1, m_2)$ and the minimum of this function (with respect to the two unknown coefficients) can be determined using simple differential calculus where, at the desired minimum, we require

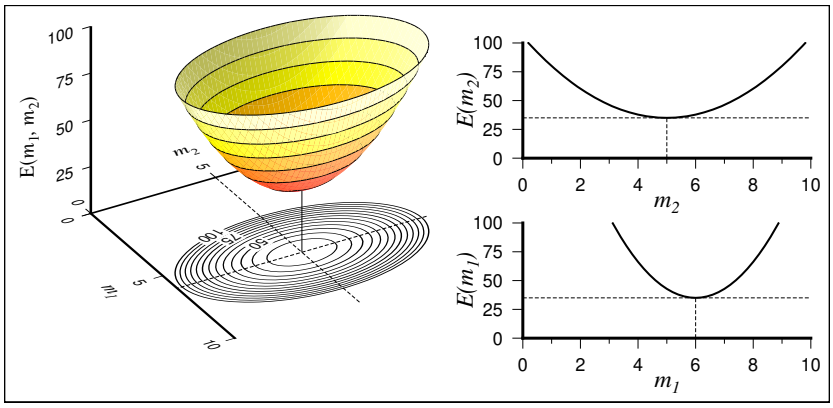


Figure 5.6: (left) The solution we seek minimizes the misfit function $E(\mathbf{m}) = E(m_1, m_2)$, which portrays a surface in 3-D. Because of the functional (quadratic) form of E we are guaranteed a unique global minimum. (right) Two orthogonal cross-sections of E along the axes m_1 and m_2. We seek the solutions for these two parameters so that the respective slopes in E are zero simultaneously.

$$\frac{\partial E(m_1, m_2)}{\partial m_1} = \frac{\partial E(m_1, m_2)}{\partial m_2} = 0. \tag{5.96}$$

Thus, the *slopes* of the misfit function with respect to each parameter must be zero (see Figure 5.6). We find

$$\begin{aligned} \frac{\partial E}{\partial m_1} &= \frac{\partial}{\partial m_1}\left(\sum_{i=1}^{n} e_i^2\right) = \frac{\partial}{\partial m_1}\left\{\sum_{i=1}^{n} [m_1 + m_2(x_i - x_0) - d_i]^2\right\} \\ &= 2\sum_{i=1}^{n} [m_1 + m_2(x_i - x_0) - d_i] = 0 \end{aligned} \tag{5.97}$$

$$\begin{aligned} \frac{\partial E}{\partial m_2} &= \frac{\partial}{\partial m_2}\left(\sum_{i=1}^{n} e_i^2\right) = \frac{\partial}{\partial m_2}\left\{\sum_{i=1}^{n} [m_1 + m_2(x_i - x_0) - d_i]^2\right\} \\ &= 2\sum_{i=1}^{n} [m_1 + m_2(x_i - x_0) - d_i](x_i - x_0) = 0. \end{aligned} \tag{5.98}$$

These two equations can now be expanded into their individual terms, forming what are known as the *normal equations*. This system of two equations with two unknowns can be uniquely solved. Rearranging, we find

$$nm_1 + m_2 \sum_{i=1}^{n} (x_i - x_0) = \sum_{i=1}^{n} d_i, \tag{5.99}$$

$$m_1 \sum_{i=1}^{n} (x_i - x_0) + m_2 \sum_{i=1}^{n} (x_i - x_0)^2 = \sum_{i=1}^{n} d_i (x_i - x_0). \tag{5.100}$$

Notice that all sums involve known values that add to simple constants. Specifically, we must compute the sums

$$S_y = \sum_{i=1}^{n} d_i,\ S_{xy} = \sum_{i=1}^{n} d_i (x_i - x_0),\ S_x = \sum_{i=1}^{n} (x_i - x_0),\ \text{and } S_{xx} = \sum_{i=1}^{n} (x_i - x_0)^2. \tag{5.101}$$

Substituting these symbols into (5.99) and (5.100), we obtain

$$nm_1 + m_2 S_x = S_y \tag{5.102}$$

$$m_1 S_x + m_2 S_{xx} = S_{xy} \tag{5.103}$$

Solving for the intercept yields

$$m_1 = \frac{1}{n} S_y - \frac{m_2}{n} S_x. \tag{5.104}$$

We substitute m_1 into (5.103) and find

$$\left[\frac{1}{n} S_y - \frac{m_2}{n} S_x \right] S_x + m_2 S_{xx} = S_{xy}. \tag{5.105}$$

Now solve for m_2:

$$\frac{1}{n} S_y S_x - \frac{m_2}{n} S_x^2 + m_2 S_{xx} = S_{xy}, \tag{5.106}$$

$$m_2 \left(S_{xx} - \frac{1}{n} S_x^2 \right) = S_{xy} - \frac{1}{n} S_y S_x. \tag{5.107}$$

Finally,

$$m_2 = \left(S_{xy} - \frac{1}{n} S_y S_x \right) / \left(S_{xx} - \frac{1}{n} S_x^2 \right) = \frac{n S_{xy} - S_x S_y}{n S_{xx} - S_x^2}, \tag{5.108}$$

and we substitute m_2 into (5.104) to find

$$m_1 = \frac{S_{xx} S_y - S_x S_{xy}}{n S_{xx} - S_x^2}. \tag{5.109}$$

In matrix form the normal equations are

$$\begin{bmatrix} n & \sum_{i=1}^{n} (x_i - x_0) \\ \sum_{i=1}^{n} (x_i - x_0) & \sum_{i=1}^{n} (x_i - x_0)^2 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} d_i \\ \sum_{i=1}^{n} d_i (x_i - x_0) \end{bmatrix}, \tag{5.110}$$

which may be simplified to

$$\begin{bmatrix} n & S_x \\ S_x & S_{xx} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} S_y \\ S_{xy} \end{bmatrix} \tag{5.111}$$

or simply $\mathbf{Nm} = \mathbf{v}$. Since $\mathbf{N}$ is square, symmetric and of full rank, this equation is solved in the standard manner:

$$\mathbf{N}^{-1}\mathbf{N}\mathbf{m} = \mathbf{m} = \mathbf{N}^{-1}\mathbf{v}. \tag{5.112}$$

This problem was simple enough (2×2) to solve for m_1 and m_2 by brute force. For larger systems, that approach becomes impractical and instead a matrix solution to the rectangular $\mathbf{G} \cdot \mathbf{m} = \mathbf{d}$ equation must be sought. We will next look at the general linear least-squares problem and find a solution in matrix notation.

5.9.2 General linear least squares method, version 1

We have looked at a few special cases where we have sought to fit a model to data in a least-squares sense. Fitting a straight line to the (x_i, d_i) points was a very simple example of this technique. We will now look at the more general problem of finding the coefficients for *any* linear combination of a chosen set of basis functions that fits a data set in a least squares sense. There are numerous situations where this is needed; some are listed in Table 5.1.

While the basis functions in these cases are all vastly different, they are all used in *linear combinations* to fit the observed data. We will therefore take time to investigate how such a problem is set up, and how the setup can be simplified with matrix algebra. Some typical basis functions are given in Table 5.2.

Consider the least squares fitting of any continuous basis of the form

$$g_1(x), g_2(x), g_3(x), \cdots, g_k(x). \tag{5.113}$$

Situation	Model Parameters	Data
Curve fitting	Coefficients of polynomials, Fourier series, etc.	Points in $x-y$ plane
Gravity modeling	Densities of subsurface polygons	Gravity observations
Hypocenter location	Small perturbations to hypocenter location	Seismic arrival times

Table 5.1: Examples of situations where linear least squares solutions are used.

Polynomial basis	Fourier sine basis
$g_1 = x^0$	$g_1 = \sin(2\pi x/T)$
$g_2 = x^1$	$g_2 = \sin(4\pi x/T)$
$g_3 = x^2$	$g_3 = \sin(6\pi x/T)$
$\vdots$	$\vdots$
$g_k = x^{k-1}$	$g_k = \sin(2k\pi x/T)$

Table 5.2: Examples of basis functions used for modeling of data.

For example, we desire to fit a model with k terms

$$d(x) = m_1 g_1(x) + m_2 g_2(x) + \cdots + m_k g_k(x) \tag{5.114}$$

to a data set of n data points, where $n > k$, by minimizing $E(\mathbf{m})$ given by

$$E(\mathbf{m}) = E(m_1, m_2, \cdots, m_k) = \sum_{i=1}^{n} (e_i)^2 = \sum_{i=1}^{n} (m_1 g_1(x_i) + m_2 g_2(x_i) + \cdots + m_k g_k(x_i) - d_i)^2, \tag{5.115}$$

or simply

$$E(\mathbf{m}) = \sum_{i=1}^{n} (m_1 g_{i1} + m_2 g_{i2} + \cdots + m_k g_{ik} - d_i)^2, \tag{5.116}$$

where d_i is the observed value and g_{ij} is the j'th basis function, evaluated at the location (or time) x_i. In other words, $g_{ij} = g_j(x_i)$.

There are *four* concepts of vital importance in a general linear least squares modeling problem:

1. The *observed data*, $(x_i, d_i), i = 1, n$, where n is the number of observations. These are all known quantities.
2. The *general linear model* (linear in $\mathbf{m} = m_j, j = 1, k$, with k unknown parameters), given by (5.114).
3. The m chosen *basis functions*, $g_j(x), j = 1, k$. We can evaluate these for any x.
4. The *least squares misfit criteria*, given by (5.116).

We can write a linear system of equations for the misfit at each data point:

$$\begin{aligned} m_1 g_{11} + m_2 g_{12} + \cdots + m_k g_{1k} - d_1 &= e_1 \\ m_1 g_{21} + m_2 g_{22} + \cdots + m_k g_{2k} - d_2 &= e_2 \\ &\vdots \\ m_1 g_{n1} + m_2 g_{n2} + \cdots + m_k g_{nk} - d_n &= e_n \end{aligned} \quad . \tag{5.117}$$

To minimize E, we require

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = 0, \quad j = 1, k. \tag{5.118}$$

Considering the first term (case $j = 1$), we see

$$\begin{aligned}\frac{\partial E(\mathbf{m})}{\partial m_1} &= \frac{\partial}{\partial m_1}\sum_{i=1}^{n}(m_1 g_{i1} + m_2 g_{i2} + \cdots m_k g_{ik} - d_i)^2 \\ &= 2\sum_{i=1}^{n}(m_1 g_{i1} + m_2 g_{i2} + \cdots m_k g_{ik} - d_i)g_{i1} = 0,\end{aligned} \tag{5.119}$$

while for the second term (case $j = 2$), we find

$$\begin{aligned}\frac{\partial E(\mathbf{m})}{\partial m_2} &= \frac{\partial}{\partial m_2}\sum_{i=1}^{n}(m_1 g_{i1} + m_2 g_{i2} + \cdots m_k g_{ik} - d_i)^2 \\ &= 2\sum_{i=1}^{n}(m_1 g_{i1} + m_2 g_{i2} + \cdots m_k g_{ik} - d_i)g_{i2} = 0.\end{aligned} \tag{5.120}$$

Consequently, for the j'th parameter,

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = 2\sum_{i=1}^{n}(m_1 g_{i1} + m_2 g_{i2} + \cdots m_k g_{ik} - d_i)g_{ij} = 0. \tag{5.121}$$

Rearranging these normal equations gives the square $k \times k$ system

$$\begin{aligned}m_1\sum_{i=1}^{n} g_{i1}^2 + m_2\sum_{i=1}^{n} g_{i2}g_{i1} + \cdots + m_k\sum_{i=1}^{n} g_{ik}g_{i1} &= \sum_{i=1}^{n} d_i g_{i1} \\ m_1\sum_{i=1}^{n} g_{i1}g_{i2} + m_2\sum_{i=1}^{n} g_{i2}^2 + \cdots + m_k\sum_{i=1}^{n} g_{ik}g_{i2} &= \sum_{i=1}^{n} d_i g_{i2} \\ &\vdots \\ m_1\sum_{i=1}^{n} g_{i1}g_{ik} + m_2\sum_{i=1}^{n} g_{i2}g_{ik} + \cdots + m_k\sum_{i=1}^{n} g_{ik}^2 &= \sum_{i=1}^{n} d_i g_{ik}\end{aligned} \tag{5.122}$$

or equivalently,

$$m_1\sum_{i=1}^{n} g_{i1}g_{ij} + m_2\sum_{i=1}^{n} g_{i2}g_{ij} + \cdots + m_k\sum_{i=1}^{n} g_{ik}g_{ij} = \sum_{i=1}^{n} d_i g_{ij}, \quad j = 1,k. \tag{5.123}$$

This setup provides a *closed system* of k normal equations. In matrix form,

$$\begin{bmatrix} \sum_{i=1}^{n} g_{i1}^2 & \sum_{i=1}^{n} g_{i2}g_{i1} & \cdots & \sum_{i=1}^{n} g_{ik}g_{i1} \\ \sum_{i=1}^{n} g_{i1}g_{i2} & \sum_{i=1}^{n} g_{i2}^2 & \cdots & \sum_{i=1}^{n} g_{ik}g_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} g_{i1}g_{ik} & \sum_{i=1}^{n} g_{i2}g_{ik} & \cdots & \sum_{i=1}^{n} g_{ik}^2 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} d_i g_{i1} \\ \sum_{i=1}^{n} d_i g_{i2} \\ \vdots \\ \sum_{i=1}^{n} d_i g_{ik} \end{bmatrix}. \tag{5.124}$$

Once again, we simply have

$$\mathbf{N}\cdot\mathbf{m} = \mathbf{v}, \tag{5.125}$$

where $\mathbf{N}$ is the (known) coefficient matrix, $\mathbf{m}$ the vector with the unknowns m_j, and $\mathbf{v}$ contains weighted sums of known (observed or computable) quantities. Solving for the $\mathbf{m}$ vector (since $\mathbf{N}$ is square, symmetric and of full rank) yields

$$\mathbf{N}^{-1}\cdot\mathbf{N}\cdot\mathbf{m} = \mathbf{m} = \mathbf{N}^{-1}\cdot\mathbf{v}. \tag{5.126}$$

The resulting m_j values are the ones which satisfy (5.118) and therefore the same ones, when combined with the chosen basis, that produce the "best" fit to the n data points such that (5.116) is minimized.

5.9.3 General linear least squares method, version 2

We will now look at a simpler approach to the same problem using matrix algebra. We have $\mathbf{e} = \mathbf{G}\cdot\mathbf{m} - \mathbf{d}$, or

$$\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1k} \\ g_{21} & g_{22} & \cdots & g_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nk} \end{bmatrix} \cdot \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{bmatrix} - \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}. \tag{5.127}$$

We note that each column vector of $\mathbf{G}$ is simply a single basis function evaluated at all our observation points. In fact, we could write $\mathbf{G}$ as

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_k \end{bmatrix}, \tag{5.128}$$

where

$$\mathbf{g}_j = \begin{bmatrix} g_j(x_1) & g_j(x_2) & \cdots & g_j(x_n) \end{bmatrix}^T. \tag{5.129}$$

We wish to find the m_j values that minimize $E = \mathbf{e}^T\mathbf{e}$. Minimizing the misfit with respect to the unknown model parameters $\mathbf{m}$ means we must solve the k linear equations that result from setting all partial derivatives of E to zero (i.e., 5.118). Using matrix algebra, we express the *predicted* solution as $\hat{\mathbf{d}} = \mathbf{G}\cdot\mathbf{m}$. We may now express the misfit between model and observations as $\mathbf{e} = \hat{\mathbf{d}} - \mathbf{d} = \mathbf{G}\cdot\mathbf{m} - \mathbf{d}$ and use this expression to evaluate the misfit as

$$E(\mathbf{m}) = \sum_{i=1}^{n}(e_i)^2 = \mathbf{e}^T\cdot\mathbf{e} = \left(\hat{\mathbf{d}} - \mathbf{d}\right)^T\cdot\left(\hat{\mathbf{d}} - \mathbf{d}\right) = (\mathbf{G}\cdot\mathbf{m} - \mathbf{d})^T\cdot(\mathbf{G}\cdot\mathbf{m} - \mathbf{d}). \tag{5.130}$$

Expanding terms, we find

$$E(\mathbf{m}) = \left(\mathbf{m}^T\mathbf{G}^T - \mathbf{d}^T\right)\cdot(\mathbf{G}\cdot\mathbf{m} - \mathbf{d}) = \mathbf{m}^T\mathbf{G}^T\mathbf{G}\mathbf{m} - \mathbf{m}^T\mathbf{G}^T\mathbf{d} - \mathbf{d}^T\mathbf{G}\mathbf{m} + \mathbf{d}^T\mathbf{d}, \tag{5.131}$$

where we have used (5.36) to handle the transpose of a matrix product. Note that as E is a *scalar* then each of these terms must evaluate to scalars as well. To find the solution, we set

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = \dot{\mathbf{m}}^T\mathbf{G}^T\mathbf{G}\mathbf{m} + \mathbf{m}^T\mathbf{G}^T\mathbf{G}\dot{\mathbf{m}} - \dot{\mathbf{m}}^T\mathbf{G}^T\mathbf{d} - \mathbf{d}^T\mathbf{G}\dot{\mathbf{m}} = 0, \quad j = 1, m, \tag{5.132}$$

where the "dot" over a vector represents the derivative of that vector with respect to m_j. We note the first and second terms are transposes of each other, as are the third and fourth terms. However, since they all evaluate to scalars the two transposes must be identical and hence this repetition simply constitutes a factor of two, which we delete by retaining only the first and third term:

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = \dot{\mathbf{m}}^T\mathbf{G}^T\mathbf{G}\mathbf{m} - \dot{\mathbf{m}}^T\mathbf{G}^T\mathbf{d} = 0, \quad j = 1, k. \tag{5.133}$$

What does the mysterious "dot"-derivative, written as

$$\dot{\mathbf{m}}^T = \frac{\partial}{\partial m_j}\left(\mathbf{m}^T\right), j = 1, k, \tag{5.134}$$

mean? We illuminate this term by trying some values of j, remembering $\mathbf{m}^T = [m_1\ m_2\ \cdots\ m_k]$:

$$\begin{aligned} \text{Case } j = 1 &: \tfrac{\partial}{\partial m_1}\mathbf{m} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T \\ \text{Case } j = 2 &: \tfrac{\partial}{\partial m_2}\mathbf{m} = \begin{bmatrix} 0 & 1 & \cdots & 0 \end{bmatrix}^T \\ &\vdots \\ \text{Case } j = k &: \tfrac{\partial}{\partial m_k}\mathbf{m} = \begin{bmatrix} 0 & 0 & \cdots & 1 \end{bmatrix}^T \end{aligned}$$

Thus, the k linear equations can be combined into a single matrix equation, noting that all these derivatives (each producing a row vector) combine to form the identity matrix, $\mathbf{I}$:

$$\frac{\partial}{\partial m_j}\left(\mathbf{m}^T\right), j = 1, k \rightarrow \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{I}. \tag{5.135}$$

Hence, we may write

$$\mathbf{I}\mathbf{G}^T\mathbf{G}\mathbf{m} - \mathbf{I}\mathbf{G}^T\mathbf{d} = \mathbf{0}, \tag{5.136}$$

or by rearranging,

$$\mathbf{G}^T\mathbf{G}\mathbf{m} = \mathbf{G}^T\mathbf{d}. \tag{5.137}$$

Because the $\mathbf{G}^T\mathbf{G}$ matrix is square and symmetric and thus can be inverted, we simply multiply by its inverse and obtain the general least squares solution as

$$\mathbf{m} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{d}. \tag{5.138}$$

Comparing (5.138) with (5.125) we see clearly that $\mathbf{N} = \mathbf{G}^T\mathbf{G}$ and $\mathbf{v} = \mathbf{G}^T\mathbf{d}$. Furthermore, given (5.128) we may write $\mathbf{G}^T\mathbf{G}$ using the product

$$\mathbf{N} = \mathbf{G}^T\mathbf{G} = \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \vdots \\ \mathbf{g}_k^T \end{bmatrix} \cdot \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_k \end{bmatrix} = \begin{bmatrix} \mathbf{g}_1^T\mathbf{g}_1 & \mathbf{g}_1^T\mathbf{g}_2 & \cdots & \mathbf{g}_1^T\mathbf{g}_k \\ \mathbf{g}_2^T\mathbf{g}_1 & \mathbf{g}_2^T\mathbf{g}_2 & \cdots & \mathbf{g}_2^T\mathbf{g}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{g}_k^T\mathbf{g}_1 & \mathbf{g}_k^T\mathbf{g}_2 & \cdots & \mathbf{g}_k^T\mathbf{g}_k \end{bmatrix}, \tag{5.139}$$

which makes it clear that each element of $\mathbf{N}$, such as n_{pq}, is the dot product between two basis vectors $\mathbf{g}_p^T$ and $\mathbf{g}_q$, and

$$\mathbf{v} = \mathbf{G}^T\mathbf{d} = \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \vdots \\ \mathbf{g}_k^T \end{bmatrix} \cdot \mathbf{d} = \begin{bmatrix} \mathbf{g}_1^T\mathbf{d} \\ \mathbf{g}_2^T\mathbf{d} \\ \vdots \\ \mathbf{g}_k^T\mathbf{d} \end{bmatrix}, \tag{5.140}$$

which shows each element of $\mathbf{v}$ is the dot product between each basis function $\mathbf{g}_j^T$ and the data vector $\mathbf{d}$. This is simply what we found the hard way earlier (i.e., 5.124). Thus, to solve a general linear least squares problem, all we have to do is to evaluate $\mathbf{G}$ via (5.128) and the rest is taken care of by (5.138).

Example 5–1. We are given a data set with four data pairs (2,1), (4,4), (6,3) and (8,4) ($n = 4$) and asked to determine the coefficients for a *quadratic* curve that best describes the data. Except for special situations, we know that the three-parameter curve will not pass through all the four points, so we decide to seek a least squares solution.

We write down the functional form for our quadratic curve as $d = m_1 + m_2x + m_3x^2$ and use it to form the matrix equation

$$\begin{aligned} m_1 + m_2x_1 + m_3x_1^2 &= d_1 \\ m_1 + m_2x_2 + m_3x_2^2 &= d_2 \\ m_1 + m_2x_3 + m_3x_3^2 &= d_3 \\ m_1 + m_2x_4 + m_3x_4^2 &= d_4 \end{aligned}, \tag{5.141}$$

which, for our data, yields the linear system

$$\begin{bmatrix} 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 6 & 36 \\ 1 & 8 & 64 \end{bmatrix} \cdot \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 4 \end{bmatrix}. \tag{5.142}$$

Using (5.138) we find the solution to be $d = -1.5 + 1.65x - 0.125x^2$. Figure 5.7 shows our data as well as the fitted quadratic curve.

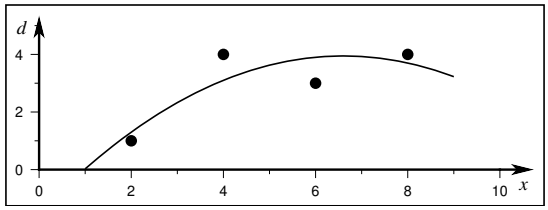


Figure 5.7: Fitting a three-parameter quadratic curve to four data points by minimizing the least squares misfit.

5.9.4 Weighted least squares solution

What if some data constraints are more reliable than others? We may simply give those residuals more weight than the others, i.e.,

$$\mathbf{e}' = \begin{bmatrix} 0.8e_1 \\ 2.1e_2 \\ \vdots \\ 0.7e_n \end{bmatrix}. \tag{5.143}$$

In general, we can assign weights w_i to each misfit so that the new weighted misfits become $e'_i = e_i \cdot w_i$. Very often, the weights will simply be $s_{ii} = 1/\sigma_i$, where σ_i is the one-sigma uncertainty in the i'th measurement, d_i. We implement such weights by introducing a diagonal weight matrix

$$\mathbf{S} = \begin{bmatrix} s_{11} & & & \\ & s_{22} & & \\ & & \ddots & \\ & & & s_{nn} \end{bmatrix}, \tag{5.144}$$

which means the weighted residuals are $\mathbf{S}\cdot\mathbf{e}$ and the sum of the squared errors, E, becomes

$$E = (\mathbf{S}\cdot\mathbf{e})^T(\mathbf{S}\cdot\mathbf{e}) = \mathbf{e}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{e} = \mathbf{e}^T\cdot\mathbf{W}\cdot\mathbf{e}, \tag{5.145}$$

where we have introduced $\mathbf{W} = \mathbf{S}^T\mathbf{S}$. Since $\mathbf{S}\cdot\mathbf{e} = \mathbf{S}\cdot(\mathbf{G}\cdot\mathbf{m}-\mathbf{d})$ we obtain

$$\begin{aligned} E(\mathbf{m}) &= (\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{S}\cdot\mathbf{d})^T\cdot(\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{S}\cdot\mathbf{d}) = (\mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{S}^T-\mathbf{d}^T\cdot S^{\mathbf{T}})\cdot(\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{S}\cdot\mathbf{d}) \\ &= \mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{d}-\mathbf{d}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}+\mathbf{d}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{d}. \end{aligned} \tag{5.146}$$

We substitute $\mathbf{W} = \mathbf{S}^T\mathbf{S}$, take the partial derivatives, and obtain

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = \mathbf{0} = \dot{\mathbf{m}}^T\cdot\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\mathbf{m}+\mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\dot{\mathbf{m}}-\dot{\mathbf{m}}^T\cdot\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{d}-\mathbf{d}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\dot{\mathbf{m}}, \quad j = 1,k. \tag{5.147}$$

Since $\mathbf{m}$ only contains the m_j, we know $\dot{\mathbf{m}}^T = \dot{\mathbf{m}} = \mathbf{I}$. We again find the k normal equations can be written more compactly as the single matrix equation

$$\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\mathbf{m}+\mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}-\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{d}-\mathbf{d}^T\cdot\mathbf{W}\cdot\mathbf{G} = \mathbf{0}. \tag{5.148}$$

As before, the second and fourth terms are the transposes of the first and third terms, and as they all represent the same terms our equation reduces to

$$\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{d} = \mathbf{0}. \tag{5.149}$$

Thus, the weighted linear least squares solution is

$$\mathbf{m} = \left[\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\right]^{-1}\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{d}. \tag{5.150}$$

This solution is universal and applies to *any* linear least squares problem one can imagine. In the particular case when all $s_{ii} = 1$ the solution reduces to (5.138).

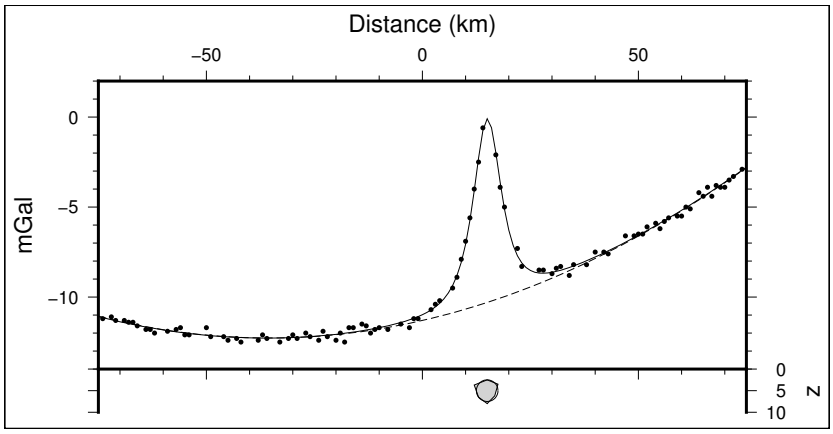


Figure 5.8: Observed and modeled gravity anomalies over a dense ore body. While the model is nonlinear in x and z, it is *linear* in the coefficients p_i.

Example 5–2. We will try the least squares machinery on an example taken from exploration geophysics. Figure 5.8 shows how observed gravity anomalies (d_i, solid circles) vary over a buried dense ore body as a function of location x_i. Based on the inferred geometry of the ore (from subsurface geology seen in mine shafts, etc.) we expect that a first-order approximation to the ore body could be a sphere buried at a depth of 5 km, with a radius of 2.5 km, and located at 15 km to the right of the origin. We would like to determine the density of that ore body. Exploration geophysics textbooks tell us that the gravity anomaly over a buried sphere of radius r and *unit* density is

$$g_{sp}(x,z,r) = \gamma \frac{\frac{4}{3}\pi r^3 z}{(x^2+z^2)^{3/2}}, \tag{5.151}$$

where z is the depth to the center of the sphere, $x = 0$ is where the sphere is located, and γ is the universal gravitational constant ($6.674 \cdot 10^{-11}$ $\mathrm{m^3 kg^{-1} s^{-2}}$). However, inspection of the data suggests that the anomaly due to the ore body is superimposed on a regional field with some curvature to it (i.e., dashed trend in Figure 5.8). Therefore, we decide to model these anomalies as a sum of a quadratic background (regional) field and the attraction of the sphere; this is accomplished with the four-parameter linear model

$$g(x) = m_1 + m_2 x + m_3 x^2 + m_4 g_{sp}(x-15,5,2.5), \tag{5.152}$$

where $m_4 = \Delta\rho$, the density contrast between the ore and the host rock. In the parlance of the previous sections, our basis functions $g_j(x)$ are $\{1,\ x,\ x^2,$ and $g_{sp}(x)\}$. To solve the problem we need to evaluate the matrix equation $\mathbf{G}\cdot\mathbf{m} = \mathbf{d}$, i.e.,

$$\begin{bmatrix} 1 & x_1 & x_1^2 & g_{sp}(x_1-15,5,2.5) \\ 1 & x_2 & x_2^2 & g_{sp}(x_2-15,5,2.5) \\ 1 & x_3 & x_3^2 & g_{sp}(x_3-15,5,2.5) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & g_{sp}(x_n-15,5,2.5) \end{bmatrix} \cdot \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{bmatrix}, \tag{5.153}$$

whose solution becomes

$$\mathbf{m} = [\mathbf{G}^T \cdot \mathbf{G}]^{-1} \mathbf{G}^T \cdot \mathbf{d}. \tag{5.154}$$

Thus, we have solved a fairly complicated least squares modeling problem (solution is the solid line in Figure 5.8).

5.10 Problems for Chapter 5

Problem 5.1. If the major product moment $\mathbf{A}^T\mathbf{A} = \mathbf{0}$, show that all elements in $\mathbf{A}$ must be zero, i.e. $a_{ij} = 0$. (Hint: what does an element in $\mathbf{A}^T\mathbf{A}$ represent?)

Problem 5.2. An analytical experiment yields the following data pairs t and $d(t)$:

t	d
-0.82	-0.86
0.23	-0.58
1.35	0.54
2.25	1.30
3.33	2.20

a) Using the solutions for least-squares regression lines, i.e., (5.109) and (5.108) above, determine the slope and intercept of the best-fitting line.

b) Plot this line and the data points.

c) Redo the problem using matrix algebra. Write down the various components in the matrix equation $\mathbf{G}\cdot\mathbf{m} = \mathbf{d}$. What does $\mathbf{G}\cdot\mathbf{m}$ represent?

d) Evaluate the matrix solution (i.e., $\mathbf{m} = [\mathbf{G}^T\mathbf{G}]^{-1}\mathbf{G}^T\mathbf{d}$) and make sure it matches your answer in (a).

e) What is the total misfit, E?

Problem 5.3. A detailed study of seafloor heat flow was conducted along a transect perpendicular to a mid-ocean ridge. The results are reported as distance (km), heat flow (mWm^{-2}) pairs in the file *heatflow.txt*. Because of hydrothermal circulation, we choose to model the decay of heat flow with distance from the ridge as

$$q(x) = q_0 + q_1 x + \frac{q_2}{x}.$$

a) Using matrix algebra, what are the least-squares values for the model coefficients?

b) Plot the data as points and the model as a solid line evaluated every km from 25 km to 250 km on the same graph.

c) What is the r.m.s (root-mean-square) misfit (r.m.s = $\sqrt{E/n}$)? Be sure to give units.

d) Using your error analysis skills, what is the estimated heat flow at a distance of $x = 25 \pm 1$ km from the ridge axis? (Hint: Consider the uncertainty in the function $q(x)$.)

Problem 5.4. The file *faultstep.txt* contains distance (m) and relief (m) for a topographic profile across an small normal fault. We want to estimate the total vertical offset across the fault. There is also a linear trend in the data because the the fault sits on a gently dipping monocline.

a) Make a linear model that includes the trend and the step. You may want to use the Heaviside step function (`heaviside` in MATLAB), defined as

$$H(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & x = 0 \\ 1, & x > 0 \end{cases}$$

What is your model equation for the topographic profile?

b) Let your best guess for the fault location be $x_0 = 145$ m and use matrix algebra to solve for the model parameters. Write the first few terms of the matrix equation $\mathbf{G} \cdot \mathbf{m} = \mathbf{d}$ so you can see the pattern. Plot your model prediction on top of the data and indicate the values of the parameters. What is the rms misfit? What is the size of the fault offset?

c) Guessing x_0 is not very robust. Determine the best choice for x_0 that minimizes the total misfit E by trying a range of x_0 (hint: evaluate $E(x_0)$ for a dense set of x_0 values covering the likely range, plot E versus x_0, and find the value of x_0 where E has its minimum). What is your final model parameters for the fault profile? Plot it on top of the data. What is your minimum r.m.s. value? What is the final size of the fault offset?

Problem 5.5. As sedimentary layers become buried they compact, reducing the porosity of the unit. The empirical relation between depth (z) and porosity (θ) in well-compacted sediments is called *Athy's law*:

$$\theta = \theta_0 e^{-\alpha z},$$

where θ_0 and α are constants that vary with sediment type and location, and z is depth of burial.

a) Given the data set below, find the weighted least-squares estimates of θ_0 and α using matrix algebra, taking the uncertainties in θ into account (note: Your first idea on how to do that is likely to be wrong!). Plot the data and your Athy's law. (Hint: Athy's law as given is not linear! You must transform it first.)

b) What is the predicted porosity ($\pm$ uncertainty) at a depth of 6 km ($\pm$200m)?

Depth (m)	**Porosity (%)**
650	38 ± 5.0
1000	35 ± 4.0
2050	24 ± 2.5
2950	18 ± 2.0
4075	14 ± 1.5
5030	9.8 ± 1.0

Problem 5.6. The rim of Halemaumau crater within Kilauea caldera has been digitized and its UTM coordinates (x_i, d_i) are listed in *halemaumau.txt*. We wish to approximate this shape by a perfect circle with parameters (x_0, d_0, r).

a) What is the misfit function, E, that we want to minimize using the least-squares criterion?

b) Determine the three parameters for the circular model by minimizing E.

c) Estimate the area of the crater and use the standard deviation of the radial residual misfits to assign one-sigma confidence bounds on the area.

Problem 5.7. The file *noisy.txt* contains observations of a phenomenon known to oscillate at a single frequency, ω. This signal is superimposed on a linear trend in the presence of some random noise.

a) Assume you know the frequency ω. What is the form of the model you could use to fit the data? Hint: $A\cos(\omega t - \phi) = a\cos(\omega t) + b\sin(\omega t)$.

b) Plot the data and eye-ball the period and use it to determine the frequency. Use least-squares to fit your model and plot it. What is the misfit, E, for this trial model?

c) Try a range of frequencies centered on your best guess and determine the frequency ω_0 that minimizes the misfit. Plot the optimal model, and in a separate diagram plot the misfits you obtained versus the frequencies you explored.

Problem 5.8. Having crash-landed on an alien planet you worry you may not have enough fuel to reach escape velocity. To find out you need to know the planet's gravitational attraction, g_p (which on Earth is 9.81 m s^{-2}). To get an estimate you decide to drop rocks off cliffs of various heights and clock the time it takes for them to reach the valley floor. Your data are listed in the file *drops.txt*. You vaguely remember your high-school physics exploits where you learned that the vertical drop in vacuum is related to drop time given by $h = \frac{1}{2} g_p t^2$. Plotting the data you realize that the height measurements obtained with your damaged laser rangefinder is biased by some unknown but constant offset. Aargh, there is always something....

a) What is a suitable linear model that will relate your observations and unknowns?

b) Determine the planet's gravity and the bias in your instrument using the linear least squares method.

Problem 5.9. An elongated sedimentary basin is estimated to have a width of 20 km and a depth of 4 km. The gravity anomalies measured across the basin are given in *basingrav.txt* and show a broad negative anomaly due to the lower density sediments relative to the surrounding higher density igneous bedrock, but some lateral regional trend is also apparent. You decide to model the anomalies as a linear regional trend plus the attraction of a prism-shaped two-dimensional basin of given dimensions. You may use the MATLAB function `g_basin.m` to compute the gravity of the prism.

a) Determine the least-squares solution for the density contrast. If the bedrock density is 2670 kg m^{-3}, what is your estimate of the sediment density?

b) A seismic crew collects data that throw some doubt on your depth estimate of 4 km. You decide to repeat your modeling for a range of depths (from 3 to 6 km in steps of 100 m) and keep track of the misfit E as a function of the chosen depth. Find the optimal depth that minimizes the misfit and report this depth and the corresponding density contrast.

Chapter 6

REGRESSION

"The invalid assumption that correlation implies cause is probably among the two or three most serious and common errors of human reasoning."

Stephen Jay Gould, Paleontologist

Regression refers to a subset of data modeling where we fit a simple model with a linear trend in one (or more) dimensions. Usually, we also include a constant (intercept) term. Entire books have been written about regression and all the various methods, norms, and misfit-minimizations possible. A summary of some of these developments will be given below.

6.1 Line-Fitting Revisited

We will again consider the best-fitting line problem $y = m_1 + m_2 x$, this time with errors s_i in the observed y-values. We want to measure how well the model agrees with the data, and for this purpose we will use the χ^2 function, i.e.,

$$\chi^2(m_1, m_2) = \sum_{i=1}^{n} \left(\frac{y_i - m_1 - m_2 x_i}{s_i} \right)^2 . \tag{6.1}$$

Minimizing χ^2 will give the best weighted least squares solution. Again, we set the partial derivatives to zero and obtain

$$\begin{aligned} \frac{\partial \chi^2}{\partial m_1} &= 0 = -2 \sum_{i=1}^{n} \left(\frac{y_i - m_1 - m_2 x_i}{s_i^2} \right), \\ \frac{\partial \chi^2}{\partial m_2} &= 0 = -2 \sum_{i=1}^{n} \left(\frac{y_i - m_1 - m_2 x_i}{s_i^2} \right) x_i. \end{aligned} \tag{6.2}$$

Let us define the following terms (unless noted, all sums go from $i = 1$ to n):

$$S = \sum \frac{1}{s_i^2}, \quad S_x = \sum \frac{x_i}{s_i^2}, \quad S_y = \sum \frac{y_i}{s_i^2}, \quad S_{xx} = \sum \frac{x_i^2}{s_i^2}, \quad S_{xy} = \sum \frac{x_i y_i}{s_i^2}. \tag{6.3}$$

Then, (6.2) reduces to

$$\begin{aligned} m_1 S + m_2 S_x &= S_y, \\ m_1 S_x + m_2 S_{xx} &= S_{xy}. \end{aligned} \tag{6.4}$$

Introducing

$$\Delta = S S_{xx} - S_x^2 \tag{6.5}$$

we find

$$\begin{aligned} m_1 &= \frac{S_{xx}S_y - S_xS_{xy}}{\Delta}, \\ m_2 &= \frac{SS_{xy} - S_xS_y}{\Delta}. \end{aligned} \tag{6.6}$$

All this is swell but we must also estimate the uncertainties in m_1 and m_2. For the same s_i we may get large differences

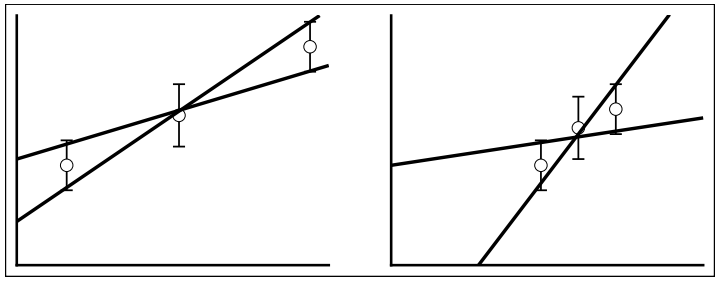


Figure 6.1: The uncertainty in the line fit depends to a large extent on the distribution of the x-positions as well as the uncertainties in the y-values.

in the uncertainties in m_1 and m_2 (e.g., Figure 6.1). As shown in Chapter 2, consideration of the propagation of errors (e.g., 2.37) implies that the variance σ_f^2 in the value of any function is

$$\sigma_f^2 = \sum \left(\frac{\partial f}{\partial y_i} \sigma_i \right)^2, \tag{6.7}$$

where we now consider f to be a function of all the n independent parameters y_i. For our model, f is either m_1 or m_2 so the partial derivatives become

$$\begin{aligned} \frac{\partial m_1}{\partial y_i} &= \frac{S_{xx} - S_x x_i}{s_i^2 \Delta}, \\ \frac{\partial m_2}{\partial y_i} &= \frac{Sx_i - S_x}{s_i^2 \Delta}. \end{aligned} \tag{6.8}$$

Inserting in turn these terms into (6.7) now gives

$$\begin{aligned} s_{m_1}^2 &= \sum s_i^2 \left[\frac{S_{xx} - S_x x_i}{s_i^2 \Delta} \right]^2 = \sum \frac{S_{xx}^2 - 2S_{xx}S_x x_i + S_x^2 x_i^2}{s_i^2 \Delta^2} \\ &= \frac{S_{xx}^2}{\Delta^2} \sum \frac{1}{s_i^2} - \frac{2S_{xx}S_x}{\Delta^2} \sum \frac{x_i}{s_i^2} + \frac{S_x^2}{\Delta^2} \sum \frac{x_i^2}{s_i^2} = \frac{S_{xx}^2 S}{\Delta^2} - \frac{2S_{xx}S_x^2}{\Delta^2} + \frac{S_{xx}S_x^2}{\Delta^2} \\ &= \frac{S_{xx}(S_{xx}S - S_x^2)}{\Delta^2} = \frac{S_{xx}}{\Delta} \end{aligned} \tag{6.9}$$

and

$$\begin{aligned} s_{m_2}^2 &= \sum s_i^2 \left[\frac{Sx_i - S_x}{s_i^2 \Delta} \right]^2 = \sum \frac{S^2 x_i^2 - 2S\, S_x x_i + S_x^2}{s_i^2 \Delta^2} \\ &= \frac{S^2}{\Delta^2} \sum \frac{x_i^2}{s_i^2} - \frac{2S\, S_x}{\Delta^2} \sum \frac{x_i}{s_i^2} + \frac{S_x^2}{\Delta^2} \sum \frac{1}{s_i^2} = \frac{S^2 S_{xx}}{\Delta^2} - \frac{2S\, S_x^2}{\Delta^2} + \frac{SS_x^2}{\Delta^2} \\ &= \frac{S(S_{xx}S - S_x^2)}{\Delta^2} = \frac{S}{\Delta}. \end{aligned} \tag{6.10}$$

Similarly, we can find the covariance s_{m_1,m_2} from

$$s_{m_1,m_2}^2 = \sum s_i^2 \left(\frac{\partial m_1}{\partial y_i} \right) \left(\frac{\partial m_2}{\partial y_i} \right) = -\frac{S_x}{\Delta}. \tag{6.11}$$

Thus, the correlation between m_1 and m_2 becomes

$$r_{m_1,m_2} = \frac{-S_x}{\sqrt{SS_{xx}}}. \tag{6.12}$$

It is therefore useful to shift the origin to $\bar{x}$ so that $r_{m_1,m_2} = 0$, leaving our estimates for slope and intercept uncorrelated.

Finally, we must check if the fit is significant. We determine critical χ^2_α for $n-2$ degrees of freedom and test if our computed χ^2 exceeds the critical limit. If it does not, then we may say the fit is *significant* at the α level of confidence.

6.1.1 Confidence interval on regression

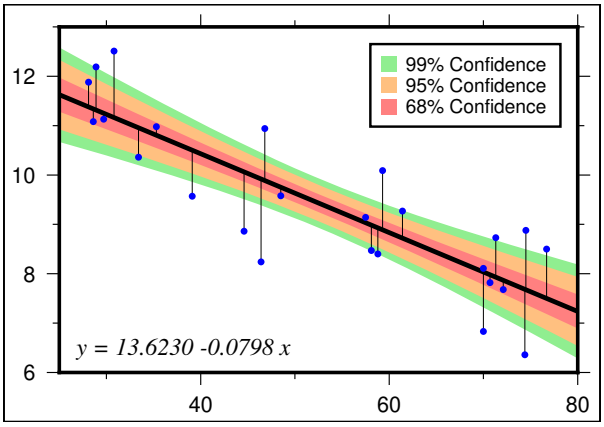


Figure 6.2: Solid line shows the least squares regression fit to the data points (blue circles), with color bands reflecting different confidence levels. Short vertical lines are the residual errors which are squared and summed in (6.16).

The formalism in the previous section allowed us to derive solutions for slope and intercept given via (6.6). Let us consider the confidence interval on the prediction. We write the least squares fit as $\hat{y} = m_1 + m_2 x$, and by substituting $m_1 = \bar{y} - m_2\bar{x}$ we obtain

$$\hat{y} = \bar{y} + m_2(x - \bar{x}). \tag{6.13}$$

Here, both the (weighted) mean y-value ($\bar{y}$) and slope (m_2) are subject to error and these in turn affect $\hat{y}$. For some chosen location x_0 the prediction for the regression would be

$$\hat{y}_0 = \bar{y} + m_2(x_0 - \bar{x}). \tag{6.14}$$

In this formulation, with a local origin at $(\bar{x},\bar{y})$, the correlation between $\bar{y}$ and m_2 is zero (e.g., 6.12) and $s^2_{m_1} = 1/S$ and $s^2_{m_2} = 1/S_{xx}$. These facts allow us to compute the expected variance V of the *mean predicted value* by summing the separate variance terms:

$$V(\hat{y}_0) = V(\bar{y}) + V(m_2)(x_0 - \bar{x})^2 = \frac{s^2}{S} + \frac{s^2(x_0 - \bar{x})^2}{S_{xx}}. \tag{6.15}$$

Here, s is our sample estimate of the (weighted) *standard deviation of the regression residuals*, $e_i = y_i - \hat{y}$, given by

$$s^2 = \frac{\sum(e_i/s_i)^2}{S(n-2)/n}, \tag{6.16}$$

where $n-2$ are the remaining degrees of freedom after computing m_1 and m_2. To obtain confidence intervals on the linear regression we simply scale s from (6.16) by a critical Student's t-value for the degrees of freedom ν and chosen confidence level (Table A.2), i.e.,

$$\hat{y}_0 \pm t(\nu,\alpha/2)s\sqrt{\frac{1}{S} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}. \tag{6.17}$$

For larger data sets the Student's t values approach the normal distribution critical values $|z_{\alpha/2}|$. Figure 6.2 illustrates how confidence bands on a least-squares regression fit takes on a parabolic shape around the best-fit line.

6.2 Orthogonal Regression

6.2.1 Major axis

It is often the case that the uncertainties in our (x,y) data affect both coordinates. Examples where this is occurs include situations where both x and y are observed quantities (and hence are known to have errors). It is also applicable when y is a function of x, but x (e.g., distance or time) itself has uncertainties. In these cases, orthogonal regression is the correct way to determine linear relationships between x and y (Figure 6.3). We will use the least squares principle and

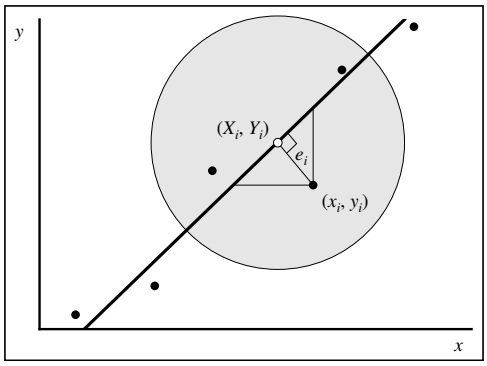


Figure 6.3: The misfit e_i is measured in the direction perpendicular to the line. Note that (x_i,y_i) denote our data points (black circles) and (X_i,Y_i) are the coordinates of their orthogonal projections (white circles; only one is shown here) onto the regression line.

minimize the sum of the squared *perpendicular* distances e_i^2 from the data points (x_i,y_i) to the regression line[1]. The function we want to minimize is

$$E = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[(X_i - x_i)^2 + (Y_i - y_i)^2 \right], \tag{6.18}$$

where lowercase (x_i,y_i) again are our observations and uppercase (X_i,Y_i) are the "adjusted" coordinates we want to find. These are, of course, required to lie on a straight line described by

$$Y_i = m_1 + m_2 X_i, \tag{6.19}$$

or equivalently

$$f_i = m_1 + m_2 X_i - Y_i = 0. \tag{6.20}$$

Thus, we cannot simply find *any* set of (X_i,Y_i) as they also have to lie on a straight line. The problem of minimizing the function (6.18) under the specified constraints (6.20) can be solved by a method known as *Lagrange's multipliers*. This method says we should form a new function F by adding the original function (6.18) and all the constraints (6.20), with each constraint scaled by an unknown (Lagrange) multiplier λ_i. Since (6.20) is actually n constraints, we find

$$F = E + \lambda_1 f_1 + \lambda_2 f_2 + \ldots + \lambda_n f_n = E + \sum_{i=1}^{n} \lambda_i f_i. \tag{6.21}$$

We may now set the partial derivatives of F to zero and solve the resulting set of equations:

$$\frac{\partial F}{\partial X_i} = \frac{\partial F}{\partial Y_i} = \frac{\partial F}{\partial m_1} = \frac{\partial F}{\partial m_2} = 0, \tag{6.22}$$

or, when viewed separately (with implied sums over $i = 1$ to n unless explicitly stated),

$$\frac{\partial F}{\partial X_i} = \sum_{j}^{n} \frac{\partial}{\partial X_i}(X_j - x_j)^2 + \sum_{j}^{n} \frac{\partial}{\partial X_i}(\lambda_j m_2 X_j) = 0 \Rightarrow 2(X_i - x_i) + m_2 \lambda_i = 0, \tag{6.23}$$

[1]This approach does not take into account the actual uncertainties in x and y (which becomes very tedious algebraically) but instead focuses on the effect of the orthogonal misfit criterion.

$$\frac{\partial F}{\partial Y_i} = \sum_j^n \frac{\partial}{\partial Y_i}(Y_j - y_j)^2 - \sum_j^n \frac{\partial}{\partial Y_i}(\lambda_j Y_j) = 0 \Rightarrow 2(Y_i - y_i) - \lambda_i = 0, \tag{6.24}$$

$$\frac{\partial F}{\partial m_1} = \sum \frac{\partial}{\partial m_1}(\lambda_i m_1) = 0 \Rightarrow \sum \lambda_i = 0, \tag{6.25}$$

$$\frac{\partial F}{\partial m_2} = \sum \frac{\partial}{\partial m_2}(\lambda_i m_2 X_i) = 0 \Rightarrow \sum \lambda_i X_i = 0. \tag{6.26}$$

Since each i represents a separate equation, we find

$$2(X_i - x_i) = -m_2\lambda_i \Rightarrow X_i = x_i - m_2\lambda_i/2, \tag{6.27}$$

$$2(Y_i - y_i) = \lambda_i \Rightarrow Y_i = y_i + \lambda_i/2. \tag{6.28}$$

Substituting these expressions for X_i and Y_i into (6.19), we obtain

$$y_i + \lambda_i/2 = m_1 + m_2\,(x_i - m_2\lambda_i/2) = m_1 + m_2 x_i - m_2^2\lambda_i/2 \tag{6.29}$$

or

$$\lambda_i = \frac{2}{1+m_2^2}\,(m_1 + m_2 x_i - y_i)\,. \tag{6.30}$$

Now, (6.25), (6.26), and (6.30) gives us $n+2$ equations in $n+2$ unknowns (all the λ_i plus m_1 and m_2). Combining (6.30) and (6.25) gives

$$\sum \frac{1}{1+m_2^2}\,(m_1 + m_2 x_i - y_i) = 0 \tag{6.31}$$

and (6.26) using (6.27) gives

$$\sum \lambda_i x_i - m_2\lambda_i^2/2 = 0, \tag{6.32}$$

into which we substitute (6.30) and find

$$\sum \frac{1}{1+m_2^2}\left(m_1 x_i + m_2 x_i^2 - y_i x_i\right) - \sum \frac{m_2}{(1+m_2^2)^2}\,(m_1 + m_2 x_i - y_i)^2 = 0. \tag{6.33}$$

These two equations (6.31 and 6.33) relate the parameters m_1 and m_2 to the given data values x_i and y_i. We find the solution by solving the equations simultaneously. Noting that the denominator in (6.31) cannot be zero, we find

$$\sum (m_1 + m_2 x_i - y_i) = 0 \Rightarrow n m_1 + m_2 \sum x_i = \sum y_i \tag{6.34}$$

or

$$m_1 = \bar{y} - m_2\bar{x}, \tag{6.35}$$

where $\bar{x}$ and $\bar{y}$ are the mean data values. This expression for the intercept can now be substituted into (6.33) so we may solve for the slope. We multiply through by $(1+m_2^2)^2$ and obtain

$$\sum (1+m_2^2)\left(\bar{y}x_i - m_2\bar{x}x_i + m_2 x_i^2 - x_i y_i\right) - m_2 \sum \left(\bar{y} - m_2\bar{x} + m_2 x_i - y_i\right)^2 = 0 \tag{6.36}$$

which simplify to

$$(1+m_2^2)\sum x_i\,(\bar{y} - y_i + m_2(x_i - \bar{x})) - m_2 \sum \left(m_2(x_i - \bar{x}) - (y_i - \bar{y})\right)^2 = 0, \tag{6.37}$$

We now introduce the residuals $u_i = x_i - \bar{x}$ and $v_i = y_i - \bar{y}$. Then

$$(1+m_2^2)\sum (u_i + \bar{x})\,(m_2 u_i - v_i) - m_2 \sum (m_2 u_i - v_i)^2 = 0,$$

$$(1+m_2^2)\sum \left(m_2 u_i^2 - u_i v_i + \bar{x} m_2 u_i - \bar{x} v_i\right) - m_2 \sum \left(m_2^2 u_i^2 - 2 m_2 u_i v_i + v_i^2\right) = 0,$$

$$\sum \left(m_2 u_i^2 + m_2^3 u_i^2 - u_i v_i - m_2^2 u_i v_i - m_2^3 u_i^2 + 2 m_2^2 u_i v_i - m_2 v_i^2\right) = 0,$$

$$\sum \left(m_2^2 u_i v_i + m_2(u_i^2 - v_i^2) - u_i v_i\right) = 0.$$

where we have used the properties $\sum u_i = \sum v_i = 0$. These steps finally give the solution for the slope as

$$m_2 = \frac{\left(\sum v_i^2 - \sum u_i^2\right) \pm \sqrt{\left(\sum u_i^2 - \sum v_i^2\right)^2 + 4\left(\sum u_i v_i\right)^2}}{2\sum u_i v_i}. \tag{6.38}$$

This equation gives two solutions for the slope and we choose the one that minimizes E. The other solution maximizes E and makes an angle of 90 degrees with the optimal solution.

6.2.2 Reduced major axis (RMA) regression

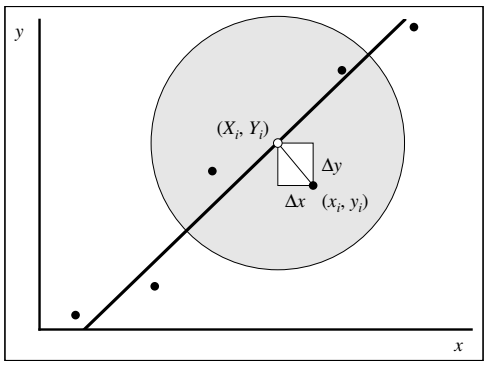


Figure 6.4: RMA regression minimizes the sum of the *areas* of the (white) rectangles defined by the data points (black circles) and their orthogonal projection points (white circles; only one is shown here) on the regression line.

In this alternative formulation of misfit we minimize the sum of the *areas* of the rectangles defined by Δx and Δy (Figure 6.4). Hence, the function to minimize is

$$E = \sum (X_i - x_i)(Y_i - y_i). \tag{6.39}$$

The constraints remain the same, i.e. $Y_i = m_1 + m_2 X_i$. The Lagrange's multiplier method leads to a system of equations similar to those discussed in the previous section (6.21), but now we find the $2n+2$ equations

$$\frac{\partial F}{\partial X_i} = \frac{\partial}{\partial X_i}\sum_j^n (X_j - x_j)(Y_j - y_j) + \frac{\partial}{\partial X_i}\sum_j^n \lambda_j m_2 X_j = 0 \Rightarrow Y_i - y_i + m_2\lambda_i = 0 \quad i = 1,n, \tag{6.40}$$

$$\frac{\partial F}{\partial Y_i} = \frac{\partial}{\partial Y_i}\sum_j^n (X_j - x_j)(Y_j - y_j) - \frac{\partial}{\partial Y_i}\sum_j^n \lambda_j Y_j = 0 \Rightarrow X_i - x_i - \lambda_i = 0 \quad i = 1,n, \tag{6.41}$$

$$\frac{\partial F}{\partial m_1} = \frac{\partial}{\partial m_1}\sum \lambda_i m_1 = 0 \Rightarrow \sum \lambda_i = 0, \tag{6.42}$$

$$\frac{\partial F}{\partial m_2} = \frac{\partial}{\partial m_2}\sum \lambda_i m_2 X_i = 0 \Rightarrow \sum \lambda_i X_i = 0. \tag{6.43}$$

Rearranging yields

$$X_i - x_i - \lambda_i = 0 \Rightarrow X_i = x_i + \lambda_i, \tag{6.44}$$

$$Y_i - y_i + m_2\lambda_i = 0 \Rightarrow Y_i = y_i - m_2\lambda_i. \tag{6.45}$$

Substituting these values into the equation for the line (i.e., $Y_i = m_1 + m_2 X_i$) gives

$$y_i - m_2\lambda_i = m_1 + m_2(x_i + \lambda_i) \tag{6.46}$$

or

$$\lambda_i = \frac{y_i - m_1 - m_2 x_i}{2m_2}. \tag{6.47}$$

Since $\sum \lambda_i = 0$, we find again

$$m_1 = \bar{y} - m_2 \bar{x}. \tag{6.48}$$

Substituting λ_i, X_i and m_1 into (6.43) gives

$$\sum \left(x_i + \frac{y_i - \bar{y} + m_2\bar{x} - m_2 x_i}{2m_2} \right) \left(\frac{y_i - \bar{y} + m_2\bar{x} - m_2 x_i}{2m_2} \right) = 0.$$

We let $u_i = x_i - \bar{x}$ and $v_i = y_i - \bar{y}$ as before and obtain

$$\sum (m_2(u_i + \bar{x}) + v_i - m_2 u_i)(v_i - m_2 u_i) = 0,$$

$$\sum (2m_2\bar{x} + m_2 u_i + v_i)(v_i - m_2 u_i) = 0,$$

$$\sum (2m_2\bar{x}v_i - 2m_2^2 \bar{x}u_i + m_2 u_i v_i - m_2^2 u_i^2 + v_i^2 - m_2 u_i v_i) = 0,$$

$$\sum v_i^2 - m_2^2 \sum u_i^2 = 0.$$

where we again have used the property that sums of u_i and v_i are zero. Finally, we obtain

$$m_2 = \pm\sqrt{\sum v_i^2 / \sum u_i^2} = \pm \frac{s_y}{s_x}, \tag{6.49}$$

i.e., the best slope equals the ratio of the y and x observations' separate standard deviations. The sign is indeterminate but is inferred from the sign of the correlation coefficient (a negative correlation means a negative slope)[2].

6.3 Robust Regression

In simple regression one assumes a relation of the type

$$y_i = m_1 + m_2 x_i + \varepsilon_i, \tag{6.50}$$

in which x_i is called the *explanatory variable* or *regressor*, and y_i is the *response variable*. Again, we seek to estimate m_1 and m_2 (intercept and slope) from the data (x_i, y_i). It is commonly assumed that the deviations ε_i are normally distributed. Fortunately, in simple regression the observations (x_i, y_i) are 2-D so they can be plotted. It is always a good idea to do that first to see if any unusual features are present and to make sure the data are roughly linear. Applying a regression estimator to the data (x_i, y_i) will result in the two regression coefficients $\hat{m}_1$ and $\hat{m}_2$. They are not the *true* parameters m_1 and m_2, but our "best" *estimates* of them. We can insert those into (6.50) and find the predicted estimate as

$$\hat{y}_i = \hat{m}_1 + \hat{m}_2 x_i, \quad i = 1, n. \tag{6.51}$$

The residual is then the difference between the observed and estimated values, yielding

$$e_i = y_i - \hat{y}_i, \quad i = 1, n. \tag{6.52}$$

Note that there is a difference between e_i (the misfit) and ε_i (the deviation), because $\varepsilon_i = y_i - m_1 - m_2 x_i$ are evaluated with the true unknown m_1, m_2. We can compute e_i, but not ε_i.

The most popular regression estimator dates back to the early 1800's (to our old friends Gauss and Legendre) and is called the "least-squares" (LS) method since it seeks to minimize

$$E = \sum_{i=1}^{n} e_i^2. \tag{6.53}$$

The rationale was to make the residuals very small. Gauss preferred the least-squares criterion to other objective functions because in this way the regression coefficients could be computed explicitly from the data (no computers back in

[2] We cheated a bit as the individual terms in 6.39 might be negative; however, a more advanced derivation still yields the same solution.

the day, at least not mechanical or electronic ones). Later, Gauss introduced the normal, or Gaussian, distribution for which least-squares is optimal.

More recently, people have realized that real data often do not satisfy the Gaussian assumption and this "failure to comply" may have a dramatic effect on the LS results. In Figure 6.5 we have five points that lie almost on a straight line. Here, the *LS* line fits the data very well. However, let us see what happens if we get a wrong value for y_4 because of a recording or copying error:

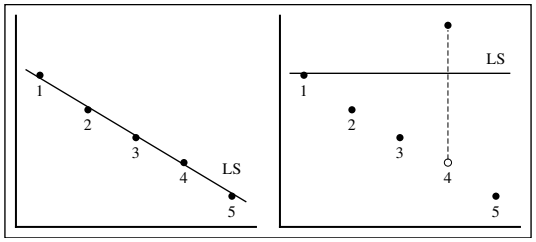


Figure 6.5: Pitfalls of least-squares regression, part I. An outlying point in the y-direction will effect the regression line considerably.

The bad point y_4 is called an *outlier in the y-direction*, and it has a dramatic effect on the *LS* line which now is tilted away from the trend of the remaining data. Such outliers have received the most attention because most experiments are set up to expect errors in y only. However, in observational studies it often happens that outliers occur in the x_i data as well.

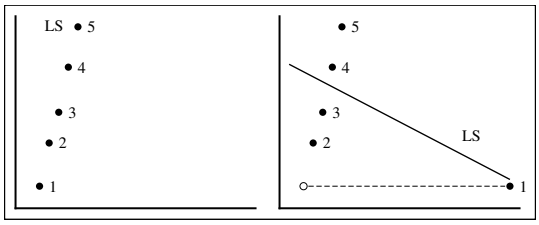


Figure 6.6: Pitfalls of least-squares regression, part II. Here, an outlying point in the x-direction can have a huge effect on the regression line.

Figure 6.6 illustrates the effect of an outlier in the x-direction. It has an even more dramatic effect on *LS* since it now is almost perpendicular to the actual trend. Because this single point has such a large influence we denote it as a *leverage* point. This is because the residual e_i (measured in the y-direction) is enormous with regard to the original *LS* fit. The second *LS* fit reduces this enormous error at the expense of increasing the errors at all other points. In general, we call the k'th point a leverage point if x_k lies far from the bulk of the x_i. Note that this definition does not take y_i into account. For instance, Figure 6.7 shows a "good" leverage point since it lies on the linear trend set by the majority of the data. Thus, a leverage point only refers to its *potential* for badly influencing the coefficients $\hat{m_1}, \hat{m_2}$. When a point (x_i, y_i) deviates from the linear relation of the majority it is called a *regression outlier*, taking into account both x_i and y_i simultaneously. As such, it is a vertical outlier or a bad leverage point.

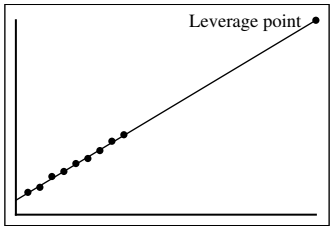


Figure 6.7: The effect of leverage points in regression can be enormous, whether the data point is a valid observation or a bad outlier.

It is often stated that regression outliers can be detected by looking at the *LS* residuals. This is not always true. The bad leverage point 1 has tilted the line so much that its residual is very small. Consequently, a rule that says "delete points with highest *LS* residual" would first find points number 2 and 5, thereby deleting the good points first. Of course, in simple $x-y$ regression we have the benefit of being able to plot the data so this is not often a problem, except when the number of data sets and points are large.

From the simple examples we have just presented, we find that the breakdown point for *LS* regression is merely $1/n$ since one point is enough to ruin the day — analogous to the breakdown point for the mean, which was also based on *LS*.

A first step toward a more robust regression was taken more than 100 years ago when Edgeworth suggested that one could instead minimize

$$E = \sum_{i=1}^{n} |e_i|, \tag{6.54}$$

which we will call L_1 regression. Unfortunately, while L_1 regression is robust with respect to outliers in y, it offers no protection toward bad leverage points. Thus, the breakdown point is still only $1/n$.

While there are many methods that offer a higher breakdown point than L_1 and L_2, we will concentrate our presentation on one particular method. Again, let us look at the *LS* formulation:

$$\underset{\hat{m_1},\hat{m_2}}{\text{minimize}} \quad E = \sum_{i=1}^{n} e_i^2.$$

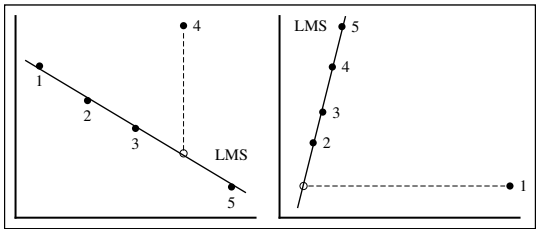


Figure 6.8: Robust regression, such as LMS, is very tolerant of outlying points in both the x and y directions.

At first glance, you would think that a better name for *LS* would be least *sum* of squares. Apparently, few people objected to removing the word "sum" as if the only sensible thing to do with n numbers would be to add them. Adding the e_i^2 terms is the same as using their mean (dividing by n does not affect the minimization). Why not replace the mean (i.e., the sum) by a median, which we know is very robust? This yields the "least median of squares" (LMS) criterion:

$$\underset{\hat{m_1},\hat{m_2}}{\text{minimize}} \quad \text{median } e_i^2. \tag{6.55}$$

It turns out the LMS fit is very robust with respect to outliers in y as well as in x. Its breakdown point is 50%, which is the most we can ask for. If more than half your data are bad then you cannot tell which part is good unless you have additional information. Figure 6.8 shows what we get using this method on the data that made the LS technique fail so miserably.

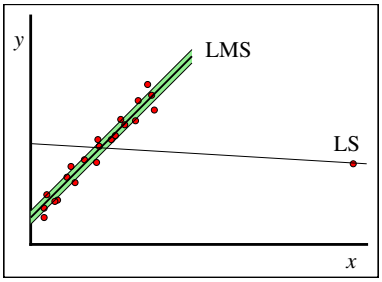


Figure 6.9: Geometrical meaning of the LMS regression: the narrowest strip that covers half the data points.

The LMS line also has an intuitive geometric interpretation: it lies at the center of the narrowest strip that covers half of the data points. By half of the points we mean $n/2+1$, and the thickness of the strip is measured in the vertical direction (i.e., Figure 6.9).

An example of LMS regression comes from astronomy. Astronomers often look for a linear relationship between the logarithm of the light intensity and the logarithm of the surface temperature of stars. A scatter plot of observed quantities may look like Figure 6.10. Here, the LMS line defines what is known as the *main sequence*; the four outlying stars turned out to be red giants that do not follow the general trend. The *LS* fit produces a rather worthless compromise solution. Here, the outliers are not so much errors as *contamination from a different population*.

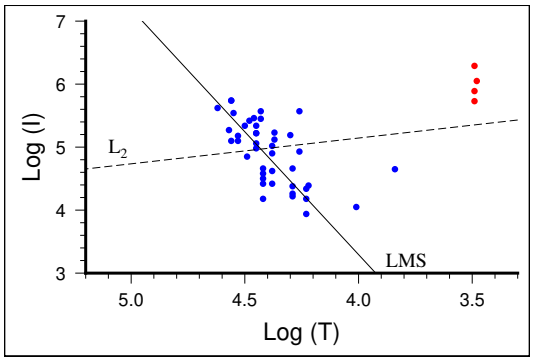


Figure 6.10: Example of robust regression in astronomy. We see a Hertzsprung-Russell diagram of the star cluster CYG OB1 with the least squares (dashed line) and LMS (solid line) fit. The red giants are distorting the LS fit. Data taken from Rousseeuw, P. J., and A. M. Leroy (1987), *Robust regression and outlier detection*, 329 pp., John Wiley and Sons, New York.

6.3.1 How to estimate LMS regression

We can rewrite the minimization criterion as follows:

$$\underset{\hat{m_1},\hat{m_2}}{\text{minimize}}\{\text{median } e_i^2\} \quad = \quad \underset{\hat{m_1},\hat{m_2}}{\text{minimize}}\{\text{median } \quad ((y_i-\hat{m_2}x_i)-\hat{m_1})^2\} \tag{6.56}$$

in the form

$$\underset{\hat{m_2}}{\text{minimize}} \left\{ \underset{\hat{m_1}}{\text{minimize median}} \quad ((y_i-\hat{m_2}x_i)-\hat{m_1})^2 \right\} . \tag{6.57}$$

We will treat the two minimizations here separately. The innermost minimization is the easy part, because for any given $\hat{m_2}$ it becomes essentially a 1-D problem, i.e., we want to find the value for $\hat{m_1}$ that minimizes the median

$$\underset{\hat{m_1}}{\text{minimize}} \left\{ \text{median } (u_i-\hat{m_1})^2 \right\}, \tag{6.58}$$

where u_i is calculated as $u_i = y_i - \hat{m_2}x_i$ (remember, we assumed that $\hat{m_2}$ was given). This minimization problem is the same one we found earlier to give a good estimate of the mode. Thus, this operation finds the mode of the u_i data set. We therefore need to find the $\hat{m_2}$ for which

$$E(\hat{m_2}) = \text{median } [(y_i-\hat{m_2}x_i)-\hat{m_1}]^2 \tag{6.59}$$

is minimal. This is simply the minimization of a 1-D function $E(\hat{m_2})$ which is continuous but not everywhere differentiable.

To find this minimum we make the observation that the slope in the $x-y$ plane must be in the $\pm 90^\circ$ range (when expressed as an angle β, with $m_2 = \tan\beta$). We then simply perform a search for the optimal angle. Starting with $\beta = -90^\circ$, we form the resulting u_i and solve the 1-D minimization problem for $\hat{m_1}$, i.e., finding the LMS mode

estimate $\hat{m_1}$. We now increment the angle β by $d\beta$ to, say, -89°, and repeat the process. At each step we keep track of what $E(\beta)$ is, and repeat these steps for all angles through $\beta = 90^\circ$.

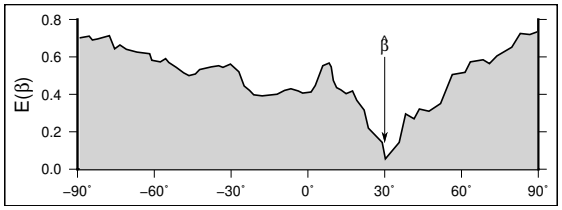


Figure 6.11: Determining the best regression slope, $\hat{m_2} = \tan\hat{\beta}$. The misfit function is not smooth but usually has a minimum for the optimal slope. It is also likely to reveal several local minima that could trick a simpler search.

Having found the slope $\hat{m_2}$ that gave the smallest misfit we may improve on this estimate by using a smaller step size $d\beta$ in this region to pinpoint the best choice for $\hat{m_2}$. A plot of $E(\beta)$, shown in Figure 6.11, is very useful since it may tell us how unique the LMS regressions are: if more than one minimum are found they may indicate a possible ambiguity as there may be two or more lines that fit the data equally well.

We will elaborate on the breakdown point for simple regression and illustrate it with a simple experiment. Consider a data set that contains 100 good data points that exhibit a strong linear relation, computed from

$$y_i = 1.0x_i + 2.0 + \varepsilon_i \quad 1 \leq x_i \leq 4, \tag{6.60}$$

with ε_i normally distributed, with $\mu = 0$ and $\sigma = 0.2$.

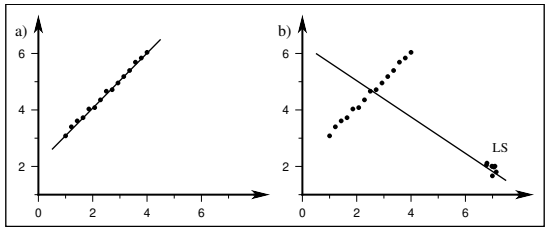


Figure 6.12: (a) Best LS fit to synthetic data set computed from a linear model with Gaussian noise. (b) Synthetic data set computed from a linear model with Gaussian noise, but now contaminated by points from another (bivariate) distribution centered on (7,2). The LS line is pulled way off by these bad leverage points. We can then plot the slope value as a function of the percentage of contamination.

Any regression technique, including L_2, will of course recover estimates of the slope and intercept that are very close to the true values 1 and 2 (Figure 6.12a). Then, we will start to contaminate the data by replacing "good" points with bad ones, the latter coming from a bivariate normal distribution with $\mu = (7,2)$ and $\sigma_r = 0.5$. We systematically substitute one bad point for a good point and recompute the regression parameters after each step. What we find is that the LS estimate goes bad right away (Figure 6.12b). The bad points are basically bad leverage points, which we know the LS process cannot handle. We keep track of the slope estimate after each substitution and graph the results in Figure 6.13.

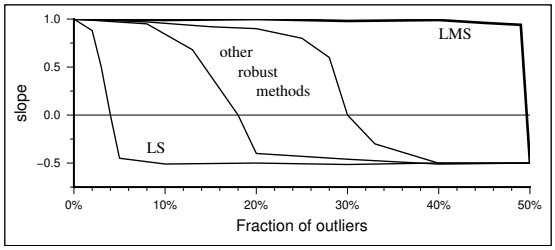


Figure 6.13: Schematic breakdown plot for several regressors. The LMS only breaks down when 50% of the data are outlying. LS breaks down immediately because every point matters.

We call the percentage at which the slope starts to deviate significantly from the true value the *breakdown point*. The clear winner of this test is the LMS regression which keeps finding the correct trend until half the points have been replaced. We should add that while LMS always have a breakdown point of 50%, it is found that the effect of the outliers often depends on the quality of the good data. In cases where the good data exhibit a strong correlation the outliers do less damage than in a case where there is little or no correlation (Figure 6.14). Of course, when the correlation among the good data is minimal it is probably not very useful to insist that the data really exhibit a linear trend in the first place.

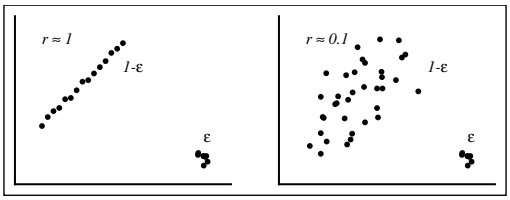


Figure 6.14: Two data sets with the same degree of contamination. However, one exhibit a much stronger correlation between the "good" points than the other.

6.3.2 How to find LMS 1-D Location (single mode)

When we discussed estimates of central location we briefly mentioned that the value $\hat{x}$ that minimized

$$\underset{\hat{x}}{\text{minimize}} \quad \left\{\text{median } (x_i - \hat{x})^2\right\} \tag{6.61}$$

was called the LMS location and that it was a good approximation to the *mode*, but how do we determine the LMS estimate? It turns out that it is rather simple. The following recipe will do:

1. Sort the data into ascending order.
2. Determine the shortest half of the sample, i.e., find the value for i that yields the smallest of the differences $(x_{h+i} - x_i)$ with $h = n/2 + 1$.
3. The LMS estimate is the midpoint of the shortest half:

$$\text{LMS } = \frac{1}{2}(x_{h+i} + x_i). \tag{6.62}$$

You can empirically try this out by letting $\hat{x}$ take on all values within the data range, compute the median of all $(x_i - \hat{x})^2$ values, and graph the curve and find its minimum (Figure 6.15).

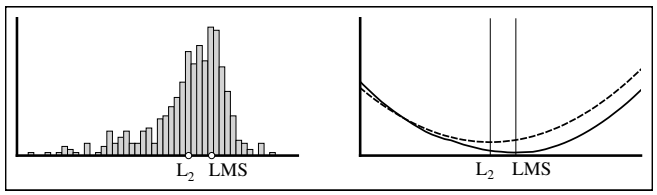


Figure 6.15: LMS defines the mode while L2 defines the mean; both are the locations where their respective objective functions are minimized.

6.3.3 Making LMS "analytical" — finding outliers

There is one problem with using the robust LMS parameters: the method is not *analytical* so it does not lend itself easily to standardized statistical tests. We will look into how we can overcome this obstacle.

The main problem comes from the fact that the outliers cause L_2 estimates to become unreliable. We will avoid this problem altogether by using the best from both worlds (L_2 *and* LMS): We will use robust LMS techniques to find the best parameters in the regression model and then use the robust residuals to detect outliers. Finally, we recompute weighted L_2 parameters with outliers given zero weights and other points given unit weights. These L_2 estimates now represent only the "good" portion of the data and confidence limits and statistical tests may be based on the behavior of these good values. We call this technique "Re-weighted Least Squares" (RLS).

First, we need a robust scale estimate for the residuals that will make them nondimensional. It is customary to choose the preliminary scale estimate

$$s^0 = 1.4826\left(1+\frac{5}{n-2}\right)\sqrt{\text{median } e_i^2}, \tag{6.63}$$

where $e_i = y_i - \hat{y}_i$ are the residuals (i.e., misfits) at each point. With this scale estimate we can evaluate the normalized residuals as

$$z_i = e_i/s^0. \tag{6.64}$$

Now use these numbers to design the weights:

$$w_i = \begin{cases} 1, & |z_i| \le 2.5 \\ 0, & \text{otherwise} \end{cases}. \tag{6.65}$$

The final LMS regression scale estimate is then given by

$$s^* = \sqrt{\sum w_i e_i^2/(\sum w_i - 2)}. \tag{6.66}$$

The RLS regression parameters are therefore obtained by minimizing the weighted, squared residuals:

$$\min E = \sum_{i=1}^{n} w_i e_i^2. \tag{6.67}$$

This is simply the L_2 solution when only the good data are used. As shown when we discussed the weighted L_2 regression problem, this technique will provide confidence intervals on both the slope and intercept and it also allows us to use the χ^2-test to check whether the RLS fit is significant or not.

The strength of the linear relationship can again be measured by the (Pearson) correlation coefficient. The LMS estimate of correlation is now given by

$$r = \sqrt{1-\left(\frac{\text{median } |e_i|}{\text{MAD } y_i}\right)^2} \tag{6.68}$$

with

$$\text{MAD } y_i = \text{median } |y_i - \tilde{y}|. \tag{6.69}$$

Compare this to the L_2 case, where

$$r = \frac{s_{xy}}{s_x s_y} = \sqrt{1-\frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y})^2}} = \sqrt{1-\left(\frac{\bar{e}}{s_y}\right)^2}. \tag{6.70}$$

This comparison shows that the robust estimates for "average" and "scale" have replaced the traditional least-squares estimates in the expression for correlation.

6.4 Multiple Regression

6.4.1 Preliminaries

Multiple regression is an extension of the simple regression problem for including more than one explanatory variable. The most straightforward extension from 1-D to 2-D results in the determination of the constants m_1, m_2, and m_3 in

$$d = m_1 + m_2 x + m_3 y, \tag{6.71}$$

which you will recognize as the equation for a *plane*. A very common application of finding the best-fitting plane is to define and remove a regional (planar) trend from 2-D data sets so that local variations can be inspected and compared. Let us say we have n data points $d_i(x_i, y_i)$ in the plane and we want to determine the regional trend using standard L_2 techniques. Using (6.71) gives

$$m_1 + m_2 x_i + m_3 y_i = d_i, \quad i = 1, n \tag{6.72}$$

or $\mathbf{G} \cdot \mathbf{m} = \mathbf{d}$. We know how to solve these L_2 problems now and quickly state that the solution is given by the linear system:

$$\mathbf{G}^T \mathbf{G}\mathbf{m} = \mathbf{G}^T \mathbf{d} \tag{6.73}$$

or

$$\mathbf{N} \cdot \mathbf{m} = \mathbf{v}. \tag{6.74}$$

We can find the $\mathbf{N}$ and $\mathbf{v}$ by performing the matrix operations (below, all sums are implicitly from $i = 1$ to n)

$$\begin{bmatrix} n & \sum x_i & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum y_i x_i \\ \sum y_i & \sum x_i y_i & \sum y_i^2 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} \sum d_i \\ \sum x_i d_i \\ \sum y_i d_i \end{bmatrix}. \tag{6.75}$$

To verify this system, note that

$$\mathbf{N} = \mathbf{G}^T \mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ y_1 & y_2 & y_3 & \cdots & y_n \end{bmatrix} \cdot \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix} = \begin{bmatrix} n & \sum x_i & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum x_i y_i \\ \sum y_i & \sum x_i y_i & \sum y_i^2 \end{bmatrix} \tag{6.76}$$

and

$$\mathbf{v} = \mathbf{G}^T \mathbf{d} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ y_1 & y_2 & y_3 & \cdots & y_n \end{bmatrix} \cdot \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = \begin{bmatrix} \sum d_i \\ \sum x_i d_i \\ \sum y_i d_i \end{bmatrix}. \tag{6.77}$$

In many situations we have a gridded data set where we have observations on an equidistant lattice. We can then find the $\mathbf{N}$ matrix analytically since the geometry is so simple.

Example 6–1. We will examine data set given in Table 6.1.

X/Y	-2	-1	0	1	2
2	-1	0	3	2	4
1	1	-1	2	2	3
0	0	0	1	2	2
-1	-2	0	1	1	2
-2	-1	-1	0	-1	1

Table 6.1: Simple data set of measured values of $d(x, y)$ on an equidistant grid.

In this case, we find x_i and y_i to be centered on 0. This means both the terms $\sum x_i = \sum y_i = 0$ by definition. Furthermore, since $\sum x_i y_i$ can be written

$$\sum_{i=1}^{n} x_i y_i = \sum_{x=-2}^{x=2} \sum_{y=-2}^{y=2} xy = \sum_{x=-2}^{2} x \sum_{y=-2}^{y=2} y = 0 \tag{6.78}$$

we are left with

$$\begin{bmatrix} 25 & 0 & 0 \\ 0 & 50 & 0 \\ 0 & 0 & 50 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} 20 \\ 38 \\ 25 \end{bmatrix}, \tag{6.79}$$

or directly

$$d = \frac{4}{5} + \frac{19}{25}x + \frac{1}{2}y. \tag{6.80}$$

The data and the best-fitting plane are shown in Figure 6.16.

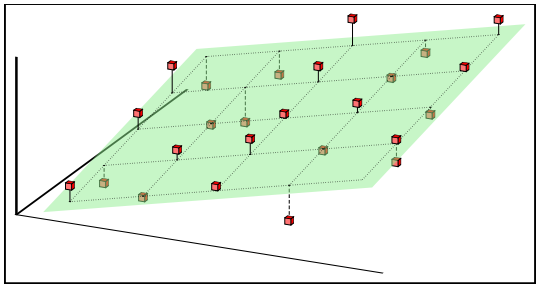


Figure 6.16: Data points (red cubes) used in Example 6.1, with the least-squares solution shown as a transparent green plane. Small black dots indicate prediction from the solution in (6.80), while vertical lines connect data points with model predictions.

It should surprise no one that in the general case (i.e., no organized grid structure) it is advantageous to translate the data to a new origin $\bar{x}, \bar{y}$ and operate on the adjusted coordinates

$$u_i = x_i - \bar{x}, v_i = y_i - \bar{y}. \tag{6.81}$$

The system to solve is then

$$\begin{bmatrix} n & 0 & 0 \\ 0 & \sum u_i^2 & \sum u_i v_i \\ 0 & \sum u_i v_i & \sum v_i^2 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} \sum d_i \\ \sum u_i d_i \\ \sum v_i d_i \end{bmatrix}. \tag{6.82}$$

Since the first equation directly gives us $m_1 = \bar{d}$, we are left with the simple 2×2 system

$$\begin{bmatrix} \sum u_i^2 & \sum u_i v_i \\ \sum u_i v_i & \sum v_i^2 \end{bmatrix} \begin{bmatrix} m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} \sum u_i d_i \\ \sum v_i d_i \end{bmatrix}. \tag{6.83}$$

Going up one more dimension means we want to find a hyper plane in a 4-D coordinate system. That situation become difficult to envision but easy to construct. As an example, consider measurements of temperature of points described by (x, y, z) triplets. One can easily find the equation of the hyper-plane

$$T = m_1 + m_2 x + m_3 y + m_4 z \tag{6.84}$$

that best fits the data in a least square sense.

6.4.2 Multiple regression

In general, we will be considering a linear regression with $k+1$ independent variables, and our regression model will now be of the form

$$m_0 + m_1 x_1 + m_2 x_2 + \cdots + m_k x_k + \varepsilon = d. \tag{6.85}$$

Our model states that the observations d_i can be explained by a constant intercept m_0 plus a linear combination of k variables, with each variable having its own "slope". Thus, m_4 describes how fast d changes when x_4 changes, holding

all other x_i fixed. As usual, we explain the misfit as being due to deviations ε_i, which we hope are normally distributed with zero mean.

In many situations it is obvious which values are the observations and which are the variables. In our example with temperatures in space we wanted to find how T varies as a function of position. However, in other cases it may be less clear-cut. E.g., if we measure (depth, porosity, permeability, water content) in a core, we will probably let depth be a variable but which of the others do we pick as our observation d? In the general case this becomes arbitrary. It also becomes arbitrary to measure the misfit in the d-direction only. The extension of orthogonal regression to multiple regression is then a solution. In that case, we would want to minimize the shortest distance between the data points and the hyper plane. To prevent unseemly bleeding from the ears, we will not attempt this approach here.

Once we have selected our d-values (and used the symbol x_{ij} to represent the i'th observation of the j'th variable), we set up the problem in matrix form as

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_k \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}, \tag{6.86}$$

or $\mathbf{G} \cdot \mathbf{m} = \mathbf{d}$. This is nothing more than our general least squares problem, whose solution is known to be

$$\mathbf{m} = [\mathbf{G}^T \mathbf{G}]^{-1} \mathbf{G}^T \mathbf{d}. \tag{6.87}$$

However, in multiple regression we want to determine the *relative importance* of the independent variables as predictors of the dependent variable d. The values of the regression coefficients tells us little, since they depend on the units chosen. Also, if the $\bar{x}_j$'s are very different in magnitude we may lose precision due to round-off errors. Consequently, we choose to transform our data into normal scores via

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \text{ and } d'_i = \frac{d_i - \bar{d}}{s_d}, \tag{6.88}$$

where

$$s_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2} \text{ and } s_d = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2} \tag{6.89}$$

are the sample standard deviations of the j'th variable and d, respectively. To populate the normal equation matrix we must form $\mathbf{G}^T \mathbf{G}$ and carry out the summations. You will remember that the form of the matrix for normal equations is

$$\begin{bmatrix} n & \sum x_1 & \sum x_2 & \cdots & \sum x_k \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 & \cdots & \sum x_1 x_k \\ \sum x_2 & \sum x_2 x_1 & \sum x_2^2 & \cdots & \sum x_2 x_k \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sum x_k & \sum x_k x_1 & \sum x_k x_2 & \cdots & \sum x_k^2 \end{bmatrix} \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_k \end{bmatrix} = \begin{bmatrix} \sum d \\ \sum x_1 d \\ \sum x_2 d \\ \vdots \\ \sum x_k d \end{bmatrix}, \tag{6.90}$$

with the sums implied to include all the data points (i.e., over $i = 1, n$). The effect of normalizing is two-fold:

1. Because the linear sums over x_i and d_i now are zero, $m'_0 = 0$, and we must find only k coefficients m'_1, ..., m'_k instead of $k+1$, as originally intended. Consequently, the first row and first column of the system (6.90) are removed.

2. The rest of the matrix becomes proportional to the correlation matrix, $\mathbf{C}$.

To verify, we examine the terms in the second equation in the normal system, which are

$$\sum \frac{(x_{i1} - \bar{x}_1)^2}{s_1^2} \quad \sum \frac{(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)}{s_1 s_2} \quad \cdots \quad \sum \frac{(x_{i1} - \bar{x}_1)(d_i - \bar{d})}{s_1 s_d} \tag{6.91}$$

and we note they are simply

$$\frac{(n-1)s_1^2}{s_1^2} \quad \frac{(n-1)s_{12}}{s_1 s_2} \quad \cdots \quad \frac{(n-1)s_{1d}}{s_1 s_d}. \tag{6.92}$$

Since the constant $(n-1)$ appears in all terms we can delete it and find

$$\begin{bmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1k} \\ r_{21} & 1 & r_{23} & \cdots & r_{2k} \\ r_{31} & r_{32} & 1 & \cdots & r_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & r_{k3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} m_1' \\ m_2' \\ \vdots \\ m_k' \end{bmatrix} = \begin{bmatrix} r_{1d} \\ r_{2d} \\ r_{3d} \\ \vdots \\ r_{kd} \end{bmatrix} \tag{6.93}$$

or $\mathbf{C} \cdot \mathbf{m}' = \mathbf{r}$. Solving $\mathbf{m}' = \mathbf{C}^{-1}\mathbf{r}$ we recover the unscaled regression parameters

$$m_j = m_j' \frac{s_d}{s_j}, \quad j = 1,k, \tag{6.94}$$

and we can reconstruct the missing intercept via

$$m_0 = \bar{d} - m_1\bar{x}_1 - m_2\bar{x}_2 - \cdots - m_k\bar{x}_k. \tag{6.95}$$

As an indicator of the goodness-of-fit we use the *coefficient of multiple regression*,

$$R^2 = SS_R/SS_T = \sum_{i=1}^{n}(\hat{d}_i - \bar{d})^2 / \sum_{i=1}^{n}(d_i - \bar{d})^2, \tag{6.96}$$

where $\hat{d}_i = d(\mathbf{x}_i)$ is the predicted value. For simple 2-D regression,

$$R^2 = r^2 = \frac{s_{xd}}{s_x^2 s_d^2}. \tag{6.97}$$

While a simple inspection of the m_j' can tell you which parameters seem to be most important, it is generally not enough to determine the best regression equation. Obviously, if any of the m_j' are close to or equal to zero we can say that the corresponding variable x_j is largely *irrelevant* to the regression and remove it from the system. At other times, some of the variables may be linearly related to each other, leading to *redundant* variables that should be removed. One approach we will consider is to evaluate all possible regressions using all combinations of x_j's and determine how many variables are necessary. We can use an ANOVA test to resolve whether a coefficient is significant or not. Suppose we have carried out a multiple regression for p independent variables. Next, we add q additional variables and repeat the regression. We would like to test whether these new variables are significant. We construct the ANOVA table given as Table 6.2.

Source of Variation	**Degrees of Freedom**	**Sums of Squares**	**Mean Square**	F
1st regression	p	SS_{R1}		
2nd regression	$p+q$	SS_{R2}		
Difference between 1st and 2nd regression	q	$\Delta SS_R = SS_{R2} - SS_{R1}$	$MSR = \Delta SS_R/q$	$\frac{MSR}{MSE}$
Residuals	$n-p-q-1$	SSE	$MSE = SSE/(n-p-q-1)$	
Total	$n-1$	SS_T		

Table 6.2: ANOVA table used to determine whether or not the addition of q extra parameters results in a significant reduction in misfit (i.e., a significant improvement in variance explained).

The observed F value is

$$F = \frac{\Delta SS_R/q}{SSE/(n-p-q-1)}. \tag{6.98}$$

With explained SS in % defined as $ESS = 100R^2$ we find

$$F = \frac{(ESS_{p+q} - ESS_p)/q}{(100 - ESS_{p+q})/(n-p-q-1)}. \tag{6.99}$$

Comparing the observed F statistic to a table of critical values (see Appendix A.4) will tell us if the new q parameters produce a significant improvement or not.

Example 6–2. We will use multiple regression to look for a linear relation between the occurrences of gold and various

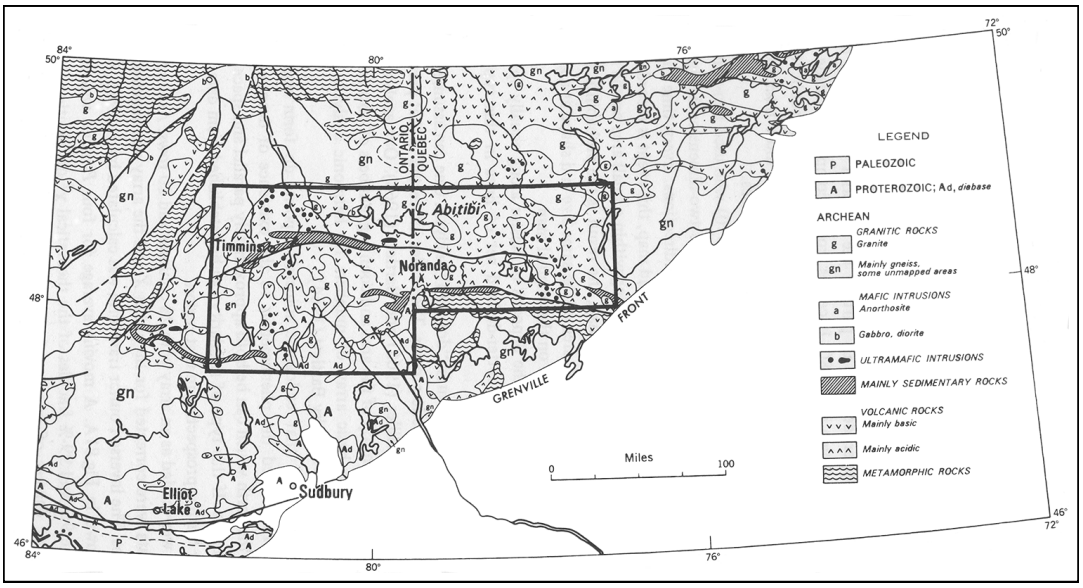


Figure 6.17: Geological map of part of Superior province, Canadian shield (After G.S.C.-Map no 1250A by R.J.W. Douglass). Area of interest is indicated by the heavy outline.

lithological units in the area west of Quebec (Agterberg, 1974). By overlying a grid on the geological map, we can estimate the density of gold occurrence per grid cell and find the area percentage of the rock units in the same cell. The gold densities are our "observed" data $d_i, i = 1,n$ and the rock types our "variables" $x_j, j = 1,k$. We chose six types of rocks ($k = 6$) for the mapping, as listed in Table 6.3.

Variable	Lithology
x_1	Granitic rocks, gneisses, acidic intrusive rocks
x_2	Mafic intrusive rocks (gabbros and diorites)
x_3	Ultramafics
x_4	Early Precambrian sedimentary rocks
x_5	Acidic volcanics (rhyolites and pyroclastics)
x_6	Mafic volcanics (basalts and andesites)

Table 6.3: List of variable names and the corresponding rock types used in the regression for gold densities.

The counting resulted in $n = 113$ cells, so $\mathbf{G}$ has dimensions $113 \cdot 7$ (i.e., six factors plus the intercept). However, because of the constraint that

$$\sum_{j=1}^{6} x_{ij} = 100 \tag{6.100}$$

for all blocks (required to equal a total of 100%), the rock type percentages are linearly dependent and we will not be able to invert $\mathbf{G}^T\mathbf{G}$. We will instead obtain all multiple regression solutions where we regress d on x_1, d on x_2, etc., for all six variables. Next, we try all possible combinations of two variables, then combinations of three, four, and finally

five (using all six is not possible since they are dependent.) We find that we need to perform N_r individual regression analyses, given by

$$N_r = \left(\begin{array}{c} 6 \\ 1 \end{array} \right) + \left(\begin{array}{c} 6 \\ 2 \end{array} \right) + \left(\begin{array}{c} 6 \\ 3 \end{array} \right) + \left(\begin{array}{c} 6 \\ 4 \end{array} \right) + \left(\begin{array}{c} 6 \\ 5 \end{array} \right) = \sum_{j=1}^{5} \left(\begin{array}{c} 6 \\ j \end{array} \right) = 62. \tag{6.101}$$

This number N_r goes up rapidly with the number of coefficients. We make a table of the results where we tabulate the parameter numbers and the associated ESS (Table 6.4).

x_j	**ESS** (%)	x_j	**ESS** (%)	x_j	**ESS** (%)	x_j	**ESS** (%)
1	8.64	3,5	16.29	2,3,5	16.43	1,3,4,6	25.61
2	1.45	3,6	0.33	2,3,6	1.79	1,3,5,6	28.37
3	0.19	**4,5**	**28.15**	2,4,5	28.18	1,4,5,6	28.61
4	10.48	4,6	10.56	2,4,6	12.46	**2,3,4,5**	**29.40**
5	**15.56**	5,6	16.07	2,5,6	16.14	2,3,4,6	12.78
6	0.17	1,2,3	8.95	**3,4,5**	**29.32**	2,3,5,6	16.84
1,2	8.85	1,2,4	14.33	3,4,6	10.96	2,4,5,6	28.18
1,3	8,76	1,2,5	19.21	3,5,6	16.72	3,4,5,6	29.33
1,4	13.54	1,2,6	22.60	4,5,6	28.15	**1,2,3,4,5**	**29.41**
1,5	18.86	1,3,4	13.79	1,2,3,4	14.56	1,2,3,4,6	29.41
1,6	22.52	1,3,5	14.40	1,2,3,5	19.83	1,2,3,5,6	29.41
2,3	1.58	1,3,6	22.58	1,2,3,6	22.65	1,2,4,5,6	29.41
2,4	12.42	1,4,5	28.19	1,2,4,5	28.24	1,2,4,5,6	29.41
2,5	15.65	1,4,6	24.87	1,2,4,6	27.45	2,3,4,5,6	29.41
2,6	1.68	1,5,6	28.34	1,2,5,6	29.37		
3,4	10.85	2,3,4	12.72	1,3,4,5	29.32		

Table 6.4: Percentage explained sum of squares (*ESS*) for 62 possible regressions, with density of gold occurrences regressed on up to five lithological variables. Bold entries are the best combinations (from *Agterberg*, 1974).

Examining all the regression results we will focus on each single combination that gave the highest ESS for that number of parameters (bold entries in Table 6.4); these are isolated in the separate Table 6.5.

x_j combination	**ESS** (%)
5	15.56
4, 5	28.15
3, 4, 5	29.32
2, 3, 4, 5	29.40
1, 2, 3, 4, 5	29.41

Table 6.5: Each row shows the best parameter combination of 1–5 variables and the corresponding percentage of explained variation (ESS).

We are now in the position to test the solutions for significance. First, we will take a look at the best one-term (x_5) solution. With $p = 0, q = 1, n = 113$, we find

$$F = \frac{(15.56 - 0)/1}{(100 - 15.56)/(113 - 1 - 1)} = 20.46. \tag{6.102}$$

From the table, $F_{0.95,1,111} = 3.93$, so it is clear that x_5 (acidic volcanics) is a significant component. Will the addition of x_4 lead to a significant improvement? Now, $p = 1, q = 1$, so

$$F = \frac{(28.15 - 15.56)/1}{(100 - 28.15)/(113 - 1 - 1 - 1)} = 19.72. \tag{6.103}$$

This also far exceeds $F_{0.95,1,110} = 3.93$. Going one step further, we check if x_3 should be included. Here, $p = 2$ and $q = 1$, so

$$F = \frac{(29.32 - 28.15)/1}{(100 - 29.32)/(113 - 2 - 1 - 1)} = 1.80. \tag{6.104}$$

The critical $F_{0.95,1,109}$ is still 3.93, so we must conclude that x_3 is an irrelevant variable, and obviously there is no need to consider any other of the remaining x_j's that gave poorer fits still. The result of this analysis is that the density of gold occurrences may be explained by a linear model that depends on the presence of acidic volcanics and Precambrian sediments in the area. However, the correlation $(r = 0.53)$ is not all that impressive $(R^2 = 28.15\%)$, suggesting that rock type alone is but one possible indicator of prospective regions. Applying this technical concept to a unexplored area may produce some disappointment.

6.5 Problems for Chapter 6

Problem 6.1. The table *hawaii.txt* contains triplets of (distance, age, Δage) values obtained along the Hawaiian seamount chain. Distance is small-circle distances from Kilauea (in km), age is radiometric seamount/island dates (in My) with Δage the associated one-sigma uncertainties (in My). You may need the `regress_ls.m` function for this problem.

a) Plot the data points with error bars (hint: see help `errorbar` if using MATLAB).

b) Find the simple least squares regression line, state the parameter values, and overlay the predictions on your plot.

c) Find the weighted least squares regression line, state the parameter values, and overlay the predictions on your plot.

d) What is the geological (or physical) meaning of the slopes and intercepts? Comment on the values.

Problem 6.2. During a calibration, nine pairs of readings were obtained with a UV spectro-photometer used to measure chemical oxygen demand in waste water (in mg/l; see file *COD.txt*) versus UV absorbance. Remove the mean from each coordinate set and determine the least-squares regression fit. Plot the data and the least-squares regression line. Compute the uncertainty in the slope using (6.10) and plot the two regression lines going through the origin that correspond to this range in slopes. Finally, determine the orthogonal regression solutions (both Major Axis and Reduced Major Axis) and plot them as well. Do these fall within the range of acceptable least-squares regression trends?

Problem 6.3. The file *c2407.txt* contains distance x (in km), bathymetry z (in m), and crustal age t (in Myr) along a segment across the Mid-Atlantic Ridge.

a) Find the deterministic component of seafloor depths given by the theoretical prediction

$$z = d_r + c\sqrt{t},$$

and report the parameters. Note that z is a function of age, not distance.

b) Plot the bathymetry and your best-fitting deterministic component on the same plot, both as functions of distance (not age).

Problem 6.4. The files *hilo.txt* and *honolulu.txt* contain tide gauge readings (the first column is time in years and the second is monthly average sea level values in mm relative to some arbitrary origin) from Hilo and Honolulu harbors, respectively. It is believed that the observations represent the sum of two phenomena: (1) eustatic (global) sea-level variations, and (2) tectonic subsidence of the Big Island.

a) Plot the data. Use `regress_ls.m` to determine the regression lines for the two series. Superimpose these lines on the graphs and label the plots with the values of the regression slopes. What do these slopes represent?

b) Assuming Oahu is tectonically stable, what is the tectonic subsidence rate at Hilo?

c) Alternatively, we only use the time period that the two data series have in common and directly subtract the Honolulu series from the Hilo series. What is the tectonic subsidence rate based on these differences, and how does this result compare to your answer in (2)?

Problem 6.5. Data sets showing the separate weights of an animal and its brain exhibit a strong correlation in log-log space (*bb_weights.txt*). Use the methodology of "reweighted least squares" (RLS) to determine this regression and calculate residual z-scores relative to the RLS regression. With a threshold of 2.5, which animals depart from the simple trend? How would you explain these outliers?

Problem 6.6. The file *sherwood.txt* contains a series of petrophysical measurements on samples of the Triassic Sherwood sandstone [from Lourenfemi, M.O., 1985, *Math Geol., 17*, 845–452.] The five columns contain permeability, porosity (ϕ), matrix conductivity, true formation factor, and induced polarization. Using multiple regression, how many of the last four parameters are significant in a regression to explain the permeability, assuming a 95% level of confidence?

Chapter 7

SEQUENCES AND SERIES ANALYSIS

"If you can't explain it simply, you don't understand it well enough."

Albert Einstein, Physicist

The geosciences are replete with observational data that can be viewed as ordered sequences. Their single most important property is that they form a *sequence*, and the *positions* where data points occur within the sequence are paramount. Contrast that arrangement to a set of repeated measurements of some quantity, say ten determinations of sandstone densities. Our goal of determining the average density is not affected by how we order our data — the order in the sequence is not important. Sequential data therefore often consist of series with *pairs* of variables: The first indicates the position in the sequence, the other gives the observation. A special family of such sequences consists of those in which an observation is given as a function of *time*. The analysis of such data is traditionally called *time series analysis*. In this book we will extend these techniques to include "space series" as well, i.e., time and distance will be considered interchangeable, and we will discuss such data and the methods used to analyze them at length in Chapter 8.

We can think of many examples of natural data that are time — or space — sequences, e.g., a series of temperatures as a function of time, tide gauge readings taken over a long period, topography measured along a transect, and much more. While such data sets will be our main concern, we should not forget that much sequential data do not have the format of a "time"-series. We might be considering a stratigraphic sequence consisting of the lithologic states encountered in a sedimentary succession. The stratigraphy might show a cyclothem of shale–sandstone–shale–sandstone–shale–coal, etc., going from top to bottom. We would like to investigate the significance of the succession, but cannot put a meaningful scale on the sequence. It is clear that the succession of lithologies represents changes over time, but we have no way of estimating the time scale. Could we use thickness as a proxy for time? Thickness is certainly related to time through sedimentation rates, but these are known to vary. Additional complications include hard-to-estimate effects like compaction and erosion. Furthermore, the thicknesses are likely to change significantly from location to location. Thus, if we use thickness or any other measure of down-hole position it may obscure the examination of the *successions*, which is the primary objective of interest. For example, consider an observation that sandstone is the second state and coal the sixth state in a sequence. Clearly, this relationship has no meaning that can be expressed numerically, e.g., "sixth" is not $3\times$ "second". Obviously, we have here a problem of a different nature than the usual time-series analysis mentioned above.

Yet another type of sequence is a series of *events*. Such data may be historical records of earthquakes in California, volcanic eruptions of Mauna Loa, or reversals of the Earth's magnetic field. In these cases, the data simply consist of the time interval *between* events or a cumulative length of time over which the events occurred, and special analysis techniques are required.

Thus, the nature of the sequential data and the type of sequence determine what questions we may hope to answer by subjecting our data to analysis. The purpose of any of these methods is to facilitate answers to questions such as these:

1. Are the data random, or do they exhibit a trend or pattern?

2. If there is a trend, what form does it have?
3. Are there any periodicities in the data?
4. What can be estimated or predicted from the data?
5. Are there other questions specific to the situation?

We shall see that while we often are concerned with the analysis of a single sequence of data, there are many instances in which we want to compare two or more sequences. One obvious example to geologists is *stratigraphic correlation*, based on either lithologic sections or well log data. Sequence correlation may speed up routine correlations and detect subtle correlations which may be hard to detect by eye.

The methods for comparing two or more sequences can be grouped into two broad classes. In the first, the exact position in a sequence matters, and a correlation is only significant if it takes place at the correct location. One example is the comparison of an X-ray diffraction chart with standard charts in attempts to recognize minerals. The comparison can only take place at certain angles. If the shape of a spectral peak centered on 20° in the data looks exactly like the bump at 30° in the standard chart it is of no significance: both peaks would have to occur at the same angle.

In the other class of methods the absolute position is not important, only relative position matters. These processes, like *cross-correlation*, are very similar to the mental process of geologic correlation. However, these methods are limited because they cannot take stretching and compression of the scale into account. Nevertheless, in many problems there is no distortion and we may use such techniques with some success.

7.1 Markov Chains

As mentioned above, many geological experiments result in data sequences consisting of ordered successions of *mutually exclusive* states. We already mentioned the lithology variations in stratigraphic sections. Other examples include:

- The changes in minerals across a line in a thin-section.
- Drill holes through zoned ore bodies.

The observations may be obtained at evenly spaced intervals, or we may simply register the position when a change of state occurs. In the first case we would expect repeated states; the latter will obviously not contain such runs since we only record changes of state.

Such data may be subjected to *cross-association* and/or *auto-association* techniques, but right now we are primarily concerned with the nature of transitions rather than the relative positions of states in the sequence. Therefore we will, for the moment, pay less attention to the positions of observations within the succession and instead concentrate on acquiring information about the *tendency* of one state to follow another.

Example 7–1.

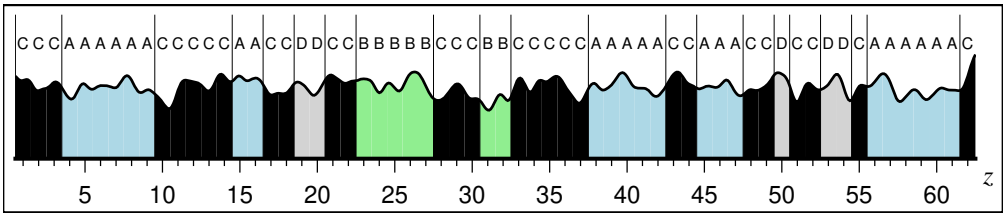


Figure 7.1: Example of a stratigraphic section with four separate lithologies. Assessing the lithology at every 1-foot interval down the section resulted in 62 states and thus we recorded 61 transitions of state.

Let us look at an example of a stratigraphic section. Here, we have determined the lithology at one foot intervals down a stratigraphic section. This exercise resulted in a sequence of states. We find four mutually exclusive states: A) sandstone, B) limestone, C) shale, and D) coal. There are 62 observed states, hence 61 transitions. We tabulate the

transitions in a 4 x 4 matrix, since we have four possible transitions for each of the four states. E.g., from sandstone (A) we may change to A, B, C, or D, since repeated states may occur. The *transition frequency matrix* **A** given in Table 7.1 expresses all the observed possibilities. We can now see that element a_{ij} reads "number of transitions from

	A	**B**	**C**	**D**	**Row Total**
A	17	0	5	0	22
B	0	5	2	0	7
C	5	2	17	3	27
D	0	0	3	2	5
Column total	22	7	27	5	61

Table 7.1: Transition frequency matrix for the transitions seen in Figure 7.1.

state i to state j". For our section, the matrix is symmetric. However, in general this will not be the case, so $a_{ij} \neq a_{ji}$.

The tendency for one state to succeed another can be made clearer by converting the frequencies to percentages or fractions. We do this by dividing each row by its row total. These percentages may be considered conditional probabilities in that they measure the probability that state j will follow *given* that the present state is i, which we write as $P(j|i)$ or $P(i \rightarrow j)$. The resulting *transition probability matrix* **P** is given by Table 7.2 and graphically illustrated in Figure 7.2.

From/To	**A**	**B**	**C**	**D**
A	0.77	0.00	0.23	0.00
B	0.00	0.71	0.29	0.00
C	0.19	0.07	0.63	0.11
D	0.00	0.00	0.60	0.40

Table 7.2: The transition probability matrix for the transitions in Table 7.1.

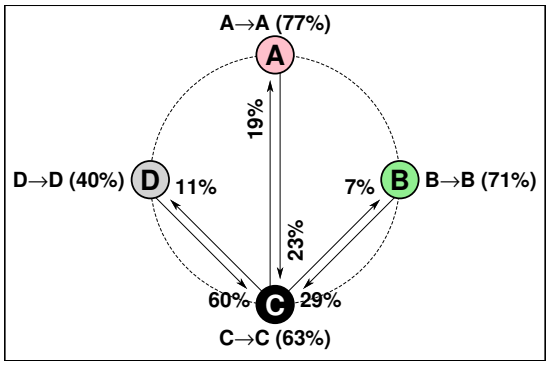


Figure 7.2: A cyclic diagram may be used to represent the frequencies of transition between the various lithologies.

This operation will usually result in an asymmetrical matrix. If we divide the row totals (the counts) by the grand total number of counts we find the relative *proportions* of the four lithologies. This is called the marginal or *fixed probability vector*:

$$\mathbf{f} = \begin{bmatrix} 0.36 & 0.12 & 0.44 & 0.08 \end{bmatrix}. \tag{7.1}$$

You may remember (if not, brush up on the probability theory in Chapter 3) that the joint probability of two events A and B is

$$P(A \cap B) = P(B|A) \cdot P(A), \tag{7.2}$$

which we can rearrange to give

$$P(B|A) = \frac{P(A \cap B)}{P(A)}. \tag{7.3}$$

The probability that state B will follow A is the probability that both states A and B will occur, divided by the probability that A occurs. Now, if A and B are independent states, then (e.g., 3.24)

$$P(A \cap B) = P(A) \cdot P(B). \tag{7.4}$$

Therefore, if there are no dependencies then the probability that B will follow A is simply the probability that B occurs. This must hold for all independent states, so

$$P(B|A) = P(B|B) = P(B|C) = P(B|D) = P(B). \tag{7.5}$$

This result provides us with an opportunity to predict what the transition probability matrix should look like if the occurrence of a lithologic state at one point were completely independent of the lithology at the underlying point. Naturally, that matrix will have rows matching the fixed probability vector. So, for our stratigraphic example, we find the expected matrix to be

	A	**B**	**C**	**D**
A	0.36	0.12	0.44	0.08
B	0.36	0.12	0.44	0.08
C	0.36	0.12	0.44	0.08
D	0.36	0.12	0.44	0.08

Finally, we are now in the position to compare the observed transition frequencies to the predicted frequencies and test the null hypothesis that all lithologic states are independent of the state immediately below it. To do so we use a χ^2-test after first converting the expected percentages back to counts or frequencies. We find

$$\begin{bmatrix} 22 & & & \\ & 7 & & \\ & & 27 & \\ & & & 5 \end{bmatrix} \cdot \begin{bmatrix} 0.36 & 0.12 & 0.44 & 0.08 \\ 0.36 & 0.12 & 0.44 & 0.08 \\ 0.36 & 0.12 & 0.44 & 0.08 \\ 0.36 & 0.12 & 0.44 & 0.08 \end{bmatrix} = \begin{bmatrix} 7.9 & 2.6 & 9.7 & 1.8 \\ 2.5 & 0.8 & 3.1 & 0.6 \\ 9.7 & 3.2 & 11.9 & 2.2 \\ 1.8 & 0.6 & 2.2 & 0.4 \end{bmatrix}. \tag{7.6}$$

The test statistic is, as usual, given by

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}, \tag{7.7}$$

where O_i is the observed number of transitions and E_i is the expectation for each transition, as given by (7.6). The degrees of freedom, ν, is $(k-1)\cdot(k-1)$, with $k = 4$. One degree of freedom is lost from each row and column because all rows of $\mathbf{P}$ must sum to 1 and we computed $\mathbf{f}$ from the row sums.

For the χ^2-test to be valid, each category should have an expected value of at least 5. Several of our categories do not fulfill that criteria. Because we only are testing whether the transition frequencies are independent (random) or not, we may combine some categories to raise the expected values above 5. Hence, we use the four largest categories $A \to A, A \to C, C \to A, C \to C$ and the combinations $B \to$ any, $D \to$ any, and $[A \to B, A \to D, C \to B, C \to D]$. We find χ^2 to be

$$\chi^2 = \frac{(17-7.9)^2}{7.9} + \frac{(5-9.7)^2}{9.7} + \frac{(5-9.7)^2}{9.7} + \frac{(17-11.9)^2}{11.9} + \frac{(7-7.0)^2}{7.0} + \frac{(5-5.0)^2}{5.0} + \frac{(5-9.8)^2}{9.8} = 19.57. \tag{7.8}$$

From Table A.3 we find the critical value of χ^2 with $\nu = 9$ and a 5% level of significance to be 16.92. Since our computed value exceeds the critical value, we must reject the hypothesis that successive states are independent. It appears there is a significant tendency for certain states to be followed by certain other states.

Sequences in which the state at one point is *partially* dependent, in a probabilistic sense, on the previous state is called a *Markov chain*. It is intermediate between deterministic (fully predictable) and completely random sequences. The section we examined has first-order Markov properties. This means there is statistical dependency between points and their immediate predecessor.

Higher-order Markov chains can exist as well. We can use the transition probability matrix to predict what the lithology might be *two* feet above a point. For example, we might want to fill in the missing part of a section in the

statistically most reasonable way. Let us say we start in state B (limestone). The probabilities of reaching the next state is then given as

$$\begin{array}{lll} B \to A & \text{(sandstone)} & 0\% \\ B \to B & \text{(limestone)} & 71\% \\ B \to C & \text{(shale)} & 29\% \\ B \to D & \text{(coal)} & 0\% \end{array} \tag{7.9}$$

Let us pretend that the next state is shale (C). Then, reaching the following state would be associated with the probabilities

$$\begin{array}{lll} C \to A & \text{(sandstone)} & 19\% \\ C \to B & \text{(limestone)} & 7\% \\ C \to C & \text{(shale)} & 63\% \\ C \to D & \text{(coal)} & 11\% \end{array} \tag{7.10}$$

Therefore, the probability that the sequence will be limestone $\to$ shale $\to$ limestone is

$$P(B \to C) \cdot P(C \to B) = 29\% \cdot 7\% = 2\%. \tag{7.11}$$

However, we can also reach limestone in two steps by way of the limestone $\to$ limestone $\to$ limestone path. Now

$$P(B \to B) \cdot P(B \to B) = 71\% \cdot 71\% = 50\%. \tag{7.12}$$

Since $B \to A$ and $B \to D$ have zero probability, we can state that the probability of finding limestone two steps up above limestone, regardless of intervening lithology, is the sum

$$P(B \to ? \to B) = P(B \to B \to B) + P(B \to C \to B) = 50\% + 2\% = 52\%. \tag{7.13}$$

One can use the same approach to calculate the probability of any lithology two steps up, but fortunately there is a more efficient way: These multiplications and additions are exactly those that define a matrix multiplication. Multiplying the transition matrix by itself (i.e., squaring it) yields $\mathbf{P}^2$, which describes the *second-order* Markov properties of the stratigraphic section:

$$\begin{bmatrix} 0.77 & 0 & 0.23 & 0 \\ 0 & 0.71 & 0.29 & 0 \\ 0.19 & 0.07 & 0.63 & 0.11 \\ 0 & 0 & 0.60 & 0.40 \end{bmatrix}^2 = \begin{bmatrix} 0.64 & 0.02 & 0.32 & 0.02 \\ 0.06 & 0.52 & 0.39 & 0.03 \\ 0.27 & 0.09 & 0.53 & 0.11 \\ 0.11 & 0.04 & 0.62 & 0.23 \end{bmatrix} \tag{7.14}$$

Note that again the rows have unit sums. If we wanted to know whether the second-order Markov properties are significant we convert the frequency percentages to counts by multiplying $\mathbf{P}^2$ by the observed row totals again. However, this time the product will approximate the second-order transition we would have *likely* observed had we measured them directly from the data. We find

$$\begin{bmatrix} 14.1 & 0.4 & 7.0 & 0.6 \\ 0.4 & 3.7 & 2.7 & 0.2 \\ 7.1 & 2.5 & 14.2 & 3.1 \\ 0.6 & 0.2 & 3.1 & 1.1 \end{bmatrix} \tag{7.15}$$

and carry out another χ^2-test. This time we obtain $\chi^2 = 7.73$ (critical value is still 16.92 since the expectations are independent of the step length). Hence, we must conclude that there are no significant second-order Markov properties present and that the lithology two steps away appears to be a random selection given the volume distribution of the rock types.

7.2 Embedded Markov Chains

The choice of a sampling interval introduces an arbitrary element into our sequence analysis. This can be avoided if we only record the transitions of state whenever they occur. It follows that the transition frequency matrix will have zeros along the diagonal since no state can follow itself. Sequences which cannot contain repeated states are called *embedded Markov chains*.

Example 7–2. Let us look at a particular example from a borehole through a sedimentary delta plain in Scotland. We have five lithologies: A (mudstone), B (shale), C (siltstone), D (sandstone), and E (coal). The analysis yields

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \left[\begin{array}{ccccc} 0 & 11 & 36 & 21 & 52 \\ 28 & 0 & 4 & 4 & 0 \\ 34 & 2 & 0 & 45 & 13 \\ 29 & 1 & 45 & 0 & 3 \\ 28 & 23 & 9 & 8 & 0 \end{array}\right] \\ \\ \end{array} \begin{array}{c} \\ = \\ = \\ = \\ = \\ = \end{array} \begin{array}{c} \Sigma \\ 120 \\ 36 \\ 94 \\ 78 \\ 68 \\ \hline 396 \end{array} \tag{7.16}$$

The fixed probability vector is found by dividing the row totals by the grand total:

$$\mathbf{f} = [0.30 \quad 0.09 \quad 0.24 \quad 0.20 \quad 0.17]. \tag{7.17}$$

To test whether the observed sequence has Markovian properties or independent states we may use the χ^2-test in a similar way to what we did above. The problem is that we cannot use the fixed vector to estimate the independent transition frequency matrix since that results in nonzero diagonal terms, which is forbidden. We must therefore use a different method to estimate the necessary matrix.

Imagine our sequence is a *censored* sample from a sequence where repeats *may* occur. Its transition matrix would be identical to the one we observed except it would have nonzero diagonal terms. If we converted this matrix to probabilities and raised it to a high power, we would find the transition probability matrix for a sequence with independent states. We could then discard the diagonal terms, adjust the off-diagonal terms (to ensure they sum to 1), and end up with P for an embedded sequence of independent states. This result is achieved by trial-and-error since we do not know the number of repeated states in the envisioned sequence. We want to find diagonal entries which do not change when the matrix is raised to higher powers. An iterative scheme is used:

1. Place arbitrary large estimates (1–2 magnitudes larger that your observations) into the diagonal positions in the observed matrix.
2. Divide row totals by the grand total to get diagonal probabilities.
3. Calculate new diagonal estimates by multiplying the diagonal probabilities from step 2 by the latest row sums.
4. Repeat process steps (2) and (3) until the diagonal terms remain unchanged, typically after 10–20 iterations.

We will try this procedure on the Scottish data. For our matrix, we try inserting 1000 first:

$$\begin{array}{c} \left[\begin{array}{ccccc} 1000 & 11 & 36 & 21 & 52 \\ 28 & 1000 & 4 & 4 & 0 \\ 34 & 2 & 1000 & 45 & 13 \\ 29 & 1 & 45 & 1000 & 3 \\ 28 & 23 & 9 & 8 & 1000 \end{array}\right] \end{array} \begin{array}{c} \\ = \\ = \\ = \\ = \\ = \end{array} \begin{array}{c} \\ 1120 \\ 1036 \\ 1094 \\ 1078 \\ 1068 \\ \hline 5396 \end{array} \tag{7.18}$$

We next obtain the new diagonal probabilities to be

$$\left[\begin{array}{ccccc} 0.208 & & & & \\ & 0.192 & & & \\ & & 0.203 & & \\ & & & 0.200 & \\ & & & & 0.198 \end{array}\right]. \tag{7.19}$$

Step (3) is to update the diagonal elements via the new row sums, hence we find

$$\begin{array}{c} \left[\begin{array}{ccccc} 233 & 11 & 36 & 21 & 52 \\ 28 & 199 & 4 & 4 & 0 \\ 34 & 2 & 222 & 45 & 13 \\ 29 & 1 & 45 & 215 & 3 \\ 28 & 23 & 9 & 8 & 212 \end{array}\right] \end{array} \begin{array}{c} \\ = \\ = \\ = \\ = \\ = \end{array} \begin{array}{c} \\ 353 \\ 235 \\ 316 \\ 294 \\ 280 \\ \hline 1478 \end{array} \tag{7.20}$$

Repeating step (2) with the new matrix gives

$$\begin{bmatrix} 0.239 & & & & \\ & 0.159 & & & \\ & & 0.214 & & \\ & & & 0.199 & \\ & & & & 0.189 \end{bmatrix}, \tag{7.21}$$

and we keep repeating this process until it stabilizes. In the end we find the marginal probability vector:

$$\begin{bmatrix} 0.335 & & & & \\ & 0.074 & & & \\ & & 0.235 & & \\ & & & 0.181 & \\ & & & & 0.155 \end{bmatrix} \tag{7.22}$$

with a corresponding grand total of 524.

Now, because all states are independent in the random case we will test for, the probability that state j will follow state i is simply $P(i \rightarrow j) = P(i) \cdot P(j)$. This allows us to construct the *expected* transition probability matrix

$$\mathbf{P}_e = \begin{bmatrix} 0.125 & 0.026 & 0.083 & 0.064 & 0.055 \\ 0.026 & 0.006 & 0.017 & 0.013 & 0.012 \\ 0.083 & 0.017 & 0.055 & 0.043 & 0.03 \\ 0.064 & 0.013 & 0.043 & 0.033 & 0.028 \\ 0.055 & 0.012 & 0.036 & 0.028 & 0.024 \end{bmatrix}. \tag{7.23}$$

Scaling these probabilities by the grand total gives the expected frequencies

$$\mathbf{E} = \begin{bmatrix} 65.6 & 13.6 & 43.5 & 33.5 & 28.8 \\ 13.6 & 3.1 & 8.9 & 6.8 & 6.3 \\ 43.5 & 8.9 & 28.8 & 22.5 & 18.9 \\ 33.5 & 6.8 & 22.5 & 17.3 & 14.7 \\ 28.8 & 6.3 & 18.9 & 14.7 & 12.6 \end{bmatrix}. \tag{7.24}$$

We strip off the diagonal elements and use the off-diagonal counts to evaluate the χ^2 statistic. In this particular case, we find $\chi^2 = 172$ which greatly exceeds the critical value of 19.68 for $v = (k-1)^2 - k = 11$ degrees of freedom; the test indicates a strong first order Markov sequence.

7.3 Series of Events

One of many types of time series that occur in the natural sciences is the *series of events*. Examples of such sequences include the historical records of earthquake occurrences, volcanic eruptions, floods, storms and hurricanes, geomagnetic reversals (Figure 7.3), landslides, and tsunamis. These series share some common characteristics:

- Events are distinguished based on *when* they occur in time.
- Events are essentially *instantaneous* in the context of your range.
- Events are so *infrequent* that they do not overlap in time.

In some cases, spatial data sequences may be considered a series of events. Consider a traverse across a thin-section. We may be interested in the occurrence of some rare mineral. Another possibility may be the occurrence of bentonite (volcanic ash layers) in a sedimentary sequence. However, when a spatial scale acts as a proxy for the actual time scale, we know that the analysis will be susceptible to errors caused by varying sedimentation rates, compaction, and erosion.

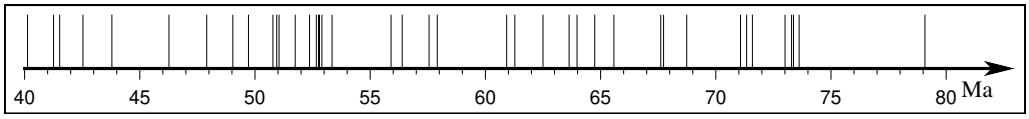


Figure 7.3: Times of reversals in the Earth's magnetic field during a 40 million year interval. While each reversal may take a few thousand years to complete, these reversals are essentially instantaneous when seen as part of the much longer geological record.

With most studies of series of events we hope to find out what the basic features of the series are and how we can relate the distribution of length intervals to a physical mechanism. We must first consider the possibility of a trend in the data. Thus, we will use a test designed to detect trends in the rate of occurrence. It works by simply comparing the mean (or centroid) of a series to its midpoint and test their separation for significance. The centroid is given by

$$\bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i. \tag{7.25}$$

We find

$$z = \frac{\bar{t} - t_{1/2}}{T/\sqrt{12n}}, \tag{7.26}$$

where $t_{1/2}$ is the half-point and T is the length of series. The denominator is obtained by taking the standard deviation of a uniform distribution over the range $\{0,T\}$ (which equals $T/\sqrt{12}$) and using the central limits theorem to give the expected standard deviation of the centroid. The test is very sensitive to changes in the rate of occurrence since the centroid is a L_2 estimate of location. If no trend is detected, then we may conclude that the series is *stationary*. Many geological processes produce events that should be uniformly distributed in time. For instance, the steady motion of the tectonic plates produce a steady increase in stress on a fault, which will slip to relieve the stress. To test for uniformity we note that the cumulative distribution of a uniform series is a straight line from 0 to 1 over the time interval. We can then compare this line to the stair-step cumulative function of our observed series of events and apply the Kolmogorov-Smirnov test on the largest discrepancy, as discussed in Chapter 4.

7.4 Run Test

The simplest sequence imaginable is a succession of observations that can take on only two mutually exclusive states or values. Consider a rock collector looking for fossils: Each time he or she opens a concretion there may or may not be a fossil present. We find true or false, yes or no, or 1 or 0. Similar sequences are generated by coin tosses. Twenty tosses may give the series

$$HTHHTHTTTHTHTHHTTHHH \tag{7.27}$$

with 11 heads and 9 tails, close to the expected 10/10. The probability of finding a given number of heads (x) in a series of n tries is given by the binomial distribution (3.70). However, there is nothing in that expression that takes the *order* in which the heads appear into account. We would find a sequence of 10H followed by 10T to be highly unlikely; the same goes for alternate HTH... Yet, the probability of the *number* of heads is unchanged. We test such binary sequences for randomness of occurrence by examining the number of *runs*, defined as *uninterrupted sequences of the same state*.

Example 7–3. In our sequence above (7.27) we have the following 13 runs:

$$\begin{array}{cccccccccccccc} H & T & HH & T & H & TTT & H & T & H & T & HH & TT & HHH \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \end{array} \tag{7.28}$$

This is a job for the U-test. We test the significance of the runs by finding all possible ways of arranging n_1 items of state 1 and n_2 items of state 2. The total number of runs is called U, and we can consult tables for critical values of U given n_1, n_2, and α (our confidence level). However, for large $n_1, n_2 > 10$ the distribution is approximated by a normal distribution with a mean of

$$\bar{U} = \frac{2n_1 n_2}{n_1 + n_2} + 1 \tag{7.29}$$

and variance

$$s_U^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}. \tag{7.30}$$

We may then simply use the z-statistic

$$z = (U - \bar{U})/s_U \tag{7.31}$$

and see if our calculated z value exceeds the $\pm z_{\alpha/2}$ interval. For our H/T series $n_1 = 11$ and $n_2 = 9$, so we find

$$\bar{U} = \frac{2 \cdot 11 \cdot 9}{11+9} + 1 = 10.9 \text{ and} \tag{7.32}$$

$$s_U = \sqrt{\frac{(2 \cdot 11 \cdot 9)(2 \cdot 11 \cdot 9 - 11 - 9)}{(11+9)^2(11+9-1)}} = \sqrt{4.6} = 2.1, \tag{7.33}$$

which gives

$$z = \frac{13 - 10.9}{2.1} = 1.0. \tag{7.34}$$

With $z_{\alpha/2} = z_{0.025} = \pm 1.96$ (i.e., Table A.2) we cannot reject the null hypothesis that the sequence appears random.

The geological application of run tests may seem somewhat obscure, since most data consist of more than two mutually exclusive states. A related procedure is a statistical method for examining runs up and down. Here we are again considering two distinct "states", i.e., whether an observation is larger or smaller than the preceding observation. Let us examine the data set illustrated in Figure 7.4.

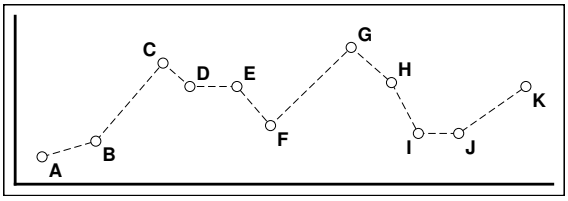


Figure 7.4: Example of how a run test can be used on continuous data. Any multipoint segment with the same sign of slope is defined as a "run".

The segment ABC is a "run up" since the slopes AB and BC are both positive. Likewise, GHI is a "run down". CDEF is also down because all slopes are negative except DE, which is zero. IJ can be part of GHIJ or IJK. For most floating point data there will be few points that exactly equal their neighbors. If we only consider the sign of the slope we get the sequence

$$+ \; + \; - \; 0 \; - \; + \; - \; - \; 0 \; +$$

Regarding the first 0 as a '$-$' and the second 0 as a '$+$', we find five runs: three of '$+$' and two of '$-$'. Then, the U-test is directly applicable. Again, we need more than 10 occurrences of each type to use the normal distribution approximation introduced earlier. It is clear that one can apply runs test by converting data to a binary series by almost any method, provided the hypothesis tested reflects the *dichotomizing* method. A common technique is to dichotomize a series by removing the median or mean value and look for randomness of runs about the central location.

Example 7–4. Consider density measurements of ore samples across a magnetite body. We want to know if the densities vary randomly about the median or if a trend is present. The data are given in Table 7.3. The median density is found to be 3.98. We subtract this value and store the signs of the deviations below the corresponding density. We observe $U = 7$, with $n_1 = 19$, $n_2 = 20$. For these numbers,

$$\bar{U} = 20.5, \;\; s_U = 3.1, \tag{7.35}$$

and we obtain $z = -4.4$. Because our observed U is far outside both the 95% and 99% confidence intervals we conclude that the variations about the median are not random but systematic.

3.57	3.63	2.86	2.94	3.42	2.85	3.67	3.78	3.86	4.02
-	-	-	-	-	-	-	-	-	+
4.56	4.62	4.31	4.58	5.02	4.68	4.37	4.88	4.52	4.80
+	+	+	+	+	+	+	+	+	+
4.55	4.61	4.93	4.60	4.51	3.98	4.22	3.52	2.91	3.87
+	+	+	+	+	+	+	-	-	-
3.52	3.77	3.84	3.92	4.09	3.86	4.13	3.92	3.54	
-	-	-	-	+	-	+	-	-	

Table 7.3: Density measurements (ρ) and their signs indicating if ρ_i is larger or smaller than the median density.

There are of course many more variants of the run test shown here. In general, such tests are *nonparametric* in that they do not require the underlying distribution to be known to us.

7.5 Autocorrelation

The main purpose of time series analysis is to take the order of the observations into account and try to learn the properties of the data set, such as discovering any periodicities, trends, or repeating patterns, and then use such characteristics to infer something about the process being observed. Repetitions and other patterns in a sequence can be found by computing a measure of the "self-similarity" of the sequence, that is, to what extent a piece of the sequence looks like another piece of the same sequence. One such measure is known as the *autocorrelation*.

In Section 3.2.6 we discussed the correlation between two variables x_i and y_i and found it to be given by

$$r = \frac{s_{xy}}{s_x s_y}, \tag{7.36}$$

where s_x, s_y are the sample standard deviations and s_{xy} the sample covariance, given by

$$s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{n-1}. \tag{7.37}$$

The concept of the autocorrelation is to let both x_i and y_i be the same signal y_i and then compare these two, identical, time-series. Of course, with $x_i = y_i$ we find

$$s_{yy} = \frac{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2}{n-1} = s_y^2 \tag{7.38}$$

and

$$r = \frac{s_y^2}{s_y s_y} = 1, \tag{7.39}$$

meaning the signal is perfectly correlated with itself.

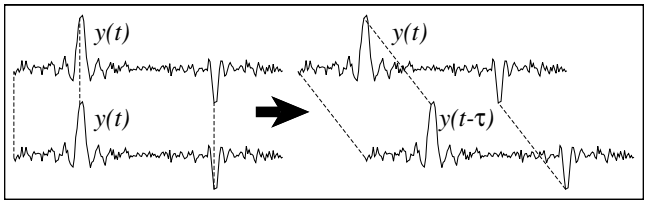


Figure 7.5: Autocorrelation is the correlation between a time-series and its identical clone at different lags, τ.

To get more useful information from the autocorrelation we will shift all values in the second copy one step to the left (Figure 7.5). So, instead of having the covariance be composed of the terms

$$s_{yy}(0) \propto y_1 \cdot y_1 + y_2 \cdot y_2 + \ldots + y_n \cdot y_n \tag{7.40}$$

we will instead have

$$s_{yy}(1) \propto y_2 \cdot y_1 + y_3 \cdot y_2 + \ldots + y_n \cdot y_{n-1}. \tag{7.41}$$

We then get

$$s_{yy}(1) = \frac{\sum_{i=2}^{n} y_i y_{i-1} - \frac{1}{n-1}\sum_{i=2}^{n} y_i \sum_{i=2}^{n} y_{i-1}}{n-2}. \tag{7.42}$$

Shifting one step further will give a different result. In general, if we shift the second sequence τ steps relative to the first sequence we find the series' *autocovariance*:

$$s_{yy}(\tau) = \frac{\sum_{i=1+\tau}^{n} y_i y_{i-\tau} - \frac{1}{n-\tau}\sum_{i=1+\tau}^{n} y_i \sum_{i=1+\tau}^{n} y_{i-\tau}}{n-\tau-1} = \frac{(n-\tau)\sum_{i=1+\tau}^{n} y_i y_{i-\tau} - \sum_{i=1+\tau}^{n} y_i \sum_{i=1+\tau}^{n} y_{i-\tau}}{(n-\tau)(n-\tau-1)}, \tag{7.43}$$

where we call the number of shifts, τ, the *lag*. It is assumed throughout our time-series discussion that the sequences are evenly spaced with spacing Δt and contain n points, so that the length of a sequence is $T = \Delta t(n-1)$. Figure 7.6 illustrates the situation for a certain lag.

Computing the autocovariance for all lags from $\tau = 0$ to about $\tau = n/4$ results in the autocovariance function $s_{yy}(\tau)$. This function will tell us if the sequence exhibits self-similarity and how much we must shift it (i.e., what is the lag) to reach a maximum in the autocovariance.

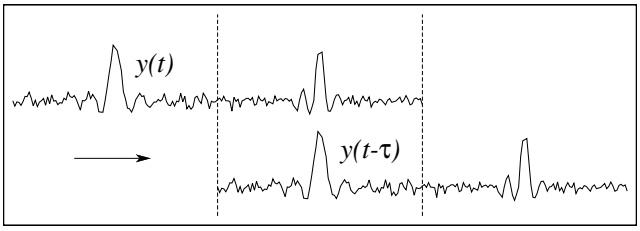


Figure 7.6: Autocorrelations can be high for large lags τ if the data have "repeating" features.

Plotting the autocovariance function for our sequence gives an *autocovariogram*, illustrated in Figure 7.7.

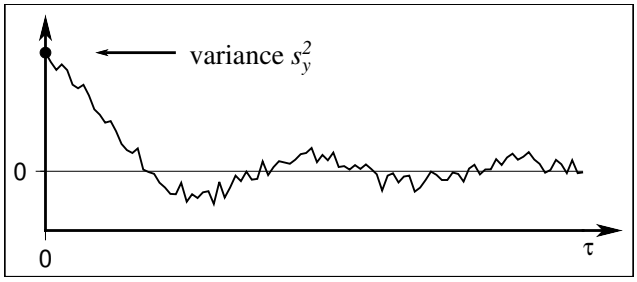


Figure 7.7: The autocovariogram for a typical time-series. At zero lag we obtain the variance of the time-series.

As was the case with the covariance for a set of paired values, the autocovariance depends on the units of the data, which makes it less useful for comparison purposes. Again, the solution is to normalize it by the variance of the

sequence. We find the variance to be given by

$$s_y^2 = \frac{\sum_{i=1+\tau}^{n} (y_i - \bar{y})^2}{n-\tau-1}. \tag{7.44}$$

If we assume the mean and variance remain unchanged by the lag τ, we obtain

$$r_\tau = \frac{(n-\tau)\sum_{i=1+\tau}^{n} y_i y_{i-\tau} - \sum_{i=1+\tau}^{n} y_i \cdot \sum_{i=1+\tau}^{n} y_{i-\tau}}{(n-\tau)(n-\tau-1)\left[\frac{1}{(n-\tau-1)}\left(\sum_{i=1+\tau}^{n} y_i^2 - (n-\tau)\bar{y}^2\right)\right]} \tag{7.45}$$

which reduces to

$$r_\tau = \frac{\sum_{i=1+\tau}^{n} y_i y_{i-\tau} - (n-\tau)\bar{y}^2}{\sum_{i=1+\tau}^{n} y_i^2 - (n-\tau)\bar{y}^2}. \tag{7.46}$$

The effect of normalizing by the variance is to obtain the *autocorrelation* which only takes on values in the $[-1,+1]$ range. This *autocorrelogram* for our sequence remains unchanged in shape but now has a maximum of $+1$ for zero lag, i.e., the series is in perfect correlation with itself.

Since it is arbitrary if we consider one copy of the sequence shifted by $-\tau$ or the other by $+\tau$, the autocorrelation function is symmetric about zero lag, i.e.

$$r_{-\tau} = r_\tau. \tag{7.47}$$

The autocorrelogram can be used to reveal characteristics of a time series. Commonly, one would like to compare the observed correlogram to predicted autocorrelograms for simple models or processes. The simplest of all models is the one in which successive observations are (1) independent and (2) normally distributed. Since each observation y_i is independent of any other observation $y_{i-\tau}$ we expect

$$r_\tau = \begin{cases} 1, & \tau = 0 \\ 0, & \text{elsewhere} \end{cases} \tag{7.48}$$

The expected autocorrelation for a totally random process (also called *white noise*) is zero, with a variance of $\sigma^2 = 1/n$, when $n > 30$ (Figure 7.8).

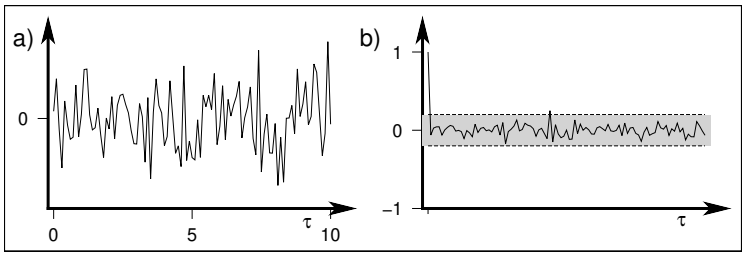


Figure 7.8: White noise and its autocorellogram. White noise is completely uncorrelated, thus the expected value of correlation is zero for nonzero lags. The finite length of a time-series leads to departures from this theoretical prediction, and the gray band indicates expected variation from zero correlation for a 95% confidence level.

Stationary time-series with short term correlations will typically have an autocorrelogram where the first few coefficients are significantly nonzero whereas the remainder are close to zero. If the time series has a tendency to alternate direction at every step, then the autocorrelogram will alternate too, with $r_1 < 0$. If our time-series is nonstationary (i.e., it includes a trend), then r_τ will not approach zero except for very large values of the lag. Other correlations tend to be completely masked by the tendency for an observation to be systematically larger (or smaller)

than its predecessor. It is therefore always necessary to remove linear trends from time-series prior to analysis of r_τ. Strict periodicities in the data will be mimicked in the correlogram: A signal $y_t = A\cos\omega t$ will have an autocorrelation of $r_\tau \sim \cos\omega\tau$ (for large n).

Outliers can wreak havoc on the estimation of autocorrelation coefficients. This is not surprising since estimating the correlation is an L_2 process. It is therefore important to suppress any outliers, using robust statistical methods, to insure good and stable coefficients.

7.6 Cross-Correlation

Rather than using two identical series, an obvious extension of the autocorrelation method is to compare two *different* time-series at various lags. From such an undertaking we would expect to learn two things: 1) The strength of the relationship between the two series, and 2) the lag that maximizes the correlation. This process is called *cross-correlation*, and it differs from autocorrelation in several ways:

1. It may not be possible to specify the zero lag position, unless the two series share a common origin and scale.
2. The cross-correlation will in general be asymmetric.
3. The two series may be of different length.

The correlation coefficient for the match position τ (relative to an arbitrary origin, unless the series have a common origin) is simply

$$r_\tau = \frac{(n-1)\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i-1}^{n} y_i}{\sqrt{\left((n-1)\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)\left((n-1)\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right)}}, \tag{7.49}$$

where x_i and y_i are the two series and the sum over n represents the point pairs that overlap for this particular lag position τ (Figure 7.9). One difference with the expression for the autocorrelation is that the denominator will depend (and thus varies) with n, while for the autocorrelation we used the variance for the entire chain. For this reason the cross-correlation is somewhat less stable. A simple t-test for the correlation r_τ to determine significance can be obtained by calculating

$$t = r_\tau \sqrt{\frac{n-2}{1-r_\tau^2}} \tag{7.50}$$

and determine if the observed t exceeds critical $t_{\alpha/2,n-2}$ as in a standard, two-sided t-test.

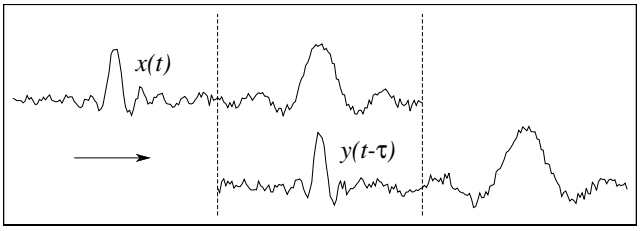


Figure 7.9: Cross-correlation between two separate time-series for an arbitrary lag.

Cross-correlation as defined here is most useful when the two signals have a common origin and time scale so that zero lag can be identified. While $\tau = 0$ always gives $r_\tau = 1$ for autocorrelation, r_0 may be zero for cross-correlation if one signal is delayed with respect to the other. After compiling the cross-correlation for all lags (positive and negative) we may find that the correlation is maximized for a particular lag τ_m. If the correlation is significantly nonzero we may draw the conclusion that there is a direct correlation between the two series, but this effect is *delayed* by some time $\Delta t \cdot \tau_m$. Examples of data pairs that may exhibit such cross-correlation include

1. Time series of the amount of water injected into a well and the intensity of seismicity. Such a cross-correlation demonstrates the importance of pore-pressure on the friction on faults. The increased pore-pressure reduces the effective normal stress on the fault and allows for more earthquakes to occur after accounting for the delaying effect of groundwater flow.

2. Glacial loading histories and land emergence since the end of the ice age. The viscous properties of the mantle retard the isostatic response and produce a delay between ice melt and land emergence.

3. Any process in which the input signal $x(t)$ is delayed and gives output $y(t)$. The optimal lag τ in the cross-correlation between x and y will tell us something about the process that caused the delay. This could be flow through a permeable medium, viscous response of the mantle, the inelastic response of the solid Earth to tidal forces, etc.

7.7 Geologic Correlation

The automated correlation of geological quantities quickly runs into trouble because cross-correlation techniques require a constant and common time scale for the two time series, which is often not the case. Depending on the particular nature of the data sets, systematic distortions such as variable sedimentation rates for sedimentary sequences in drill holes and variable seafloor spreading rates for magnetic anomalies will render cross-correlation problematic. Also, at other times the two series may have nominal values such as lithologic states and therefore cannot be assigned a numerical value. The extensions of autocorrelation and cross-correlation techniques to deal with nominal data are named *autoassociation* and *cross-association*, respectively. Correlation of such data can be enumerated by sliding the two data sets by one another and counting the number of matching states, e.g., sandstone at the same position as sandstone, and divide the result by the number of comparisons. Plotting this ratio r_τ as a function of match position τ may reveal a preferred location where the match is optimal.

S	S	C	C	L	L	C	L	S	S	S	C	S	C	C	L	L
			|		|			|	|				|		|	
			C	C	L	L	S	S	S	C	S	C	C	L	L	S	S

Table 7.4: Two lithologic sequences of sandstone (S), clay (C), and limestone (L) are compared using the cross-association technique. For the match position shown, there are six exact matches out of 14 possible, so $r_\tau = 6/14 = 0.43$.

Simple binomial probability theory may then be used to test if this ratio is significant or if it is what one can expect from two random sequences of the same composition.

7.8 Problems for Chapter 7

Problem 7.1. The stratigraphic column below represents the lithologic successions in a sequence taken from a delta plain where we find the lithologies sandstone (light blue), siltstone (light gray), clay (light green), and coal (black). Examine and record the transitions between lithologies every meter.

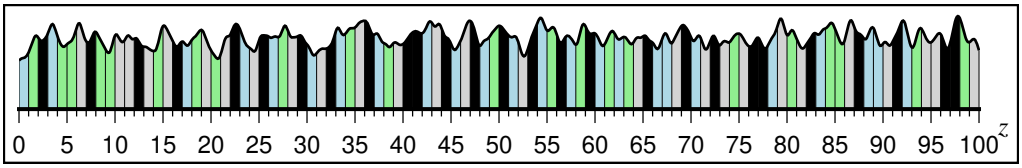


Figure 7.10: Observed stratigraphic section for Problem 7.1. Distances are in meters.

a) What is the transition frequency matrix, **A**, for this sequence?

b) Determine the fixed probability vector, **f**.

c) Evaluate the transition probability matrix, **P**.

d) At the 95% level of confidence, are the transitions random?

e) $\mathbf{S} = \mathbf{P} \cdot \mathbf{P}$ gives the second order Markov transition matrix. At the 95% level, are there significant second-order properties in the sequence?

Problem 7.2. Same questions as above, but this time using the transitions between just three lithologies: mudstone (light blue), siltstone (beige), and coal (black).

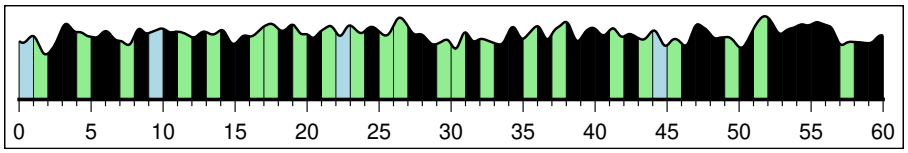


Figure 7.11: Observed stratigraphic section for Problem 7.2. Distances are in meters.

Problem 7.3. The file *embedded.txt* contains the embedded Markov chain transitions between four lithologies: A = shale with fossils, B = siltstone, C = sandstone, and D = coal. Determine if there is evidence for a first-order cyclicity at the 95% level of confidence using the test for embedded Markov chains.

Problem 7.4. The file *aso.txt* lists the years the Japanese volcano Aso has erupted during the period 1229–1962. Use a Kolmogorov-Smirnov test to determine (at the 95% level of confidence) whether the events are uniformly distributed over the time period. Plot the two cumulative distributions involved in the test.

Problem 7.5. The file *GK2007.txt* lists all magnetic reversals and the duration of each chron from the last 155 million year. Use a series of events test (at the 95% level of confidence) to determine whether the reversals are uniformly distributed over the time period.

Problem 7.6. The data set *limestone.txt* contains the thickness of successive limestone beds in the Lower Jurassic from a formation in Wales. Using the run test, is there a pattern in this sequence of thicknesses?

Problem 7.7. The 3-km long Vostok ice core from Antarctica resolves temperature variations relative to the present via oxygen isotopes. These data are given in table *vostok.txt*, which contains equidistant depths (in meter), the corresponding times (in year), and the relative change in temperature (in $^\circ$C).

a) Since the autocorrelation calculation requires an equidistant interval we must compute the autocorrelation of the temperature changes as a function of depth. At what lag > 0 is the autocorrelation maximized? What does this lag represent and how is it related to time?

b) Because of compaction, depth is not a good proxy for time, especially for the deeper (older) sections. To analyze the temporal periodicities we thus need an equidistant time-series. Use MATLAB's `spline` or other software to interpolate the data onto an equidistant interval in time ($\Delta t = 25$) and compute the autocorrelation of the resampled time-series. How do the two autocorrelations differ? What period would you now select for the dominant periodicity?

Problem 7.8. We will be revisiting Problem 6.3 so make sure you solve that problem first.

a) Determine the residual bathymetry $r(x)$ and calculate their first differences, $n(x)$, using MATLAB's `diff` operator, and plot these values versus distance.

b) Compute the autocorrelation of $n(x)$ for lags $\tau = 0$ through 5 (using MATLAB's `xcorr` function). The 99% confidence interval for white noise is known to be $\pm 3/\sqrt{n}$. For these lags, are the data compatible with a null hypothesis that states $n(x)$ may be considered white noise?

Problem 7.9. An engineer generates a two-pulse, noisy signal that he sends through a "black box" filtering operator. The experimental setup records the time (in seconds) and both the input and output magnitudes, reproduced in file *blackbox.txt*. The black box filter seems to both smooth the output and delay it in time. Use the cross-correlation technique to determine the lag induced by the filter.

Problem 7.10. Two stratigraphic sections (A and B) separated by a few hundred meters have been obtained (Figure 7.12). Use the cross-association technique to determine for what shift the second section B best fits the first section A. Plot your calculated match ratios r_τ versus the vertical offset of B relative to A and find the best fit. Draw your new tie-lines between the various layers.

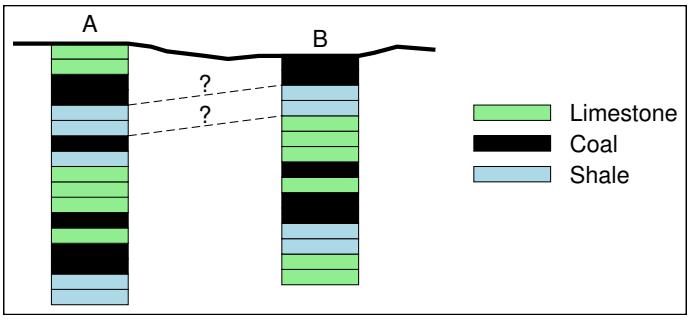


Figure 7.12: Observed stratigraphic sections for Problem 7.10.

Chapter 8

SPECTRAL ANALYSIS

"Science is spectral analysis. Art is light synthesis."

Karl Kraus, Writer

In Chapter 7 we were preoccupied with the topic of time-series analysis in the time-domain and learned a few things about the autocorrelation and cross-correlation techniques. In this chapter, we will take the different perspective of studying the *periodicities* present in a time series. At the heart of spectral analysis lies the notion of a signal's *frequency content*. This concept is utilized to decompose an observed signal into simpler components of known shape. Because many real observations in fact contain periodic components that fluctuate in a predictable way (e.g., yearly, monthly, daily), it is desirable to use periodic functions as the basic building blocks of the time-series. The most obvious choices are the trigonometric functions *sine* and *cosine*. The use of sines and cosines to approximate data and functions goes back to the early 1700s but was given mathematical rigor and extensive treatment by Joseph Fourier late in the 18th century. Fourier proved that any continuous, single-valued function could be represented by a series of sinusoids — today we know such series by the name *Fourier series*. Thus, spectral analysis involves finding the components of the Fourier series and interpreting the frequency content represented by the series. Spectral analysis also goes by other names, such as frequency analysis and harmonic analysis. Before we get into the details we must review some terminology and basic trigonometry.

8.1 Basic Terminology

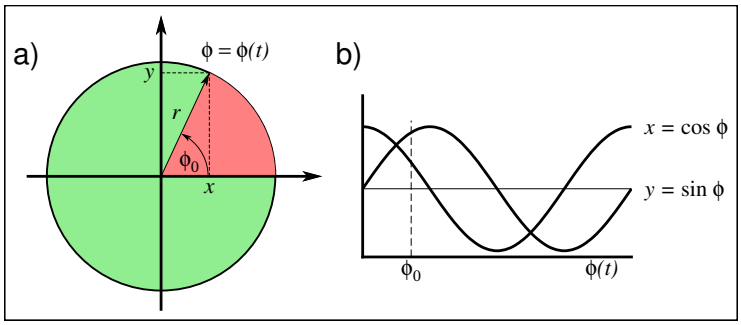


Figure 8.1: The periodic functions cosine and sine are defined as the x and y-components of the counter-clockwise spinning unit vector r as a function of the rotation angle ϕ. a) Spinning vector at a specific time t_0, yielding an angle ϕ_0 and the corresponding x and y values indicated by the dashed lines. b) Over time, these components trace the sine and cosine functions.

Consider a unit vector that rotates counterclockwise (Figure 8.1). The time it takes to complete one cycle is called the *period*, T. The y and x coordinates are then periodic functions of t and are given by the sine and the cosine,

respectively. The *radial frequency*, f, of the signal is the number of complete revolutions per second. Hence,

$$x(t) = \cos(2\pi ft) = \cos(\omega t) \quad y(t) = \sin(2\pi ft) = \sin(\omega t) \tag{8.1}$$

The period $T = 1/f$ has units of seconds per cycle. Instead of radial frequency we may use the *angular frequency*, $\omega = 2\pi f$, which has units of radians/sec. For spatial data, the period T corresponds to the *wavelength*, λ, and the angular frequency is referred to as the *wavenumber*, $k = 2\pi/\lambda$. The *amplitude*, A, of the signal is the length of the radial vector, r.

Instead of requiring that the sine curve go through zero at an even number of π, we can shift it horizontally by subtracting a constant ϕ from the argument.

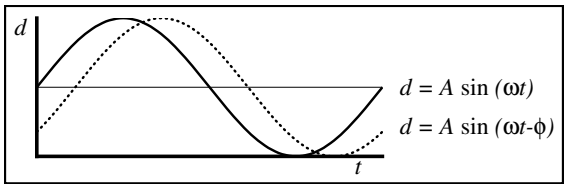


Figure 8.2: A phase-shifted sine curve (dotted line) is shifted along the t-axis.

The constant ϕ is called the *phase* of the signal (Figure 8.2). It is clear that the cosine and sine are out of phase by 90°. Let us assume that a particular time-series has one single periodic component with angular frequency ω. An example of such a series is shown in Figure 8.3, where we would need to find both A and ϕ. Unfortunately, while this model is linear in A it is *nonlinear* in ϕ. However, using the trigonometric identity for the cosine of a difference between two angles we find

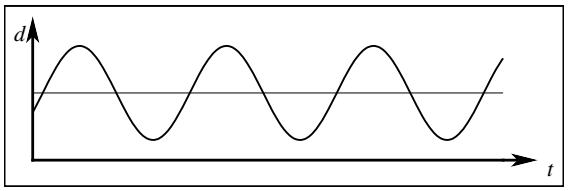


Figure 8.3: Sinusoid with arbitrary phase can be considered a sum of a sine and a cosine, both with zero phase.

$$d(t) = A\cos(\omega t - \phi) = A[\cos\phi\cos\omega t + \sin\phi\sin\omega t] = a\cos\omega t + b\sin\omega t, \tag{8.2}$$

which *is* linear in both a and b. Thus, instead of finding one amplitude and a phase we instead find the two amplitudes a and b for a cosine and sine pair, respectively, each with *no* phase shift. We may then easily recover the original parameters A and ϕ using

$$A = \sqrt{a^2 + b^2}, \quad \phi = \tan^{-1} b/a. \tag{8.3}$$

More often than not, the observed signal will contain many different sinusoids of different periods, phases, and amplitudes. We can use the subscript j to indicate the j'th component of the series. Perhaps a complete Fourier series for a signal $y(t)$ could therefore be written

$$d(t) = \sum_{j=0}^{\infty} a_j \cos\omega_j t + b_j \sin\omega_j t, \tag{8.4}$$

where ω_j represent the various angular frequency components? The a_j and b_j coefficients could then be found by standard least squares techniques. However, if we try to solve this system for many components we quickly run into computational problems. To avoid this problem we must look at *harmonics*.

8.1.1 Harmonics

The *fundamental* frequency of a signal has period T (scaled to 2π; see Figure 8.4) and reproduces a full cycle corresponding to the length of the data signal. Consequently,

$$f_F = 1/T, \quad \omega_F = 2\pi f_F = 2\pi/T. \tag{8.5}$$

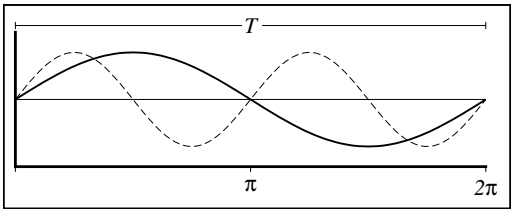


Figure 8.4: The fundamental frequency (solid line) has period T and typically represents the length of our data. Also shown is the second harmonic (dashed line).

The first harmonic is the sinusoid which makes two complete oscillations over the period T. However, because $f_1 = 2f_F$, this first harmonic sinusoid is usually called the *second harmonic* (though some people still refer to it as the first harmonic). Using that notation,

$$\begin{array}{ll} f_1 = f_F = 1/T & \omega_1 = 2\pi f_F = 2\pi/T \\ f_2 = 2f_F = 2/T & \omega_2 = 4\pi f_F = 4\pi/T \\ \vdots & \vdots \\ f_n = nf_F = n/T & \omega_n = 2n\pi f_F = 2n\pi/T \end{array} \tag{8.6}$$

Superposition of harmonics will always produce a new periodic function with period T.

8.1.2 Beats

Nearby frequency components can interact in interesting ways. Consider

$$d(t) = \cos\omega_1 t + \cos\omega_2 t, \tag{8.7}$$

where

$$\omega_1 = \omega + \delta\omega \quad \omega_2 = \omega - \delta\omega, \tag{8.8}$$

and $\delta\omega$ is small. Using trigonometric identities,

$$d(t) = 2\cos(\delta\omega t)\cos(\omega t). \tag{8.9}$$

Thus, the component $\cos(\omega t)$ (with $T = 2\pi/\omega$) has a slowly varying amplitude according to $\cos(\delta\omega t)$, which is the *modulation* function. This phenomenon is referred to as a *beat* (Figure 8.5). As $\delta\omega$ gets smaller, the period of the

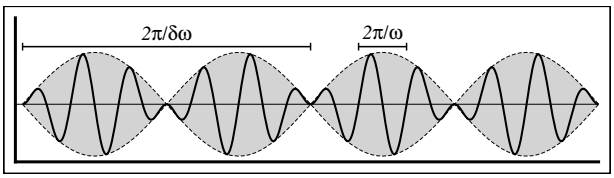


Figure 8.5: Graphical representation of a "beat" or amplitude modulation. Modulating the amplitude of a constant frequency carrier wave is the basis for AM radio, while FM broadcasts use a constant amplitude carrier wave and modulate the frequency instead.

beat curve gets longer and the phenomenon becomes more noticeable. In acoustics, the two frequencies ω_1 and ω_2 are often too high to hear but the beat is within an audible range. *Modulation* is the general phenomenon of a sinusoid with a varying amplitude (of any functional form).

8.2 Fitting the Fourier Series

Consider fitting a set of data, d_i, using a set of sine and cosines as the basis functions. At each observation time, t_i, our model predictions would be

$$\hat{d}_i = m_0 + m_1 \sin 2\pi t_i/T + m_2 \cos 2\pi t_i/T + m_3 \sin 4\pi t_i/T + m_4 \cos 4\pi t_i/T + \ldots, \tag{8.10}$$

where $T = n\Delta t$ if t_i are evenly spaced. We want to interpolate the values exactly, hence at our observations,

$$\begin{array}{l} m_0 + m_1 \sin 2\pi t_1/T + m_2 \cos 2\pi t_1/T + m_3 \sin 4\pi t_1/T + \cdots = d_1 \\ m_0 + m_1 \sin 2\pi t_2/T + m_2 \cos 2\pi t_2/T + m_3 \sin 4\pi t_2/T + \cdots = d_2 \\ \qquad\qquad \vdots \\ m_0 + m_1 \sin 2\pi t_n/T + m_2 \cos 2\pi t_n/T + m_3 \sin 4\pi t_n/T + \cdots = d_n \end{array}, \tag{8.11}$$

and this gives us n equations in n unknowns. Written as a matrix equation,

$$\begin{bmatrix} 1 & \sin 2\pi t_1/T & \cos 2\pi t_1/T & \sin 4\pi t_1/T & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \\ 1 & \sin 2\pi t_n/T & \cos 2\pi t_n/T & \sin 4\pi t_n/T & \dots \end{bmatrix} \times \begin{bmatrix} m_0 \\ \vdots \\ m_{n-1} \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}. \tag{8.12}$$

This standard matrix equation, $\mathbf{G} \cdot \mathbf{m} = \mathbf{d}$, can now be solved for the n coefficients of the sine and cosine series (i.e., the Fourier series) in the usual (matrix) way (i.e., $\mathbf{m} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d}$). However, because of orthogonality relationships between harmonics, the coefficients can be found analytically (this was fist shown by Lagrange in the 1800's using just the sine components over a range of x from 0 to π).

First, we rewrite the Fourier series as

$$\hat{d}_i = a_0 + \sum_{j=1}^{\leq n/2} \left[a_j \cos \frac{2\pi j t_i}{n\Delta t} + b_j \sin \frac{2\pi j t_i}{n\Delta t} \right], \quad i = 1, n, \tag{8.13}$$

where $\leq n/2$ means the largest whole integer resulting from the division $n/2$. Thus, the unknown vector $\mathbf{m}$ would now contain the renamed components

$$\mathbf{m}^T = \begin{bmatrix} a_0 & a_1 & \dots a_{\leq n/2} & b_1 & b_2 & \dots & b_{\leq n/2} \end{bmatrix}. \tag{8.14}$$

The number of data points n may be any integer value, but we will see it makes a minor difference if n is or is not divisible by two. If n is an even number, then for $j = n/2$ we find the two last terms in (8.13) to be

$$a_{n/2} \cos \frac{2\pi n t_i}{2n\Delta t} + b_{n/2} \sin \frac{2\pi n t_i}{2n\Delta t} = a_{n/2} \cos \frac{\pi t_i}{\Delta t} + b_{n/2} \sin \frac{\pi t_i}{\Delta t}. \tag{8.15}$$

Since $t_i = (i-1)\Delta t$, where $i = 1, \cdots, n$, we have

$$a_{n/2} \cos(i-1)\pi + b_{n/2} \sin(i-1)\pi = [-1]^{(i-1)}\, a_{n/2}. \tag{8.16}$$

This is true, since $\sin(i-1)\pi = 0$ for all i, hence the $b_{n/2}$ term drops out and we find

$$\mathbf{m}^T = \begin{bmatrix} a_0 & a_1 & \dots & a_{n/2} & b_1 & b_2 & \dots & b_{(n/2)-1} \end{bmatrix}, \tag{8.17}$$

which gives us $1 + (n/2) + (n/2) - 1 = n$ coefficients for n unknowns, and thus a solvable system results.

On the other hand, if n is *odd* then for $j < (n/2) = (n-1)/2$, both the sine and cosine terms remain and we have

$$\mathbf{m}^T = \begin{bmatrix} a_0 & a_1 & \dots & a_{(n-1)/2} & b_1 & b_2 & \dots & b_{(n-1)/2} \end{bmatrix}, \tag{8.18}$$

which again yields $1 + (n-1)/2 + (n-1)/2 = n$ coefficients for n unknowns. Since, for $j = 0$,

$$a_0 \cos 0 + b_0 \sin 0 = a_0, \tag{8.19}$$

and thus there is never a b_0 term. Using the convention $a_0 = \frac{a_0}{2}$ and $a_{n/2} = \frac{a_{n/2}}{2}$, the Fourier series may be written as

$$\hat{d}_i = \sum_{j=0}^{\leq n/2} \left[a_j \cos \frac{2\pi j t_i}{n\Delta t} + b_j \sin \frac{2\pi j t_i}{n\Delta t} \right], \quad i = 1, n, \text{ with } a_0 = \frac{a_0}{2}, \quad a_{n/2} = \frac{a_{n/2}}{2} \tag{8.20}$$

(the $a_0/2$ and $a_{n/2}/2$ convention is only a convenience that will become obvious later). The frequency

$$\omega_j = \frac{2\pi j}{n\Delta t} = \frac{2\pi j}{T} \tag{8.21}$$

is called the j'th *Fourier frequency*. Notice if n is odd, then $j < n/2$, so $\omega_j < \pi/\Delta t$ and hence $\pi/\Delta t$ is *not* a Fourier frequency, otherwise it *is* a Fourier frequency and the sine term is zero. Also, the highest frequency $f_{n/2} = (n/2)/(n\Delta t) = 1/(2\Delta t)$ is called the *Nyquist* frequency, which we will return to later.

Lagrange took advantage (in a brute force and laborious manner) of the following five relationships of harmonic components (which are easily shown using the corresponding integral relationships):

$$\sum_{i=1}^{n} \cos\omega_j t_i = \begin{cases} 0, & j \neq 0 \\ n, & j = 0 \end{cases}, \tag{8.22}$$

$$\sum_{i=1}^{n} \sin\omega_j t_i = 0, \tag{8.23}$$

$$\sum_{i=1}^{n} \cos\omega_j t_i \cos\omega_k t_i = \begin{cases} n/2, & j = k \neq 0, n/2 \\ n, & j = k = 0, n/2 \\ 0 & j \neq k \end{cases}, \tag{8.24}$$

$$\sum_{i=1}^{n} \sin\omega_j t_i \cos\omega_k t_i = 0, \tag{8.25}$$

and

$$\sum_{i=1}^{n} \sin\omega_j t_i \sin\omega_k t_i = \begin{cases} n/2, & j = k \\ 0 & j \neq k \end{cases}. \tag{8.26}$$

For a proof, consider the integral corresponding to (8.22):

$$\int_0^T \cos\omega_j t dt = \frac{1}{\omega_j}\int_0^{T\omega_j} \cos u du = \frac{1}{\omega_j}\sin u \Big|_0^{T\omega_j} = \frac{T}{2\pi j}\left(\sin 2\pi j - \sin 0\right) = 0\,(j \neq 0). \tag{8.27}$$

For $j = 0$,

$$\int_0^T \cos 0t dt = \int_0^T dt = T. \tag{8.28}$$

It is also easy to visualize these relationships. For instance, see Figure 8.6 for the a graphical proof of (8.25).

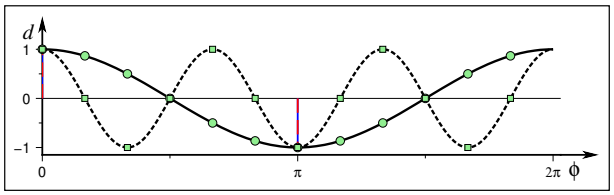


Figure 8.6: Orthogonality of two cosine harmonics drawn as solid (ω_1) and dashed (ω_3) lines, with green circles and squares as the hypothetical discrete samples. The products in (8.24) are represented by the product of the red (negative) and blue (positive) signed lengths. For this pair, you can see that for each red–blue line there is another of opposite orientation, thus canceling each other and yielding a final sum of zero.

Returning to the Fourier series, we have

$$\hat{d}_i = \sum_{j=0}^{\leq n/2} \left[a_j \cos\omega_j t_i + b_j \sin\omega_j t_i\right], \quad i = 1, n \tag{8.29}$$

and once again we want to minimize the misfit between data and model:

$$E = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[d_i - \hat{d}_i\right]^2 = \sum_{i=1}^{n} \left[d_i - \sum_{j=0}^{\leq \frac{n}{2}} \left(a_j \cos\omega_j t_i + b_j \sin\omega_j t_i\right)\right]^2. \tag{8.30}$$

Our unknowns are the n coefficients a_j and b_j. Taking the partial derivatives of E with respect to those parameters gives

$$\frac{\partial E}{\partial a_k} = 2\sum_{i=1}^{n}\left[d_i - \sum_{j=0}^{\leq n/2}(a_j\cos\omega_j t_i + b_j\sin\omega_j t_i)\right]\cos\omega_k t_i = 0, \quad k = 0,\ldots,\leq \frac{n}{2} \tag{8.31}$$

$$\frac{\partial E}{\partial b_k} = 2\sum_{i=1}^{n}\left[d_i - \sum_{j=1}^{<n/2}(a_j\cos\omega_j t_i + b_j\sin\omega_j t_i)\right]\sin\omega_k t_i = 0, \quad k = 1,\ldots,< \frac{n}{2}. \tag{8.32}$$

Eliminating the factor of 2 and rearranging, we obtain

$$\sum_{i=1}^{n} d_i\cos\omega_k t_i = \sum_{j=0}^{\leq n/2}\left[a_j\sum_{i=1}^{n}\cos\omega_j t_i\cos\omega_k t_i + b_j\sum_{i=1}^{n}\sin\omega_j t_i\cos\omega_k t_i\right], \tag{8.33}$$

$$\sum_{i=1}^{n} d_i\sin\omega_k t_i = \sum_{j=1}^{<n/2}\left[a_j\sum_{i=1}^{n}\cos\omega_j t_i\sin\omega_k t_i + b_j\sum_{i=1}^{n}\sin\omega_j t_i\sin\omega_k t_i\right]. \tag{8.34}$$

The five orthogonality relationships can now be employed. For $k = 0$, (8.33) becomes

$$\sum_{i=1}^{n} d_i\cos\omega_0 t_i = \sum_{j=0}^{\leq n/2}\left[a_j\sum_{i=1}^{n}\cos\omega_j t_i\cos\omega_0 t_i + b_j\sum_{i=1}^{n}\sin\omega_j t_i\cos\omega_0 t_i\right], \tag{8.35}$$

and since $\omega_0 = 0$, $\cos\omega_0 t_i$ is unity. Furthermore, we already determined that $b_0 = 0$, thus

$$\begin{aligned}\textstyle\sum_{i=1}^{n} d_i = a_0\textstyle\sum_{i=1}^{n}\cos\omega_0 t_i + a_1\sum_{i=1}^{n}\cos\omega_1 t_i + a_2\sum_{i=1}^{n}\cos\omega_2 t_i \\ + \cdots + b_1\textstyle\sum_{i=1}^{n}\sin\omega_1 t_i + b_2\sum_{i=1}^{n}\sin\omega_2 t_i + \cdots.\end{aligned} \tag{8.36}$$

Using the relationships (8.24) and (8.25), we find

$$\sum_{i=1}^{n} d_i = a_0 n + a_1 0 + a_2 0 + \cdots b_1 0 + b_2 0 + \cdots, \tag{8.37}$$

and with our convention $a_0 = a_0/2$ we obtain

$$\sum_{i=1}^{n} d_i = \frac{n}{2}a_0 \quad \Rightarrow \quad a_0 = \frac{2}{n}\sum_{i=1}^{n} d_i = 2\bar{d}, \tag{8.38}$$

i.e., the first coefficient a_0 is simply twice the mean of the data (a consequence of our convention for a_0 that includes the factor of 1/2). Using the same approach for $k = 1$, we first obtain

$$\begin{aligned}\sum_{i=1}^{n} d_i\cos\omega_1 t_i = a_0\sum_{i=1}^{n}\cos\omega_0 t_i\cos\omega_1 t_i + a_1\sum_{i=1}^{n}\cos\omega_1 t_i\cos\omega_1 t_i + a_2\sum_{i=1}^{n}\cos\omega_2 t_i\cos\omega_1 t_i + \cdots \\ + b_1\sum_{i=1}^{n}\sin\omega_1 t_i\cos\omega_1 t_i + b_2\sum_{i=1}^{n}\sin\omega_2 t_i\cos\omega_1 t_i + \cdots.\end{aligned} \tag{8.39}$$

Using the orthogonality relationships, we find

$$\sum_{i=1}^{n} d_i\cos\omega_1 t_i = a_0 0 + a_1\frac{n}{2} + a_2 0 + \cdots + b_1 0 + b_2 0 + \cdots. \tag{8.40}$$

Therefore,

$$a_1 = \frac{2}{n}\sum_{i=1}^{n} d_i\cos\omega_1 t_i. \tag{8.41}$$

Since the orthogonality relationships are the same for all $k = 1,2,\cdots < n/2$ as they were for $k = 1$ in (8.41), we find

$$a_k = \frac{2}{n}\sum_{i=1}^{n} d_i\cos\omega_k t_i, \quad k = 0,1,\cdots,< \frac{n}{2}. \tag{8.42}$$

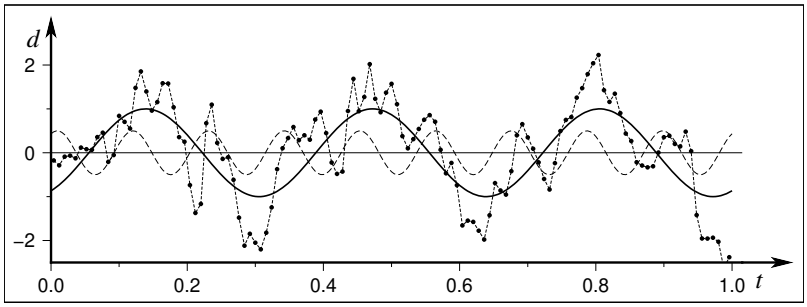


Figure 8.7: Example of a data set (thin line with connected dots) and two fitted Fourier components. Here, we show the least-squares solution for the two harmonics ω_3 (solid line) and ω_9 (dashed line). When these and all other harmonics are evaluated they will sum to equal the original data set.

Finally, we need to look at the last case, $k = n/2$ (for even n). For $k = n/2$,

$$\begin{aligned}\sum_{i=1}^{n} d_i \cos\omega_{n/2} t_i = a_0 \sum_{i=1}^{n} \cos\omega_0 t_i \cos\omega_{n/2} t_i + a_1 \sum_{i=1}^{n} \cos\omega_1 t_i \cos\omega_{n/2} t_i + a_2 \sum_{i=1}^{n} \cos\omega_2 t_i \cos\omega_{n/2} t_i + \cdots \\ + b_1 \sum_{i=1}^{n} \sin\omega_1 t_i \cos\omega_{n/2} t_i + b_2 \sum_{i=1}^{n} \sin\omega_2 t_i \cos\omega_{n/2} t_i + \cdots .\end{aligned} \tag{8.43}$$

Again, using the orthogonality relationships,

$$\sum_{i=1}^{n} d_i \cos\omega_{n/2} t_i = a_0 0 + a_1 0 + \cdots + n a_{n/2} + \cdots + b_1 0 + b_2 0 + \cdots . \tag{8.44}$$

Since we required $a_{n/2} = a_{n/2}/2$, we find

$$a_{n/2} = \frac{2}{n} \sum_{i=1}^{n} d_i \cos\omega_{n/2} t_i. \tag{8.45}$$

Therefore, with the convention that $a_0 = a_0/2$ and $a_{n/2} = a_{n/2}/2$, and interchanging the dummy subscripts k and j, the a_j can be defined as

$$\boxed{a_j = \frac{2}{n} \sum_{i=1}^{n} d_i \cos\omega_j t_i, \quad 0 \le j \le \frac{n}{2}.} \tag{8.46}$$

We now turn our attention to the other half of the normal equations (8.34) involving the $\sin\omega_j t_i$ terms. Previously, we have shown that $b_0 = b_{n/2} = 0$ so we will only need to look at one case $(k = 1)$ and generalize the result for all k. We find

$$\begin{aligned}\sum_{i=1}^{n} d_i \sin\omega_1 t_i = a_0 \sum_{i=1}^{n} \cos\omega_0 t_i \sin\omega_1 t_i + a_1 \sum_{i=1}^{n} \cos\omega_1 t_i \sin\omega_1 t_i + a_2 \sum_{i=1}^{n} \cos\omega_2 t_i \sin\omega_1 t_i + \cdots \\ + b_1 \sum_{i=1}^{n} \sin\omega_1 t_i \sin\omega_1 t_1 + b_2 \sum_{i=1}^{n} \sin\omega_2 t_i \sin\omega_1 t_i + \cdots .\end{aligned} \tag{8.47}$$

Thus,

$$\sum_{i=1}^{n} d_i \sin\omega_1 t_i = a_0 0 + a_1 0 + \cdots + b_1 \frac{n}{2} + b_2 0 + \cdots , \tag{8.48}$$

and

$$b_1 = \frac{2}{n} \sum_{i=1}^{n} d_i \sin\omega_1 t_i. \tag{8.49}$$

Since the orthogonality relationships hold for all k, we again interchange k and j and find

$$\boxed{b_j = \frac{2}{n} \sum_{i=1}^{n} d_i \sin\omega_j t_i, \quad 0 < j < \frac{n}{2}.} \tag{8.50}$$

The formulae for a_j and b_j are called the *Discrete Cosine and Sine Transforms* and combined they define the *Discrete Fourier Transform*. Figure 8.7 shows an example of a data set and the determination of two Fourier components.

Example 8–1. Consider the time-series $d = [1\ 0\ \ -2\ \ -1\ \ 1\ 2\ 1\ 0.5]^T$ with $\Delta t = 1, n = 8, T = 8$. The Fourier frequencies are therefore

$$\begin{array}{lll} \omega_j = 2\pi j/T, & \omega_0 = 0, & \omega_1 = \pi/4, \\ \omega_2 = \pi/2, & \omega_3 = 3\pi/4, & \omega_4 = \pi. \end{array} \tag{8.51}$$

Solving for the coefficients, we find

$$\begin{array}{lllll} a_0 = 0.3125, & a_1 = -0.0884, & a_2 = 0.7508, & a_3 = 0.0804, & a_4 = -0.0625, \\ b_1 = -1.3687, & b_2 = 0.625, & b_3 = 0.1313. & & \end{array} \tag{8.52}$$

Thus, our Fourier series is

$$\begin{aligned} d(t) \quad = \quad & 0.3125 - 0.0884\cos\frac{\pi}{4}t - 1.3687\sin\frac{\pi}{4}t + 0.7508\cos\frac{\pi}{2}t \\ & +0.625\sin\frac{\pi}{2}t + 0.0804\cos\frac{3\pi}{4}t + 0.1313\sin\frac{3\pi}{4}t - 0.0625\cos\pi t \end{aligned} \tag{8.53}$$

Note that all the terms have periods that are multiples of the fundamental frequency; hence the Fourier representation must itself be periodic with the same fundamental period T. We will later see that the assumption of periodic data may have grave consequences for our coefficients.

8.2.1 The power of orthogonality

Let us pause and lament the demise of our dear friend from Chapter 5, the design matrix $\mathbf{G}$. What just happen to it in our analysis? Well, recall equations (5.139) and (5.140), both overflowing with dot-products of the basis vectors $\mathbf{g}_j$. Yet, with our choice of harmonics we found that these basis vectors were in fact orthogonal and thus their dot-products yielded zero except along the matrix diagonal (where we got the simple constant $n/2$). With $\mathbf{G}$ being diagonal, the formidable $[\mathbf{G}^T\mathbf{G}]^{-1}$ collapsed to the identity matrix $\mathbf{I}$ scaled by $2/n$, as we just witnessed. Now *that* is the power of orthogonal functions (of which sine and cosine are just two possibilities) and why they are so widely used in data analysis as well as for modeling in the physical sciences.

8.3 The Periodogram

We determined that the Fourier series expansion of our observed time-series d_i could be written

$$\hat{d}_i = \sum_{j=0}^{\leq n/2} [a_j \cos\omega_j t_i + b_j \sin\omega_j t_i]. \tag{8.54}$$

Remember that (8.2) started out by trying to fit a cosine of arbitrary amplitude A_j and phase ϕ_j, but that we could rewrite this single term as a sum of a cosine and sine components with different amplitudes and zero phases. We found

$$a_j = A_j \cos\phi_j, \quad b_j = A_j \sin\ \phi_j. \tag{8.55}$$

From these expressions we readily find a component's full amplitude and phase. Dividing the b_j by a_j gives

$$\tan\phi_j = b_j/a_j \quad \Rightarrow \quad \phi_j = \tan^{-1}(b_j/a_j). \tag{8.56}$$

Squaring a_j and b_j and adding them gives

$$A_j^2 = a_j^2 + b_j^2. \tag{8.57}$$

The *periodogram* is constructed by plotting A_j^2 versus j, f_j, ω_j, or P_j. While often called the *power spectrum*, it is strictly speaking a raw, discrete periodogram. The true spectrum is a smoothed periodogram showing frequency components of statistical regularity. However, the periodogram is the most common form of output of a Fourier transform. Figure 8.8 shows the periodogram for the function

$$d(t) = \frac{1}{2}\cos\omega_1 t + \frac{3}{4}\cos\omega_2 t + \frac{1}{2}\sin\omega_3 t + \frac{1}{4}\cos\omega_3 t + \frac{1}{3}\cos\omega_4 t + \frac{1}{5}\sin\omega_4 t + \frac{1}{3}\sin\omega_6 t - \frac{3}{5}. \tag{8.58}$$

Let us look, for a moment, at the variance of the time series expansion. Recall, the variance is given by

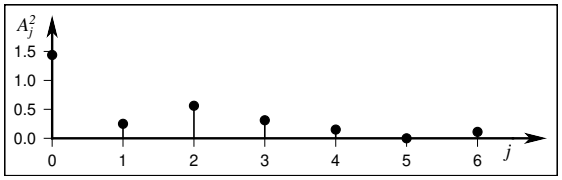


Figure 8.8: Raw periodogram of the function given in (8.58). The peak corresponds to the $A_0^2 = a_0^2$ term defined to be twice the mean (-0.6) squared.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\hat{d}_i - \bar{d})^2. \tag{8.59}$$

We shall write the Fourier series as

$$\hat{d}_i = \bar{d} + \sum_{j=1}^{\leq \frac{n}{2}} \left(a_j \cos\omega_j t_i + b_j \sin\omega_j t_i\right), \tag{8.60}$$

by pulling the constant (mean) term out separately. Since the two means cancel, we find

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left\{\left[\sum_{j=1}^{\leq \frac{n}{2}}\left(a_j\cos\omega_j t_i + b_j \sin\omega_j t_i\right)\right]\left[\sum_{q=1}^{\leq \frac{n}{2}}\left(a_q\cos\omega_q t_i + b_q \sin\omega_q t_i\right)\right]\right\}. \tag{8.61}$$

Also recall that, because of orthogonality, all the cross terms $(q \neq j)$ resulting from the full expansion of the two squared expressions will be zero when summed over i, while the remaining terms will sum to $n/2$ (since $j, q > 0$). Hence, we are left with

$$s^2 = \frac{n}{2(n-1)}\sum_{j=1}^{\leq \frac{n}{2}}(a_j^2 + b_j^2) \sim \frac{1}{2}\sum_{j=1}^{\leq \frac{n}{2}} A_j^2. \tag{8.62}$$

Therefore, the power spectrum (periodogram) of $(a_j^2 + b_j^2)$ versus ω_j is a plot showing the contribution of individual frequency components to the total variance of the signal. For this reason, the power spectrum is often called the variance spectrum. However, most of the time it is simply called "the spectrum." Hence, the Fourier transform converts a signal from the time domain to the frequency domain (or wavenumber domain), where the signal can be viewed in terms of the contribution of the different frequency components of which it is made. The phase spectrum (ϕ_j versus ω_j) shows the relative phase of each frequency component. In general, phase spectra are more difficult to interpret than amplitude (or power) spectra.

8.3.1 Aliasing of higher frequencies

We mentioned before that the highest frequency (or shortest period, or wavelength) that can be estimated from the data is called the Nyquist frequency (or period, or wavelength), given by

$$f_N = f_{n/2} = \frac{1}{2\Delta t}, \quad \omega_N = 2\pi f_N = \frac{\pi}{\Delta t} \quad P_{n/2} = 2\Delta t. \tag{8.63}$$

Higher frequencies, whose wavelengths are less than twice the spacing between sample points *cannot be detected*. However, when we sample a signal every Δt and the original signal has higher frequencies than $f_{n/2}$, we introduce

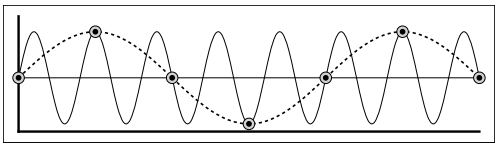


Figure 8.9: Aliasing: A short-wavelength signal that is not sampled at the Nyquist frequency or higher will instead appear as a longer-wavelength component that does not exist in the actual data.

aliasing. Aliasing means that some frequencies will leak power into other frequencies. This concept is readily seen by sampling a high-frequency signal at a spacing larger than the Nyquist interval.

Sampling of the high-frequency signal actually results in a longer-period signal (Figure 8.9). When Clint Eastwood's wagon wheels seem to spin backwards in an old Western movie — that's aliasing: The 24 pictures/sec rate is simply too slow to capture the faster rotation of the wheels.

8.3.2 Significance of a spectral peak

In some applications we may be interested in testing whether a particular component is dominant or if its larger amplitude is due to chance. The statistician R. A. Fisher devised a test that calculates the probability that a spectral peak s_j^2 will exceed the value σ_j^2 of a hypothetical time series composed of independent random points. We must evaluate the ratio of the variance contributed by the maximum peak to the entire data variance:

$$g = \frac{s_j^2}{2s^2}, \tag{8.64}$$

where s_j^2 is the largest peak in the periodogram (we divide by two to get its variance contribution) and s^2 is the variance of the entire series. For a prescribed confidence level, α, the critical value that we wish to compare to our observed g is

$$g_{\alpha,k} \approx 1 - \exp\left(\frac{\ln\alpha - \ln k}{k-1}\right), \tag{8.65}$$

with $k = n/2$ (for even n) or $k = (n-1)/2$ (for odd n). Should our observed g (obtained via 8.64) exceed this critical value we decide that the dominant component is real and reflects a true characteristic of the phenomenon we are observing. Otherwise, s_j^2 may be large simply by chance.

8.3.3 Estimating the continuous spectrum

The power spectrum or periodogram obtained from the Fourier coefficients is discrete, yet we do not expect the power at frequency ω_j to equal the underlying continuous $P(\omega)$ at exactly ω_j, since the discrete spectrum must necessarily represent some average value of power at all frequencies between ω_{j-1} and ω_{j+1}. In other words, the computed power at ω_j also represents the power from nearby frequencies not among the chosen harmonic frequencies ω_j. Furthermore, the uncertainty in any individual estimate p_j^2 is very large; in fact, it is equal to $\pm p_j^2$ itself.

Can we improve (i.e., reduce) the uncertainties in p_j^2 by using more data points or sample the data more frequently? The unpleasant answer is that the periodogram estimates do not become more accurate at all! The reason for this is that adding more points simply produces power estimates at a greater number of frequencies ω_j. The only way to reduce the uncertainty in the power estimates is to smooth the periodogram over nearby discrete frequencies. This can be achieved in one of two ways:

1. Use a time-series that is M times longer (so $f_1' = f_1/M$) and *sum* the M power estimates p_k^2 straddling each original ω_j frequency to obtain a smooth estimate $p_j^2 = \sum p_k^2$.

2. Split the original data into M smaller series, find the p_j^2 for each series, and take the *mean* of the M estimates for the same j (i.e., the same frequency).

In both cases the variance of the power spectrum estimates drop by a factor of M, i.e., $s_j^2 = p_j^2/M$. The exact way the smoothing is achieved may vary among analysts. Several different types of weights or spectral *windows* have

been proposed, but they are all relatively similar. These windows arose because, historically, the power spectrum was estimated by taking the Fourier transform of the *autocorrelation* of the data; hence many windows operated in the lag-domain. The introduction of the Fast Fourier Transform made the FFT the fastest way to obtain the spectrum, which then is simply smoothed over nearby frequencies. The FFT is a very rapid algorithm for doing a discrete Fourier transform, especially if n is a power of 2. It can be shown that one can always split the discrete transform into the sum of two discrete, scaled transforms of subsets of the data. Applying this result recursively, we eventually end up with a sum of transforms of data sets with one entry, whose transform equals itself. While mathematically equivalent, there is a huge difference computationally: While the discrete Fourier transform's execution time is proportional to n^2, the FFT only takes $n \cdot \log(n)$. For a data set of 10^6 points, the speed-up is a factor of $> 75,000$.

By doing a Fourier Analysis, we have transformed our data from one domain (time or space) to another (frequency or wavenumber). A physical analogy is the transformation of light sent through a triangular prism. White light is composed of many frequencies, and the prism acts as a frequency analyzer that separates the various frequency components, here represented by colors. Each color band is separated from its neighbor by an amount proportional to their difference in wavelength, and the intensity of each band reflects the amplitude of that component in the white light. We know that by examining the spectrum we can learn much about the composition and temperature of the source and the material the light passed through. Similarly, examining the power spectra of other processes may tell us something about them that may not be apparent in the time domain. Consequently, spectral analysis remains one of the most powerful techniques we have for examining temporal or spatial sequences.

8.3.4 First-Order Spectrum Interpretation

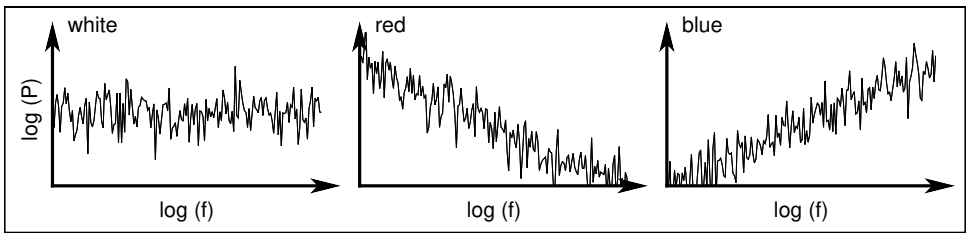


Figure 8.10: Simplified representations of typical spectra that are called "white" (left; equal power at all frequencies), "red" (middle; power falling off with increasing frequency), and "blue" (right; power increasing with frequency).

Per Section 8.3, the raw power spectrum, or *periodogram*, is obtained by plotting the squared amplitude A_j^2 versus frequency. Often, a spectrum will fall into one of three categories (see Figure 8.10):

white: This is a spectrum that shows little or no amplitude variation with frequency. Random values such as independent samples drawn from a normal distribution will have a white spectrum.

red: This spectrum is dominated by long-wavelength (low-frequency) signals, with the spectrum tapering off for higher frequencies. This is very common behavior in observed data, such as topography and potential fields (gravity, magnetics). It may also be indicative of data that represent an integrated phenomenon.

blue: This spectrum is dominated by short-wavelength (high-frequency) signal, with the spectrum tapering off for lower frequencies. Data that depend on derivatives, such as slopes and curvatures, might behave this way, being higher-order derivatives of a red-spectrum topography signal.

One reason for the prevalence of red or blue spectra for natural phenomena can be understood if we consider what effect a temporal derivative (e.g., d/dt) has in the frequency domain. Given that the Fourier series representation of data can be written

$$\hat{d}(t) = \sum_{j=0}^{\leq n/2} A_j \cos\left(\omega_j t - \phi_j\right), \tag{8.66}$$

taking the derivative yields

$$\frac{d}{dt}\hat{d}(t) = \sum_{j=0}^{\leq n/2} -\omega_j \cdot A_j \sin\left(\omega_j t - \phi_j\right), \tag{8.67}$$

In effect, we multiply each Fourier amplitude by its corresponding frequency, hence amplitudes at higher frequencies are preferentially enhanced while those at lower frequencies are attenuated. This scaling will make the spectrum more "blue". By analogy, integration in the temporal domain has the effect of *dividing* the spectrum by the frequency, conversely "reddening" the spectrum. This frequency effect is what we allude to when we say that taking a derivative typically make data noisier (it amplifies the short-wavelength or high-frequency components in the data) while integration tends to make data smoother by attenuating the same components. Of course, these statements assume that the uncertainties in the data are mostly at high frequencies, but some data have more uncertainty at low frequencies, in which case the situation is reversed. These simple considerations may be useful when interpreting your observed spectra. Finally, note that the derivative also introduces a phase change of $\pi/2$ (90°) since the cosine and the negative sine are shifted by 90°. A second multiplication (i.e., for a second-derivative result) leads to a 180° change in phase since we are essentially multiplying by -1 (this does not affect power, which is proportional to amplitude squared).

8.4 Convolution

Convolution represents one of the most fundamental operations of time series analysis and is one of the most physically meaningful. Consider the passage of a signal through a linear filter, where the filter (a "black box") will modify a signal passing through it (Figure 8.11). For instance, it may

1. Amplify, attenuate or delay the signal.
2. Modify or eliminate specific frequency components.

Figure 8.11: Example of convolution between an input signal and a filter.

Consider the propagation of a seismic pulse through the upper layers of the Earth's crust, as illustrated in Figure 8.12. The generated pulse may be sharp and thus have high frequencies, yet the recorded signal that traveled through the crust may be much smoother and include repeating signals that reflect internal boundaries.

Figure 8.12: Convolving a seismic pulse with the Earth gives a seismic trace that may reflect changing properties of the Earth with depth.

Convolution is this process of linearly modifying one signal using another signal. In Figure 8.12 we convolved the seismic pulse with the "Earth filter" to produce the observed returned seismogram. Symbolically, we write the convolution of a signal $d(t)$ by a filter $p(t)$ as the integral

$$h(t) = d(t) * p(t) = \int_{-\infty}^{+\infty} d(u) \cdot p(t-u) du, \tag{8.68}$$

where $*$ represents the convolution operator. *Deconvolution*, or *inverse filtering*, is the process of unscrambling the convolved signal to determine the nature of the filter *or* the nature of the input signal. Consider these two cases:

1. If we knew the exact shape of our seismic pulse $d(t)$ and seismic signal received, $h(t)$, we could deconvolve the data with the pulse to determine the (filtering) properties of the upper layers of the Earth through which the pulse passed (i.e., $p(t) = d^{-1}(t) * h(t)$).

2. If we wanted to determine the exact shape of our pulse $d(t)$, we could pass it through a known filter $p(t)$ and deconvolve the output with the shape of the filter (i.e., $d(t) = p^{-1}(t) * h(t)$).

The hard work here is to determine the inverse functions $d^{-1}(t)$ or $p^{-1}(t)$, which is akin to matrix inversion. Other examples of convolution include:

1. Filtering data — using running means, weighted means, removing specific frequency components, etc.
2. Recording a phenomenon with an instrument that responds slower than the rate at which the phenomenon changes, or which produces a weighted mean over a narrow interval of time, or which has lower resolving power than the phenomenon requires.
3. Conduction and convection of heat.
4. Deformation and the resulting gravity anomalies caused by the flexural response of the lithosphere to a volcano.

Convolution is most easily understood by examining its effect on discrete functions. First, consider the discrete impulse $d(t)$ sent through the filter $p(t)$, as illustrated in Figure 8.13:

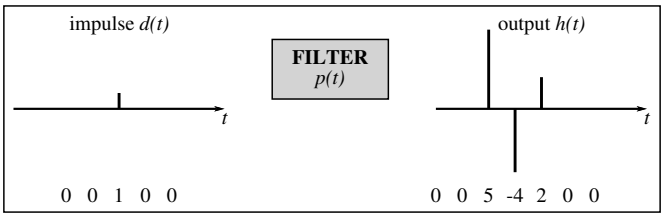


Figure 8.13: A filter's impulse response is obtained by sending an impulse $d(t)$ through the filter $p(t)$.

The output $h(t)$ from the filter is known as the *impulse response function* since it represents the response of the filter to an impulse, $d(t)$. It represents a fundamental property of the filter $p(t)$. Next, consider a more complicated input signal convolved with the filter, as shown in Figure 8.14:

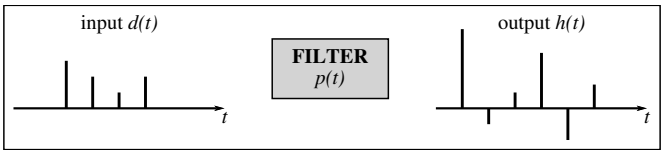


Figure 8.14: Filtering seen as a convolution.

Since the filter is linear, we may think of the input as a series of individual impulses. The output is thus the sum of several impulse responses scaled by their amplitudes and shifted in time. Calculating convolutions is a lot like calculating cross-correlations, except that the second time-series must be reversed. Consider the two signals as finite sequences on separate strips of paper (Figure 8.15).

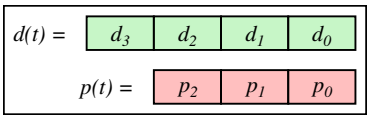


Figure 8.15: Graphical representation of a convolution. We write the discrete values of $d(t)$ and $p(t)$ on two separate strips of paper.

We obtain the zero lag output by aligning the paper strips as shown in Figure 8.16, after reversing the red strip.

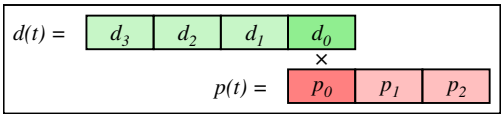


Figure 8.16: Convolution, zero lag. Reverse one strip and arrange them to yield a single overlap.

The zero lag result h_0 is thus simply $d_0 \cdot p_0$. Moving on, the first lag results from the alignment shown in Figure 8.17.

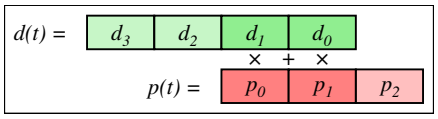


Figure 8.17: Convolution, first lag. We shift one strip by one to increase the overlap.

This simple process is repeated, and for each lag k we evaluate h_k as the sum of the products of the overlapping signal values. This is a graphic (or mechanical) representation of the discrete convolution equation (compare this operation to the integral in 8.68). Consider the convolution of the two functions shown in Figure 8.18.

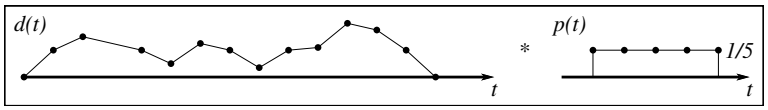


Figure 8.18: Moving averages is obtained by the convolution of data with a rectangular function of unit area.

If we look at this convolution with the moving strips of paper approach, we get the setup illustrated in Figure 8.19:

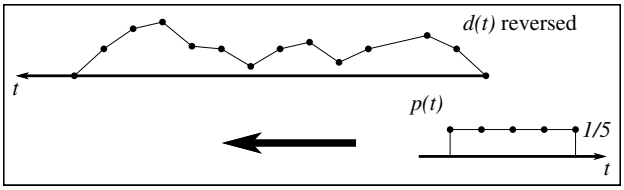


Figure 8.19: The mechanics of convolutions, this time without the paper strips.

Given the simple nature of $p(t)$, we can estimate the values of h_k directly:

$$\begin{aligned} h_0 &= d_0/5 \\ h_1 &= \tfrac{1}{5}(d_0+d_1) \\ &\vdots \\ h_4 &= \tfrac{1}{5}(d_0+d_1+d_2+d_3+d_4) \\ h_5 &= \tfrac{1}{5}(d_1+d_2+d_3+d_4+d_5) \\ &\vdots \\ h_{18} &= d_{14}/5 \end{aligned} \tag{8.69}$$

This is simply a five-point running (or moving) average of $d(t)$, and the result is shown in Figure 8.20. An n-point average would be the result if $p(t)$ consisted of n points, each with a value of $1/n$.

8.4.1 Convolution theorem

Although not shown here, it can be proven that a convolution of two functions $p(t)$ and $d(t)$ in the time-domain is equivalent to the product of $P(f)$ and $D(f)$ in the frequency domain (here, uppercase letters indicate the Fourier

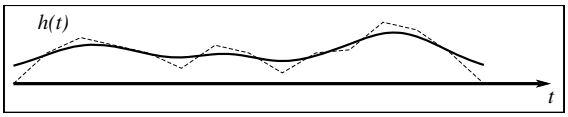


Figure 8.20: The final result of the convolution is a smoothed data set since any short-wavelength signal will be greatly attenuated.

transforms of the lowercase, time-domain functions). The converse is also true, thus

$$\begin{aligned} p(t) * d(t) &= h(t) \quad \leftrightarrow \quad P(f) \cdot D(f) = H(f), \\ p(t) \cdot d(t) &= z(t) \quad \leftrightarrow \quad P(f) * D(f) = Z(f). \end{aligned} \tag{8.70}$$

Because convolution is a slow calculation it is often advantageous to transform our data from one domain to the other, perform the simpler multiplication, and transform the data back to the original domain. The availability of *fast Fourier transforms* (FFTs) makes this approach practical.

8.5 Sampling Theory

The *sampling theorem* states that if a function is *band-limited* (i.e., the transform is zero for all radial frequencies $f > f_N$), then the continuous function $d(t)$ can be uniquely determined from knowledge of its sampled values given a sampling interval $\Delta t \leq 1/(2f_N)$. From distribution theory, we have

$$d_t = \sum_{j=-\infty}^{+\infty} d(t)\delta(t - j\Delta t) = \sum_{j=-\infty}^{\infty} d(j\Delta t)\delta(t - j\Delta t) = d(t) \cdot \Delta(t), \tag{8.71}$$

where

$$\Delta(t) = \sum_{j=-\infty}^{+\infty} \delta(t - j\Delta t) \tag{8.72}$$

is the sampling or "comb" function in the time domain (Figure 8.21). Thus, d_t is the continuous function $d(t)$ sampled at the discrete times $j\Delta t$. Consequently, it is true that the original signal $d(t)$ can be reconstructed exactly from its sampled values d_t via the *Whittaker-Shannon* interpolation formula

$$d(t) = \sum_{j=-\infty}^{+\infty} d_j \,\mathrm{sinc}\left(\frac{t - j\Delta t}{\Delta t}\right), \tag{8.73}$$

where $d_j = d(j\Delta t)$ are the sampled data values and the sinc function is defined as

$$\mathrm{sinc}(x) = \frac{\sin \pi x}{\pi x}. \tag{8.74}$$

Recall that the multiplication of two functions in the time domain is equivalent to the convolution of the their Fourier transforms in the frequency domain, hence

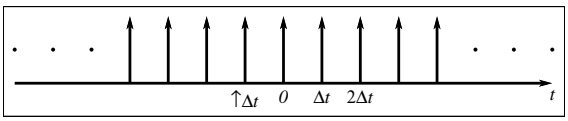


Figure 8.21: The sampling or "comb" function, $\Delta(t)$, represents mathematically what we do when we sample a continuous phenomenon $d(t)$ at discrete, equidistantly spaced times.

$$d(t) \cdot \Delta(t) \leftrightarrow D(f) * \Delta(f). \tag{8.75}$$

The time-domain expression is visualized in Figure 8.22.

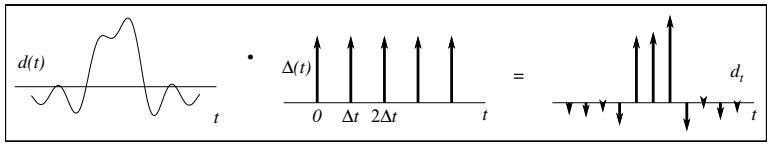


Figure 8.22: Sampling equals multiplication of a continuous signal $d(t)$ with a comb function $\Delta(t)$ in the time-domain.

The transformed function $\Delta(f)$ can be shown to be a series of impulses as well (Figure 8.23).

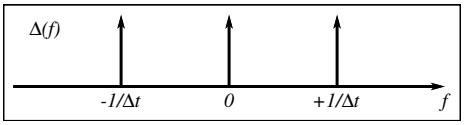


Figure 8.23: The Fourier transform of the comb function, $\Delta(t)$, is another comb function, $\Delta(f)$, with a spacing of $1/\Delta t$ between impulses.

In the frequency domain, $d(t)$ is represented as $D(f)$ and illustrated in Figure 8.24. We note that while the time-domain comb function $\Delta(t)$ is a series of impulses spaced every Δt, the frequency-domain comb function $\Delta(f)$ is also a series of impulses, but spaced every $1/\Delta t$. The time and frequency domain spacings of the comb functions are thus reciprocal: A finer sampling interval leads to a larger distance between the impulses in the frequency domain.

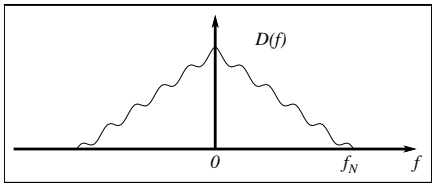


Figure 8.24: The Fourier transform of our continuous phenomenon, $d(t)$. We assume it is band-limited so that the transform goes to zero beyond the highest frequency, f_N.

Given $D(f)$ and $\Delta(f), D(f) * \Delta(f)$ is schematically shown in Figure 8.25.

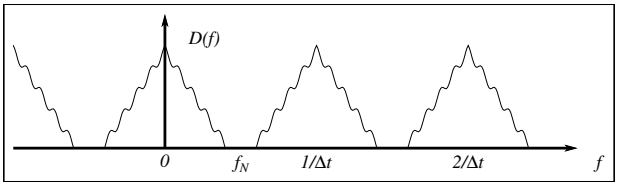


Figure 8.25: Replication of the transform, $D(f)$, due to its convolution with the comb function, $\Delta(f)$.

If the impulses in $\Delta(f)$ are spaced closer than $1/\Delta t$ then there will be some overlap between the $D(f)$ replicas that are centered at the location of each impulse (see Figure 8.26).

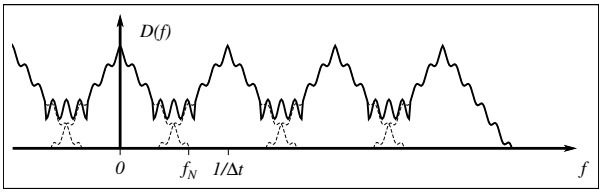


Figure 8.26: Aliasing in the frequency domain occurs when the sampling interval Δt is too large.

This overlap introduces *aliasing* (which we shall discuss more later). To prevent aliasing, we must ensure $\Delta t \leq 1/(2f_N)$, where f_N is the highest (radial) frequency component present in the time series. As mentioned earlier, we call f_N the *Nyquist frequency* and the Nyquist sampling interval is $\Delta t = 1/(2f_N)$, hence $f_N = 1/(2\Delta t)$. As long as we follow the sampling theorem and select $\Delta t \leq 1/(2f_N)$, with f_N being the highest frequency component, there will be no spectral overlap in $D(f) * \Delta(f)$ and we will be able to recover $D(f)$ completely. Therefore (and to prove the sampling theorem) we recover $D(f)$ by truncating the signal:

$$D(f) = [D(f) * \Delta(f)] \cdot H(f), \tag{8.76}$$

which is illustrated in Figure 8.27 as a multiplication of the replicating spectrum with a *gate* function, $H(f)$.

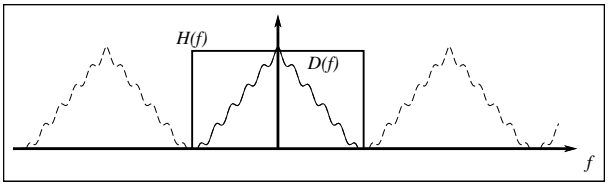


Figure 8.27: Truncation of the Fourier spectrum via multiplication with a rectangular gate function, $H(f)$.

8.5.1 Aliasing, again

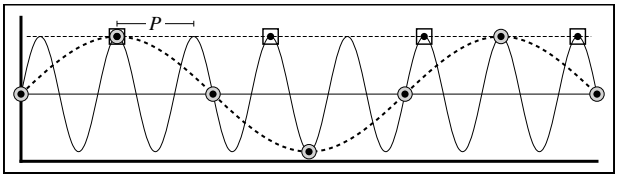


Figure 8.28: Aliasing as seen in the time domain. Thin line shows a phenomenon with period P. The circles and heavy dashed line show the signal obtained using a sampling rate of $1.25P$, while the squares and dashed line show a constant signal ($f = 0$) obtained with a sampling rate of $2P$.

Aliasing can be viewed from several angles. Conceptually, if $\Delta t > 1/(2f_N)$ (where f_N is the highest frequency component in phenomenon of interest), then a high frequency component will *masquerade* in the sampled series as a lower, artificial frequency component, as shown in Figure 8.28. If Δt is a multiple of P (e.g., see the squares in Figure 8.28), then this frequency component is indistinguishable from a horizontal line (i.e., a constant, with frequency $f = 0$). If $\Delta t = 5P/4$ (see circles in Figure 8.28) then this frequency component is indistinguishable from a component with frequency $1/5P$ (i.e., period of $5P$). Therefore, the under-sampled frequency components manifest themselves as lower frequency components (hence the word alias). In fact, every frequency *not* in the range

$$0 \leq f \leq 1/(2\Delta t) \tag{8.77}$$

has an alias in that range — this is its *principal alias*. Furthermore, any frequency $f_H > f_N$ will be indistinguishable from its principal alias. That is, the actual frequency $f_H = f_N + \Delta f$ will appear as the aliased frequency $f_L = f_N - \Delta f$.

Because of this relationship, the Nyquist frequency (f_N) is often called the *folding frequency* since the aliased frequencies ($f > f_N$) will appear at their principal aliases folded back into the range $\leq f_N$ (Figure 8.29). Therefore, when computing the transform of a data set, any frequency components in the phenomenon with true frequencies $f > f_N$ have been folded back into the resolved frequency range during sampling. Consequently, we must carefully choose Δt so that the powers at frequencies $f' > f_N$ are either small or nonexistent, or we must ensure that f_N is high enough so that the aliased part of the spectrum only affects frequencies higher than those of interest ($f \leq f_I$, see Figure 8.29).

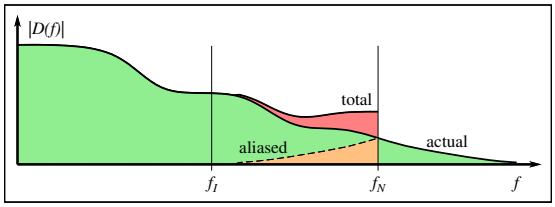


Figure 8.29: Aliasing and folding frequency. Power at higher frequencies than the Nyquist (f_N) will reappear as power at lower frequencies, "folded" around f_N. This extra power (orange) is then added to the actual power and the result is a distorted, total power spectrum (red). Selecting the Nyquist frequency so that aliasing only affects frequencies higher than the frequencies of interest ($f \leq f_I$). In this case, the extra power (orange) that is folded around f_N does not reach into the lower frequencies of interest, and consequently the total spectrum is unaffected for frequencies lower than f_I.

8.6 Aliasing and Leakage

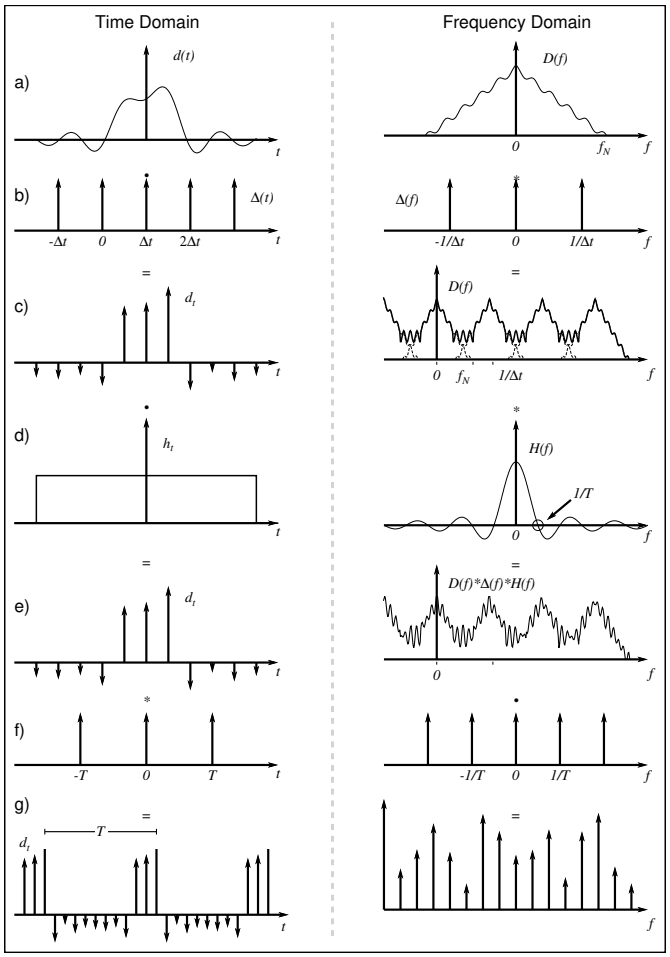


Figure 8.30: The continuous and band-limited phenomenon of interest, represented both in the time and frequency domains. Left column represents the time domain and the right column represents the frequency domain, separated by a vertical dashed gray line. The multiply, convolve, and equal signs indicate the operations that are being performed. a) Continuous phenomenon, b) Sampling function, c) Infinite discrete observations, d) Gate function, e) Truncated discrete observations, f) Assumed periodicity T, g) Aliasing and leakage of signal.

We were exploring the relationship between the continuous and discrete Fourier transform and found that we could illustrate the process graphically. First, we found that we had to sample the time-series $d(t)$ (Figure 8.30a). The sampling of the phenomenon by the sampling function $\Delta(t)$ (Figure 8.30b) is a multiplication in the time-domain, which implies a convolution in the frequency domain. This sampling yields discrete observations in the time domain, but the multiplication in the time domain equals a convolution in the frequency domain, enforcing periodicity of the spectrum (Figure 8.30c). Depending on the chosen sampling interval we may or may not have spectral overlap (aliasing). This discrete infinite series must then be truncated to contain a finite number of observations. The truncation is conceptually performed by multiplying our infinite time series with a finite gate function. This truncation of the infinite and periodic signal amounts to a multiplication in the time domain with a gate function, $h(t)$, whose transform is

$$H(f) = \text{sinc}(fT) = \frac{\sin \pi fT}{\pi fT}, \tag{8.78}$$

with both functions displayed in Figure 8.30d. This process results in the finite discrete observations shown in Figure 8.30e. It is this truncation that is responsible for introducing *leakage*.

Leakage arises because the truncation implicitly assumes that the time-series is periodic with period T (Figure 8.30f). Consequently, the discretization of frequencies is equivalent to enforcing a periodic signal (Figure 8.30g). Because both the time and frequency domain functions have been convolved with a series of impulses (by $\Delta(t)$ in time and $\Delta(f)$ in frequency), both functions are periodic in n discrete values, so the final discrete spectrum (for a real series as shown here) between 0 and f_N represents the discrete transform of the series on the left (which is periodic over T).

If the procedure in Figure 8.30 is followed mathematically, it is seen that the continuous Fourier transform is related to the discrete Fourier transform by the steps outlined graphically above. These show that a discrete Fourier transform will *differ* from the continuous transform by two effects:

1. Aliasing — from discrete time domain sampling.

2. Leakage — from finite time domain truncation.

Aliasing can be prevented by choosing $\Delta t \leq 1/(2f_N)$ or reduced as discussed previously. Leakage is always a problem for most observed (and hence truncated) time series. As discussed, leakage arises from truncation in the time domain, which corresponds to a convolution with a sinc function in the frequency domain. Conceptually, consider the effect of time domain truncation (Figure 8.31). Fourier analysis is essentially fitting a series of sines and cosines (using the harmonics of the fundamental frequency $1/T$) to the series $d(t)$. Since the Fourier series is necessarily periodic, it follows that

$$d(T/2 + \Delta t) = d(-T/2). \tag{8.79}$$

In other words, the transform is equivalent to that of a time series in which $d(t)$ is repeated every T (Figure 8.32).

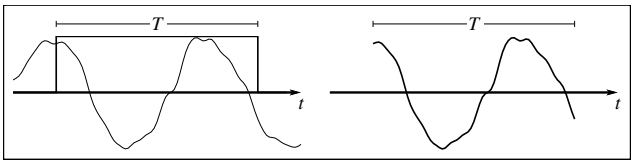


Figure 8.31: Truncation of a continuous signal, the equivalent of multiplying the signal with a gate function $h(t)$, determines the fundamental frequency, $f = 1/T$.

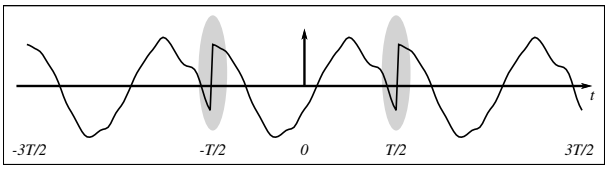


Figure 8.32: Artificial high frequencies are introduced due to the forced periodicity of a truncated time-series, which produces a discontinuous signal (highlighted by the gray regions).

The leakage (conceptually) thus results from the frequency components that must be present to allow the discontinuity, occurring every T, to be fit by the Fourier series. If the series $d(t)$ is perfectly periodic over T then there is no leakage because $d(T+\Delta t) = d(\Delta t)$ and the transition will be continuous and smooth across T.

To minimize leakage we attempt to minimize the discontinuity (between $d(0)$ and $d(T)$) or minimize the lobes of the $\text{sinc}(fT)$ function convolving the spectrum. This is accomplished by truncating the time series with a more gently sloping gate function (called a taper, fader, window, etc.). In other words, we use a smoother function that has fewer high frequency components (Figure 8.33).

The triangular function is the *Bartlett* window, which is the rectangle function convolved with itself (hence its transform is $\text{sinc}^2(fT)$). The dashed line is the split cosine-bell window. Other windows include: *Hanning* (a cosine taper), *Parzen* (similar to Hanning but decays sooner and more steeply, *Hamming* (like Hanning), and *Bartlett-Priestley* (which is quadratic and has "optimal" properties, satisfying specific error considerations.) All of these tapers have transforms that are less oscillatory than the sinc function but they are also wider. Therefore, multiplication of the time series with one of these gate functions results in a convolution whose transform in the frequency domain will smear spectral peaks more than the sinc function did. In return, it will not introduce ripples far away from these spectral peaks.

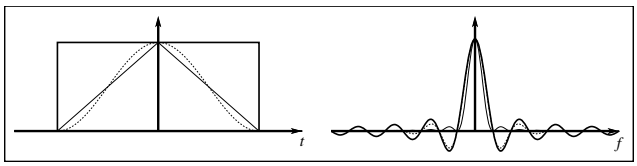


Figure 8.33: Alternative gate functions and their spectral representations. The less abrupt a gate function is in the time domain the less ringing it will introduce in the frequency domain.

Note that multiplying by, say, a Hanning window will make $d(T/2+\Delta t) \sim d(-T/2)$, so the bothersome discontinuity is eliminated — however damping of all $d(t)$ away from $d(T)$ acts like a modulation, which accounts for the smearing of spectral peaks. Hence, leakage is still not completely eliminated.

8.7 Complex Fourier Series

As mentioned, Fourier series combine *even* (cosine) and *odd* (sine) components. It is common to simplify these expressions by using *complex notation*, in which a *complex number* z is written

$$z = x + iy. \tag{8.80}$$

Here, x is considered the *real* part, y is the *imaginary* part, and $i = \sqrt{-1}$ is the *imaginary number*. The concept of considering a complex number as a point (x,y) in the complex plane was first presented by Caspar Wessel[1] in 1799, but like many mathematical inventions others had independently dabbled with this both before and after 1800. In the end, our old friend K. F. Gauss made many contributions to complex number theory.

Simple rules govern elementary operations on complex numbers. For instance, addition and subtraction follow

$$\begin{aligned} z &= (a+ib)+(c+id) = (a+c)+i(b+d), \\ z &= (a+ib)-(c+id) = (a-c)+i(b-d), \end{aligned} \tag{8.81}$$

while multiplication becomes

$$z = (a+ib)(c+id) = ac+iad+ibd+i^2bd = (ac-bd)+i(bc+ad). \tag{8.82}$$

Division is converted to a multiplication by a complex number and a division by a real number simply by multiplying both numerator and denominator by the denominator's *complex conjugate*, for which the imaginary part changes sign:

$$z = \frac{(a+ib)}{(c+id)} = \frac{(a+ib)(c-id)}{(c+id)(c-id)} = \frac{(ac+bd)+i(bc+ad)}{(c^2+d^2)}. \tag{8.83}$$

[1] No relation!

The use of the complex notation simplifies much of the algebra associated with Fourier analysis and is therefore mathematically more convenient to use. This is especially true in two and higher dimensions. We will employ *Euler's formula* for complex numbers in our expressions of the sine and cosine transform. Thus, we digress to discuss Euler's formula.

8.7.1 Euler's formula

Figure 8.34: Leonard Euler (1707–1783) was one of the most productive mathematicians of his era. You know you did something right when your mugshot ends up on stamps, even in former countries like DDR and the Soviet Union. His famous equation $e^{i\pi}+1=0$, relating the five most important numbers in mathematics, follows from his complex relation given in (8.88) for $\omega t=\pi$.

Using a Taylor series expansion, we can write

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots, \tag{8.84}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots, \tag{8.85}$$

and

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots. \tag{8.86}$$

Now, let us introduce

$$x = i\theta = \sqrt{-1}\theta.$$

Inserting this expression into (8.86) yields

$$\begin{aligned} e^{i\theta} &= 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} - \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right) = \\ &\quad \cos\theta + i\sin\theta. \end{aligned} \tag{8.87}$$

For $x=-i\theta$ we instead get

$$\begin{aligned} e^{-i\theta} &= 1 - i\theta - \frac{\theta^2}{2!} + \frac{i\theta^3}{3!} - \frac{\theta^4}{4!} - \frac{i\theta^5}{5!} - \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) - i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right) = \cos\theta - i\sin\theta. \end{aligned}$$

We will associate $\cos\theta$ with the positive x-axis and $i\sin\theta$ with the positive y-axis. Thus, $e^{i\theta}$ is a *unit vector* rotated an angle θ counter-clockwise from the x-axis (examine Figure 8.1 one more time). For the angular frequencies in the Fourier series (where $\theta=\theta(t)=\omega t$), the Euler relation is

$$\boxed{e^{\pm i\omega t} = \cos\omega t \pm i\sin\omega t,} \tag{8.88}$$

while the *inverse* relationships are

$$\begin{aligned} \cos\omega t &= \tfrac{1}{2}(e^{i\omega t} + e^{-i\omega t}), \\ \sin\omega t &= \tfrac{1}{2i}(e^{i\omega t} - e^{-i\omega t}). \end{aligned} \tag{8.89}$$

We will refresh (or introduce) a few concepts that apply to complex numbers. The *complex conjugate* (denoted by *) of a function $f(x)$ is written $f^*(x)$, with components

$$\begin{aligned} f(x) &= R(x) + iI(x), \\ f^*(x) &= R(x) - iI(x). \end{aligned} \tag{8.90}$$

Here, $R(x)$ is the real part and $I(x)$ is the imaginary part of $f(x)$, respectively. The *magnitude* or amplitude of a complex number is given by

$$A(x) = |f(x)| = \sqrt{f(x)\cdot f^*(x)} = \sqrt{R^2(x) + I^2(x)}, \tag{8.91}$$

while the *phase* is obtained via

$$\phi(x) = \tan^{-1}\frac{I(x)}{R(x)}. \tag{8.92}$$

8.7.2 Using the complex notation

By using Euler's formula the *five* orthogonality relations discussed in Section 8.2 become just *one*:

$$\sum_{\ell=1}^{n} e^{i\omega_j t_\ell}\cdot e^{-i\omega_k t_\ell} = \begin{cases} n, & j = k \\ 0, & \text{otherwise} \end{cases} \tag{8.93}$$

where we now use ℓ to indicate the point number (since i here represents the imaginary number). The equivalent integral relationship is obviously

$$\int_0^T e^{i\omega_j t}\cdot e^{-i\omega_k t}dt = \begin{cases} T, & j = k \\ 0, & \text{otherwise} \end{cases} \tag{8.94}$$

Any real-valued function can be represented as a complex-valued function with a zero imaginary part. Using the inverse Euler relations, the Fourier series (8.29) can now be rewritten in complex form as follows, again assuming that n is an even number:

$$\begin{aligned} d_\ell &= \sum_{j=0}^{n/2}\left[a_j\cos\omega_j t_\ell + b_j\sin\omega_j t_\ell\right] = \sum_{j=0}^{n/2}\left[\frac{a_j}{2}\left(e^{i\omega_j t_\ell} + e^{-i\omega_j t_\ell}\right) + \frac{b_j}{2i}\left(e^{i\omega_j t_\ell} - e^{-i\omega_j t_\ell}\right)\right] \\ &= \tfrac{1}{2}\sum_{j=0}^{n/2}\left[\left(a_j + \frac{b_j}{i}\right)e^{i\omega t_j} + \left(a_j - \frac{b_j}{i}\right)e^{-i\omega_j t_\ell}\right] \\ &= \frac{1}{2}\sum_{j=0}^{n/2}\left[(a_j - ib_j)\,e^{i\omega_j t_\ell} + (a_j + ib_j)\,e^{-i\omega_j t_\ell}\right]. \end{aligned} \tag{8.95}$$

Since the second term contains $e^{-\omega_j t_\ell}$, we shall consider the effect of introducing *negative* frequencies (i.e., let j take on negative values) in simplifying this expression further. In general, for negative j (i.e., $-j$),

$$(a_{-j} - ib_{-j})\,e^{i\omega_{-j} t_\ell} = (a_j + ib_j)\,e^{-i\omega_j t_\ell}, \tag{8.96}$$

which we see are related by the complex conjugate definition

$$J_{-j} = J^*{}_j, \tag{8.97}$$

because

$$a_{-j} = \frac{2}{n}\sum_{\ell=1}^{n} d_\ell \cos\omega_{-j}t_\ell = \frac{2}{n}\sum_{\ell=1}^{n} d_\ell \cos\frac{-2\pi j}{T}t_\ell = \frac{2}{n}\sum_{\ell=1}^{n} d_\ell \cos\frac{2\pi j}{T}t_\ell = a_j,$$

$$b_{-j} = \frac{2}{n}\sum_{\ell=1}^{n} d_\ell \sin\omega_{-j}t_\ell = \frac{2}{n}\sum_{\ell=1}^{n} d_\ell \sin\frac{-2\pi j}{T}t_\ell = \frac{-2}{n}\sum_{\ell=1}^{n} d_\ell \sin\frac{2\pi j}{T}t_\ell = -b_j.$$

Consequently, a_j is an *even* function, b_j is an *odd* function, and obviously

$$e^{i\omega_{-j}t_\ell} = e^{-i\omega_j t_\ell}.$$

Therefore, the second term of (8.95) can be dropped if we merely extend the sum over $-n/2 \le j < n/2$:

$$\boxed{d_\ell = \sum_{j=-n/2}^{<n/2} \frac{1}{2}(a_j - ib_j)e^{i\omega_j t_\ell} = \sum_{j=-n/2}^{<n/2} J_j e^{i\omega_j t_\ell}, \quad \ell = 1,n.} \tag{8.98}$$

Notice that the complex form has *twice* as many coefficients as the real form, reflecting the fact that each value of d_ℓ contains a real and an imaginary component. Thus, there are actually *twice* as many d_ℓ values as before, even though for most observations the imaginary components will all be zero. Equation (8.98) represents the general complex form of the Fourier series, with complex coefficients

$$J_j = \frac{1}{2}(a_j - ib_j). \tag{8.99}$$

By substituting the expressions for a_j (8.46) and b_j (8.50) into the expression above (for J_j) *or* by multiplying Equation (8.98) by $e^{-i\omega_k t_\ell}$, summing over all t_ℓ, and using the orthogonality relation (8.93), we can derive the general complex form for J_j via

$$\sum_{\ell=1}^{n} d_\ell e^{-i\omega_k t_\ell} = \sum_{j=-n/2}^{<n/2} J_j \sum_{\ell=1}^{n} e^{i\omega_j t_\ell} e^{-i\omega_k t_\ell},$$

yielding

$$\sum_{\ell=1}^{n} d_\ell e^{-i\omega_k t_\ell} = nJ_k. \tag{8.100}$$

Hence, and swapping the dummy indices j and k, we find

$$\boxed{J_j = \frac{1}{n}\sum_{\ell=1}^{n} d_\ell e^{-i\omega_j t_\ell}, \quad -\frac{n}{2} \le j < \frac{n}{2}.} \tag{8.101}$$

Equation (8.101) is the *complex discrete Fourier transform* of the series d_ℓ and equation (8.98) is the *complex discrete inverse Fourier transform*. Since $J_{n/2} = J_{-n/2}$ (and both are real since $b_{n/2} = 0$) we need not compute it for each index, hence $-n/2 \le j < n/2$. This arrangement also eliminates the awkward need to define a_0 and $a_{n/2}$ as one-half their values.

When d_ℓ is a *real-valued* series, then $d_\ell = d_\ell^*$ because all the imaginary parts are zero, so

$$J_j = \frac{1}{n}\sum_{\ell=1}^{n} d_\ell e^{-i\omega_j t_\ell} = \frac{1}{n}\sum_{\ell=1}^{n} d_\ell(\cos\omega_j t_\ell - i\sin\omega_j t_\ell) = \frac{1}{n}\sum_{\ell=1}^{n} d_\ell \cos\omega_j t_\ell - \frac{i}{n}\sum_{\ell=1}^{n} d_\ell \sin\omega_j t_\ell$$

and hence

$$J_{-j} = J_j^*.$$

Therefore, the transform (i.e., J_j) is completely determined by the positive values of j, as developed earlier, i.e., they are completely determined by the n values of a_j and b_j. Specifically, if we need to recover a_j and b_j from the J_j coefficients we simply note that

$$J_{-j} + J_j = J_j^* + J_j = \frac{1}{2}(a_j + ib_j) + \frac{1}{2}(a_j - ib_j) = a_j, \tag{8.102}$$

and likewise

$$J_{-j} - J_j = J_j^* - J_j = \frac{1}{2}(a_j + ib_j) - \frac{1}{2}(a_j - ib_j) = ib_j. \tag{8.103}$$

The negative frequencies confuse many practitioners of spectral analysis. What are they? What do they represent? Are they components going backwards in time or space? It is important to remember that we simply introduced these as a *mathematical convenience* in order to arrive at a simple and compact transform, i.e., (8.101). Physically, there are only positive frequencies.

8.7.3 The continuous Fourier transform in 1-D

For a continuous function $g(t)$, equivalent expressions make up the *forward* and *inverse Fourier transform* pair, typically written in complex (and symmetric) form as

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-i2\pi t f} dt \tag{8.104}$$

and

$$g(t) = \int_{-\infty}^{\infty} G(f) e^{+i2\pi t f} df, \tag{8.105}$$

respectively, for radial frequency f. As noted, the main differences between the continuous transforms and the discrete transforms (which we necessarily must use to operate on observed data) are:

1. Our data have a finite, nonzero sampling interval, whereas the theory considers continuous distributions. This difference opens up the possibility of *aliasing*.
2. Our data have a finite length (i.e., a fundamental period), whereas the theoretical distributions exist over all time. This limitation may lead to issues involving *leakage*.

8.8 Computing Fourier Transforms

Most practitioners of spectral analysis will find themselves performing the Fourier transform via a special algorithm known as the *Fast Fourier Transform* (FFT). This algorithm as well as practical information on how to use it and how to prepare our data for it is the focus of this section.

8.8.1 The Fast Fourier Transform (FFT)

The discrete Fourier transform involves n coefficients that each requires n multiplications to be determined. Hence, the entire calculations is $O(n^2)$ in computer time. Can we do better than that? Since $\omega_j = 2\pi j/T$, $T = n\Delta t$, and $t_\ell = \ell\Delta t$ for $\ell = 0, ..., n-1$ it follows that the exponential term in (8.101) is

$$e^{-i\omega_j t_\ell} = e^{-\frac{i2\pi j}{n\Delta t}\ell\Delta t} = \left[e^{-\frac{2\pi i}{n}}\right]^{j\ell} = W^{j\ell}. \tag{8.106}$$

Note what happened to time here; it cancels, leaving just a ratio of integers. We may now write

$$J_j = \frac{1}{n}\sum_{\ell=0}^{n-1} y_\ell W^{j\ell}. \tag{8.107}$$

It is common in FFT implementations to perform the division by n separately (at the end), so let us just consider

$$J_j = \sum_{\ell=0}^{n-1} y_\ell W^{j\ell}. \tag{8.108}$$

You may note that if $j = 0$ we obtain $J_0 = \bar{y}$ (apart from the $1/n$ term, that is). Now, consider this rearrangement:

$$J_j = \sum_{\ell=0}^{n-1} y_\ell e^{-\frac{2\pi i j\ell}{n}} = \sum_{\ell=0}^{n/2-1} y_{2\ell} e^{-\frac{2\pi i j(2\ell)}{n}} + \sum_{\ell=0}^{n/2-1} y_{2\ell+1} e^{-\frac{2\pi i j(2\ell+1)}{n}}. \tag{8.109}$$

Here, we have simply rewritten the sum as two separate series containing the even (2ℓ) and odd ($2\ell+1$) terms of the observations, respectively. This expression can be manipulated further to yield

$$J_j = \sum_{\ell=0}^{n/2-1} y_{2\ell} e^{-\frac{2\pi i j \ell}{(n/2)}} + e^{-2\pi i j/n} \sum_{\ell=0}^{n/2-1} y_{2\ell+1} e^{-\frac{2\pi i j \ell}{(n/2)}}, \tag{8.110}$$

which we can write as

$$J_j = J_j^e + W^j J_j^o. \tag{8.111}$$

The first term is the Fourier transform of the even-numbered observations (i.e., considering only the 0th, 2nd, 4th, etc.,

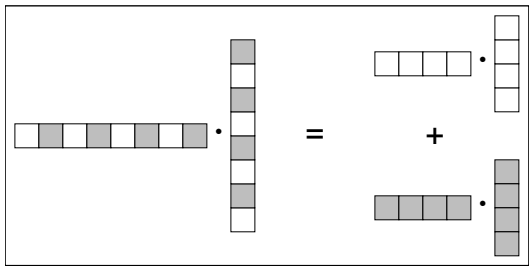


Figure 8.35: Each dot-product making up a Fourier transform for a single frequency can be split in two by considering the odd (gray) and even (white) observations separately, and by recursively applying this partitioning until we reach a single element we can speed of the Fourier transform enormously.

values), while the second term is the Fourier transform of just the odd-numbered observations (i.e., the 1st, 3rd, 5th, etc. values), scaled by the constant W^j. Since each transform only deals with $n/2$ points we find the calculation time to be proportional to $O(2(n/2)^2) = O(n^2/2)$. So while (8.110) is mathematically identical to (8.101), the partitioning into two sums leads to a 50% reduction in computation time. But wait, there is more! The partitioning idea can be continued recursively on the two separate J_j^e and J_j^o sums as well. For instance, the two terms in the transform for J_j^e requiring $n/2$ data values each can themselves be written as

$$\begin{aligned} J_j^e &= J_j^{ee} + W^j J_j^{eo}, \\ J_j^o &= J_j^{oe} + W^j J_j^{oo}, \end{aligned} \tag{8.112}$$

with each term being a transform requiring just $n/4$ data points. If n is initially some power of 2, then carrying this splitting all the way to a single number (whose transform is itself) yields a computational workload that is $O(n \log n)$. These are huge savings compared to our initial estimate of $O(n^2)$! For instance, with $n = 10^6$, the difference is a factor of 50,000, while for $n = 10^7$ the operations are 400,000 times faster. For each magnitude in n, this ratio increases by another factor of ~ 8. This clever result is essentially the *Cooley-Tukey* FFT algorithm from the 1960s. However, its origin goes further back to the 1940s with *Danielson* and *Lanczos* and their work on x-ray scattering as well as further back to *Gauss*. Apparently, every neat mathematical idea can eventually be traced back to Gauss or Euler...

8.8.2 FFT implementations

Both MATLAB and Octave implement the fast Fourier transform in a similar fashion, as do most mathematical function libraries. There are three functions we will need to become familiar with:

fft: The forward fast Fourier transform.

ifft: The inverse fast Fourier transform.

fftshift: Rearranging the order of spectral coefficients.

To see how these work and what they expect as input we will look at the arguments of data and frequencies. Let our data be represented by the data array

$$\mathbf{d} = [\quad d_1, \quad d_2, \quad \ldots, \quad d_n \quad], \tag{8.113}$$

where n is even and represents the number of data points. Then, taking the transform [i.e., `J = fft(d);` in MATLAB or Octave] yields the array

$$\mathbf{J} = [\quad J_0,\quad J_1,\quad J_2,\quad \ldots,\quad J_{\frac{n}{2}-1},\quad J_{-\frac{n}{2}},\quad J_{-\frac{n}{2}+1},\quad \ldots,\quad J_{-1}\quad]. \tag{8.114}$$

Clearly, the amplitudes are split between the positive (first half) and negative (second half) sets of the coefficients, i.e.,

$$|J_k| = |J_{-k}|. \tag{8.115}$$

Plotting $J(f)$ is awkward since the order is discontinuous with respect to frequency, f. To place J_0 symmetrically in the middle requires shifting of the coefficients. This is done by *fftshift* [i.e., `J = fftshift(J);`], which results in the array

$$J = [\quad J_{-\frac{n}{2}},\quad J_{-\frac{n}{2}+1},\quad \ldots,\quad J_{-1},\quad J_0,\quad J_1,\quad J_2,\quad \ldots,\quad J_{\frac{n}{2}-1}\quad], \tag{8.116}$$

where the very first item ($J_{-\frac{n}{2}}$) is the real coefficient associated with the Nyquist frequency (since the sine term is identically zero), and J_0 reflects the mean (real) value of d. These are the only two coefficients that *do not* appear twice in (8.114). Note that your particular time coordinates t_ℓ play *no part* in the transform. It is implicitly assumed that $t_\ell = \ell\Delta t$ and that $t_0 = 0$, so make sure you are careful in selecting your origin time.

8.8.3 Detrending and windowing

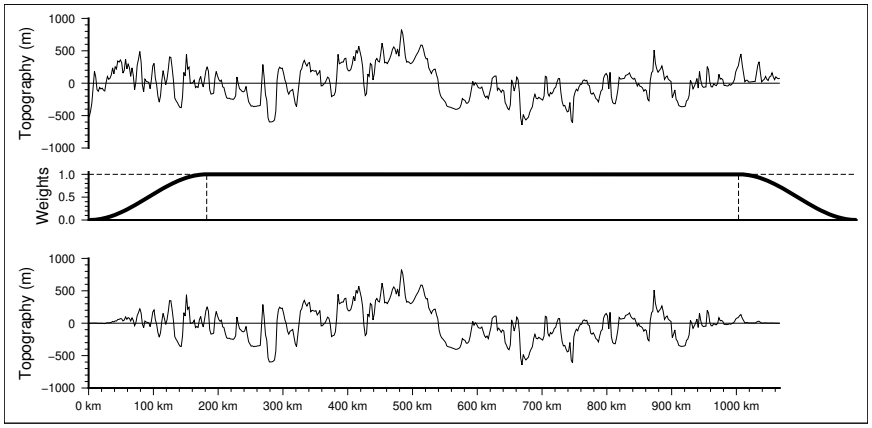


Figure 8.36: (top) Original topographic profile. If this data set were subjected to Fourier analysis, then the jump discontinuity due to the forced periodicity would introduce leakage. (middle) We reduce the leakage by using a smooth rather than rectangular window function. Here, a *Tukey* window is shown, which is a cosine taper between zero at the ends and unity in the middle. (bottom) Multiplying our data with the window function yields a data set that no longer has a jump discontinuity.

Prior to computing spectral estimates with an FFT we should do our utmost to reduce the influence of leakage. As we discussed, leakage arises due to truncation of a hypothetically infinite data series to a finite length data set (i.e., when making our observation), and the range of the observed data becomes our fundamental (and probably artificial) period. Since spectral analysis uses periodic functions (sines and cosines) to represent the data, we are in effect forcing our data to be periodic as well, with the presumably arbitrary data range as the fundamental period. Unless the data happen to start and end on the same value we may have a potentially large offset between the two values, which essentially constitutes a step function (e.g., upper panel of Figure 8.36). Decomposing the data into sines and cosines means we will approximate this step using Fourier building blocks, thus leaking energy over a wide range of frequencies. The standard approach to minimizing this problem is to use a gentler gate function than the rectangle we implicitly used when we captured our data. Figure 8.36 (middle panel) shows what is called a Tukey window, which essentially is a half cosine wavelength that connects the area outside the window (zero weight) to the central portion of the window (unit weight). By multiplying our observed data by this window we eliminate the step mismatch between the start and end of the data (e.g., lower panel of Figure 8.36), and this data modification dramatically reduces the effect of leakage. Note, however, that leakage is not completely eliminated since our tapering still changes the

observed signal in some way. For one thing, it effectively shortens the fundamental period by an amount proportional to the ramp margin.

8.8.4 Zero-padding

These are two common and seemingly related problems in spectral analysis:

1. For many data sets, the frequency resolution $\Delta f = 1/T$ may not be small enough to resolve spectral components that are closely spaced (remember, all the Fourier frequencies we use are integer multiples of Δf).

2. The discrete points in the raw periodogram may be too widely spaced to resolve the actual frequencies of the peaks in question.

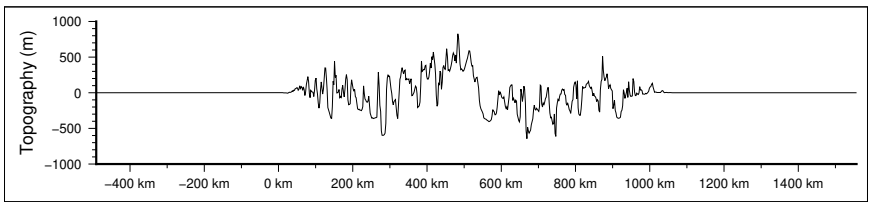


Figure 8.37: Zero-padding our data means to extend the range of the windowed data by adding zeros until the length (i.e., the number of data points) reaches the next power of 2.

Unfortunately, the first issue can only be addressed by collecting a longer time series so as to decrease Δf to a point where the two peaks are clearly separated in the periodogram. Without this added resolution such peaks will be seen as one wider (blurred) peak. The second problem, however, can be addressed by extending the finite time series with zeros prior to taking the FFT (Figure 8.37). By adding zeroes, the length of the data series (n) increases, effectively adding additional frequency components in between those that would be obtained for the original, non-padded series. In essence, we obtain an *interpolation* of the spectral density estimates. Note that zero-padding helps fill in the shape of the spectrum but of course there is no improvement in the fundamental frequency resolution. Nevertheless, zero-padding is widely used for several reasons:

1. It smooths the shape of the periodigram via spectral interpolation.

2. It may resolve potential ambiguities where the frequency difference between line spectra is greater than the fundamental frequency resolution.

3. It may help define the exact frequency of the peaks of interest by reducing the quantization of the power.

4. It may increase n to an integer power of 2, thus speeding up the analysis when the FFT is used.

While these benefits are all good, we note that adding zeros does not help us to distinguish closely spaced frequency components that could not be resolved in the original time series.

8.9 Filtering

Filtering data is a major data processing procedure that is used throughout the natural sciences. Such procedures are applied to reduce "undesired" features in the observations or to enhance the "desired" features. What the desired and undesired features are may change completely from application to application and may even be interchangeable. Filters may be expressed as convolutions (and thus may take advantage of the speed-up provided by the convolution theorem) or they must be executed relatively slowly in the time (or space) domains. Both types of filters play important roles in data processing and interpretation. Before we discuss a range of well-known filters we will digress to revisit convolution and examine various properties of the Fourier transform.

8.9.1 Convolution and the Fourier Transform

Before revisiting convolution, we need to take a closer look at the analytical Fourier transform as given by (8.104). Note that this general (continuous) form differs from the discrete form (8.101) in that it is no longer dependent upon discrete Fourier frequencies nor is it restricted to a finite period, T. Of course, for an actual transform of a discrete time series the discrete Fourier transform is always used, thus in practice we will consider the Fourier frequencies only. The general integral form is used to transform *analytic* expressions, such as may arise in the solution to differential equations, and is thus useful in other ways as well.

Using the notation of (8.104), Fourier transform pairs are usually designated by

$$g(t) \Leftrightarrow G(f),$$

i.e., the transform of a function $g(t)$, which is a function of time (or distance) and traditionally denoted by lowercase letters, is represented by the corresponding uppercase version of the function, $G(f)$, and varies as a function of frequency f (or wavenumber k for spatial data).

Example 8–2. We will look at an important analytic Fourier transform pair. Consider the instantaneous impulse or delta function

$$g(t) = \delta(t).$$

Per the definition of a *Dirac delta function*, its transform will be

$$G(f) = \int_{-\infty}^{\infty} \delta(t) e^{-i2\pi ft} dt = e^0 = 1. \tag{8.117}$$

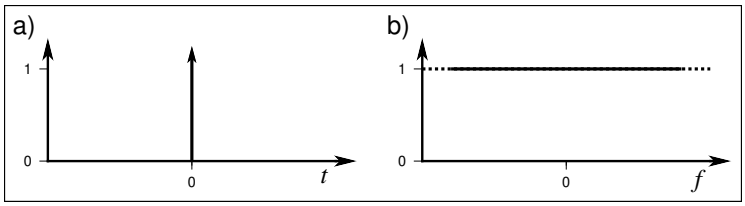


Figure 8.38: (a) A delta-function (or spike) in one domain (here, the time domain) will be transformed (b) to a constant in the other domain. Note that the phase spectrum is zero since there are no imaginary part to the transform.

Thus, the narrowest possible function in the time domain has the broadest possible spectral expression, reflecting the reciprocal nature of time and frequency (Figure 8.38). Another way to look at this result is to say that an impulse is simply a sum of cosines of unit amplitude for all frequencies. This equivalence is used in geophysical exploration for oil and gas on land. Instead of setting off explosions (i.e., an impulse), *Vibroseis* trucks generate sweeping harmonic signals of increasing frequency. In post-processing, the Earth's response to all these harmonics can be combined into the equivalent response to an actual impulse. Because we cannot use frequencies approaching infinity the resulting impulse will have a finite width, such as indicated in Figure 8.39.

Other uses of analytic Fourier transforms include obtaining insight into the nature of spectra obtained from real processes by studying those of very simple models.

8.9.2 The scale theorem

What happens if we contract or expand the time axis? We find the rule

$$\text{If } g(t) \leftrightarrow G(f) \text{ then } g(at) \leftrightarrow \frac{1}{|a|} G\left(\frac{f}{a}\right). \tag{8.118}$$

Proof, using $u = at$:

$$\int_{-\infty}^{\infty} g(at) e^{-i2\pi ft} dt = \frac{1}{a} \int_{-\infty}^{\infty} g(u) e^{-i2\pi u/a} du = \frac{1}{|a|} G\left(\frac{u}{a}\right).$$

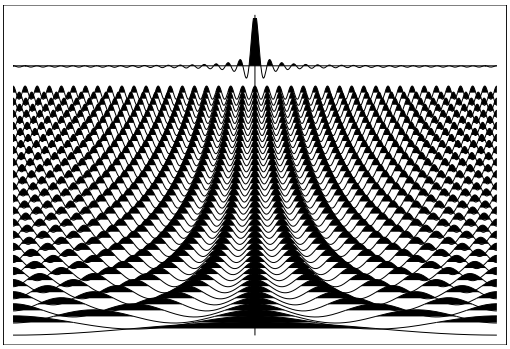


Figure 8.39: Per (8.117), if we sum up all the harmonics of uniform amplitude we should obtain a delta-function. Here we add the first 40 harmonics, yielding an approximate impulse, which illustrates the principle behind Vibroseis.

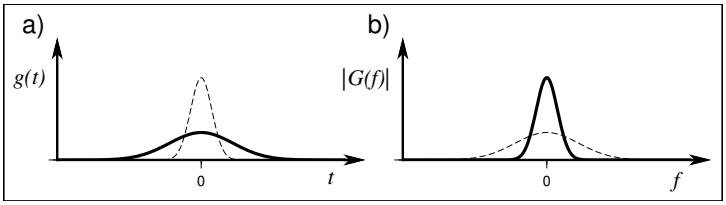


Figure 8.40: (a) A wide (heavy line) and narrow (dashed line) function in the time domain will transform to (b) a narrow and wide function in the frequency domain, respectively.

If a is negative, then the limits of integration are reversed, so the frequency domain amplitudes are still scaled positively, hence the $|a|$. Therefore, if a time scale is contracted by a, the frequency is expanded $(1/|a|)$; this is expected since time is inversely proportional to frequency. The amplitude of the frequency is decreased, which maintains a constant area relative to the unscaled function.

8.9.3 The shift theorem

If we apply a shift to the time axis, we find another rule:

$$\text{If } g(t) \leftrightarrow G(f) \text{ then } g(t-t_0) \leftrightarrow G(f)e^{-i2\pi f t_0}. \tag{8.119}$$

In polar form,

$$G(f)e^{-i2\pi f t_0} = [A(f)e^{i\theta(f)}]e^{-i2\pi f t_0} = A(f)e^{i(\theta(f)-2\pi f t_0)}.$$

Therefore, a shift in the time axis (a *translation*) does not affect the amplitude spectrum, it only affects the *phase* spectrum. It linearly shifts the phase by $\theta_0 = 2\pi f t_0$, thus the slope of the phase shift is directly proportional to the time shift t_0 (i.e., if phase were plotted against f, then the slope would equal t_0). This means the transform of a shifted impulse $\delta(t-t_0)$ is $e^{i(\theta-\theta_0)}$, showing that the transform of random noise $n(t)$ has a constant amplitude spectrum $N(f) = 1$ but a random phase spectrum.

8.9.4 The gate function

Finally, we will look at the transform of a very useful construction known as a *gate* function:

$$g(t) = \begin{cases} 1, & |t| \le T/2 \\ 0, & \text{elsewhere} \end{cases}. \tag{8.120}$$

It transforms as follows:

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-i2\pi f t}dt = \int_{-T/2}^{T/2} e^{-i2\pi f t}dt = \int_{-T/2}^{T/2} (\cos 2\pi f t - i\sin 2\pi f t)dt = \int_{-T/2}^{T/2} \cos 2\pi f t dt.$$

Here we have used the fact that the sine term, being odd, will integrate to zero. Using the substitution $u = 2\pi ft$ we obtain

$$G(f) = \frac{1}{2\pi f}\int_{-\pi fT}^{\pi fT} \cos u du = \frac{1}{2\pi f}\sin u \mid_{-\pi fT}^{\pi fT} = \frac{1}{2\pi f}\left(\sin(\pi fT) - \sin(-\pi fT)\right).$$

We rearrange this expression to give

$$G(f) = T\frac{\sin(\pi fT)}{\pi fT} = T\,\mathrm{sinc}(fT), \tag{8.121}$$

which is the definition of the sinc-function.

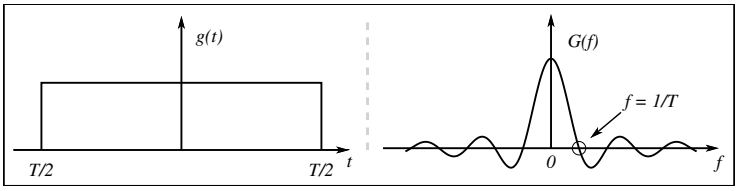


Figure 8.41: Truncating an infinite and periodic signal amounts to multiplying it in time with a gate function, $g(t)$, whose transform $G(f) = T\,\mathrm{sinc}(fT)$.

8.9.5 The convolution theorem

The *convolution theorem* simply states that a convolution in the time domain is equivalent to a multiplication in the frequency domain, and vice-versa, written as

$$s(t) * g(t) \leftrightarrow S(\omega)\cdot G(\omega). \tag{8.122}$$

If $h(t) = s(t) * g(t)$, then

$$H(\omega) = \int_{-\infty}^{\infty} h(t)e^{-i\omega t}dt = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} s(\tau)g(t-\tau)d\tau\right]e^{-i\omega t}dt.$$

Notice the lag is defined as t in the innermost integral. Interchanging the order of integration (this assumes continuity of the integrand over $\pm\infty$) gives

$$H(\omega) = \int_{-\infty}^{\infty} s(\tau)\left[\int_{-\infty}^{\infty} g(t-\tau)e^{-i\omega t}dt\right]d\tau.$$

Let $u = t - \tau$, so $dt = du$. Then,

$$\begin{aligned} H(\omega) &= \int_{-\infty}^{\infty} s(\tau)\left[\int_{-\infty}^{\infty} g(u)e^{-i\omega(u+\tau)}du\right]d\tau = \int_{-\infty}^{\infty} s(\tau)e^{-i\omega\tau}\left[\int_{-\infty}^{\infty} g(u)e^{-i\omega u}du\right]d\tau \\ &= G(\omega)\int_{-\infty}^{\infty} s(\tau)e^{-i\omega\tau}d\tau = S(\omega)\cdot G(\omega). \end{aligned}$$

The result of this theorem provides the basic tool from which (among other things) the continuous Fourier transform is related to the discrete Fourier transform. Also, because of symmetry,

$$s(t)\cdot g(t) \leftrightarrow S(\omega) * G(\omega). \tag{8.123}$$

8.9.6 Parseval's theorem

Parseval's theorem is a statement relating the variance of a data set to its power spectrum. For a continuous function $g(t)$, the statement becomes

$$\int_{-\infty}^{+\infty} \mid g(t) \mid^2 dt = \int_{-\infty}^{+\infty} \mid G(f) \mid^2 df. \tag{8.124}$$

In Section 8.3 we derived a discrete version of Parseval's theorem by using a Fourier series expansion for $g(t)$, removing its mean, and computing the variance of the finite time series. Using our complex notation we may now simply write

$$y_\ell = \sum_{j\geq -n/2}^{<n/2} J_j e^{i\omega_j t_\ell}, \quad \ell = 1,n \tag{8.125}$$

and form the discrete version of (8.124) via (8.101) to yield

$$\sum_{\ell=1}^{n} y_\ell^2 = \sum_{\ell=1}^{n} \left\{ \left[\sum_{j\geq -n/2}^{<n/2} J_j e^{i\omega_j t_\ell} \right] \left[\sum_{k\geq -n/2}^{<n/2} J_k e^{i\omega_k t_\ell} \right] \right\}. \tag{8.126}$$

Because of orthogonality (8.93), only the $j = k$ cross-terms will be nonzero and we find

$$\sum_{\ell=1}^{n} y_\ell^2 = n \sum_{j\geq -n/2}^{<n/2} J_j^2. \tag{8.127}$$

By normalizing by $(n-1)$ we see the partitioning of data variance into its frequency contributions:

$$s^2 = \frac{1}{n-1} \sum_{\ell=1}^{n} y_\ell^2 = \frac{n}{n-1} \sum_{j\geq -n/2}^{<n/2} J_j^2 \approx \sum_{j\geq -n/2}^{<n/2} J_j^2. \tag{8.128}$$

8.9.7 Convolution filters

Filtering of data is typically performed in order to either smooth the signal or suppress power at particular frequencies or wavenumbers. So far, we have learned that filtering can be considered an example of a convolution between the data $d(t)$ and the filter $p(t)$, i.e.,

$$h(t) = d(t) * p(t). \tag{8.129}$$

Thus, we can immediately take advantage of the convolution theorem and write

$$H(f) = D(f) \cdot P(f). \tag{8.130}$$

Hence, we can simply take the Fourier transform of $d(t)$ and multiply its spectral components $D(f)$ by $P(f)$. For example, we learned previously that a simple MA filter consists of convolving the signal with a rectangle function of width W (Figure 8.42). Figure 8.43 shows how this operation might look like in the frequency domain.

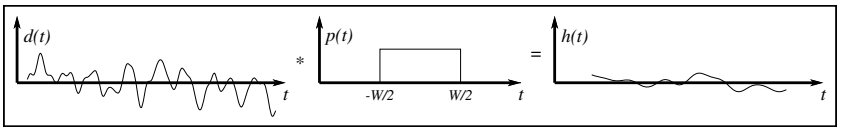


Figure 8.42: Moving average filter convolution seen in the time domain.

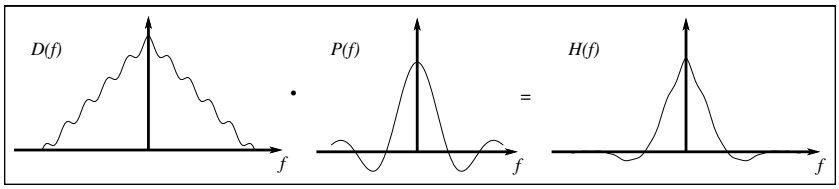


Figure 8.43: Moving average filtering seen in the frequency domain.

Thus, the MA filter output is obtained by multiplying the signal's Fourier coefficients by a sinc function. This means that some of the coefficients will have their signs reversed, which we know equates to a phase change of 180°. Also, while the filter coefficients do fall off with increasing f, they do so in a very oscillatory way and take a long time to

approach zero. This fact would suggest that the sinc function is a poor choice for a filter if you really wanted to cut out, say, high frequency information. Perhaps we should instead design $P(f)$ directly so that it truncates all power at frequencies higher than f_{cut}? Surely, by simply removing all power from frequencies of no interest we should retain the part of the signal that we are most interested in. This alternative operation, which is simply carried out in the frequency domain, is illustrated in Figure 8.44. The filtered output is finally obtained by transforming $H(f)$ back to the time domain (Figure 8.45).

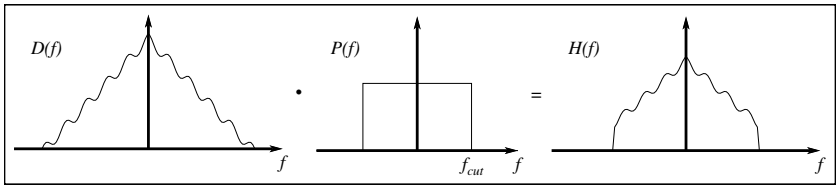


Figure 8.44: Truncation of high frequencies is straightforward in the frequency domain.

We discover that the output has a lot of "ripples" (here somewhat exaggerated) at approximately the period corresponding to $1/f_{\text{cut}}$. This "ringing" is caused by our convolving of $d(t)$ with $p(t)$, which in this case is a sinc function. Because all power at higher frequencies are completely eliminated, the power at the remaining highest frequency f_{cut} stands out. This effect is referred to as "Gibbs' phenomenon".

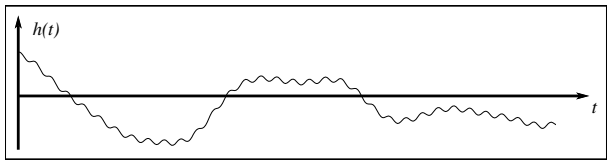


Figure 8.45: "Ringing" in the time domain due to truncation in the frequency domain.

It is clear from these two examples of the use of a rectangular-shaped function that the price we pay for having a sharp truncation of filter coefficients in one domain is excessive "ringing" in the other domain. This, of course, is simply the convolution theorem at work. A rectangle function is necessarily a poor choice for a filter because of the slowly decaying oscillatory nature of the sinc function.

Gaussian filter

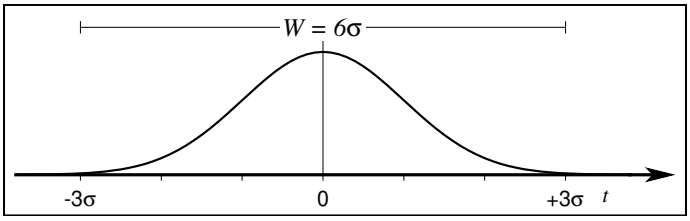


Figure 8.46: Gaussian filter is typically specified by its "full width", defined to be $W = 6\sigma$. While a Gaussian curve never reaches zero weight, its amplitude will have fallen to ~ 1 percent of peak value at $t = \pm 3\sigma$.

While we do desire to find a filter that rapidly tapers off power, say, beyond a certain frequency, we know that in the limit, when the gradual taper approximates a step-function, our transform will approximate a sinc-function. How do we find a good filter? It will depend on the application. Because of the inverse relationship between time and frequency (i.e., $f = 1/T$), we saw that the *scale theorem* (8.118) stated

$$g(at) \leftrightarrow \frac{1}{|a|} G\left(\frac{f}{a}\right).$$

Thus, making a filter broader in one domain narrows it in the other. In the limit ($T \to \infty$) we recover the transform pair

$$1 \leftrightarrow \delta(t).$$

Clearly, somewhere in the middle there must be a function that looks similar in both domains, i.e.,

$$g(t) \leftrightarrow G(f) = g(f).$$

There is such a function; here we will simply state that the Gaussian normal distribution behaves this way. Let

$$g(t) = e^{-\pi t^2}.$$

To find its transform we must follow (8.104) and integrate

$$G(f) = \int_{-\infty}^{+\infty} e^{-\pi t^2} e^{-2\pi i f t} dt = \int_{-\infty}^{+\infty} e^{-(\pi t^2 + 2\pi i f t)} dt.$$

Notice that the exponent is almost of the form $(a+b)^2$. We complete the square by adding and subtracting the missing term:

$$G(f) = \int_{-\infty}^{+\infty} e^{-(\pi t^2 + 2\pi i f t + i^2\pi f^2 - i^2\pi f^2)} dt = \int_{-\infty}^{+\infty} e^{-(\sqrt{\pi}t + i\sqrt{\pi}f)^2} e^{i^2\pi f^2} dt.$$

With $u = \sqrt{\pi}t + i\sqrt{\pi}f$ and $du = \sqrt{\pi}dt$ we get

$$G(f) = \frac{1}{\sqrt{\pi}} e^{-\pi f^2} \int_{-\infty}^{+\infty} e^{-u^2} du = e^{-\pi f^2} = g(f), \tag{8.131}$$

since the definite integral equals $\sqrt{\pi}$. Filters based on the Gaussian function are some of the most used filters in data analysis and data processing. Because of their smooth transform properties we know that they will minimize ringing in the other domain.

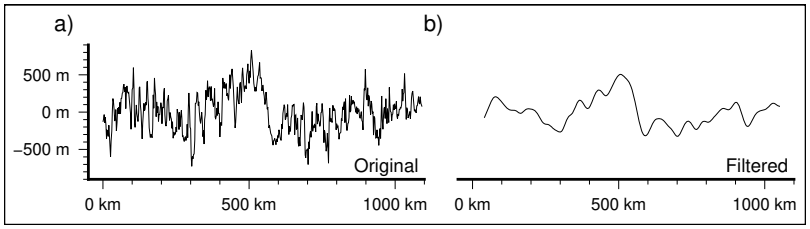


Figure 8.47: (a) Raw topography profile shows considerable short-wavelength noise. (b) Topography after application of a 80-km full-width Gaussian filter. Note we lose 40 km of data at either end, corresponding to the filter's half-width.

Butterworth filter

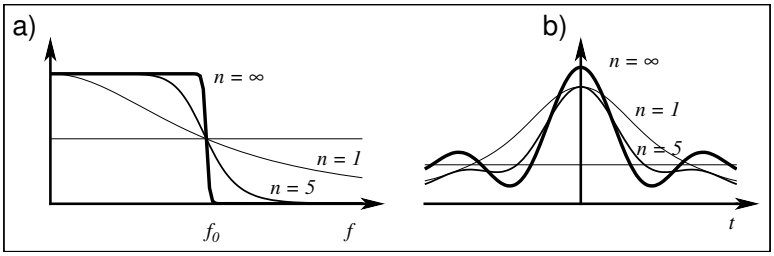


Figure 8.48: Various Butterworth filters for different orders. The higher the order (a), the more ringing will be apparent in the time-domain (b).

Another well-known set of frequency-domain filters has the functional form

$$G(f) = \sqrt{\frac{1}{1+(f/f_0)^{2n}}}. \tag{8.132}$$

These are the so-called *Butterworth* filters and are often used for low-pass filtering purposes. The frequency f_0 defines the halfway point of the filter: The power of $G(f_0)$ is always 1/2, regardless of the exponent $(2n)$. The value of n determines how fast the filter falls off. The higher the value of n, the faster the filter will drop off beyond f_0. We see the sharper the drop off, the more ringing in the time domain, again reflecting Gibbs' phenomenon caused by truncating the spectrum too rapidly. As always, there will have to be a trade-off between how sharply you want to reduce the spectrum and how much ringing you are willing to tolerate.

Wiener filter

Most observed time series can be thought of as a sum of two components:

1. The signal, $u(t)$, that we want to analyze,
2. The noise, $n(t)$.

The measured signal is therefore corrupted by the noise. We will call this the *observed signal*, $c(t)$. In addition, the measuring process may not be able to record all frequencies present in the phenomenon being observed. Hence, the true signal $u(t)$ is "blurred" or smeared, which we can describe as a convolution of $u(t)$ with the known instrument response, $r(t)$. The smeared signal is then

$$s(t) = r(t) * u(t) = \int_{-\infty}^{\infty} r(t-\tau)u(\tau)d\tau \quad \leftrightarrow \quad S(f) = R(f)\cdot U(f). \tag{8.133}$$

Add in the noise component, $n(t)$, and we have

$$c(t) = s(t) + n(t) \quad \leftrightarrow \quad C(f) = S(f) + N(f). \tag{8.134}$$

In the absence of noise we can easily invert (or deconvolve) this equation by solving $U(f) = S(f)/R(f)$. However, with noise a different method must be applied. We want to find the optimal filter, $\phi(t)$ [or $\Phi(f)$] which, when applied to the measured signal $c(t)$ [or $C(f)$], and then deconvolved by $r(t)$ [or $R(f)$], produces a signal $\hat{u}(t)$ [or $\hat{U}(f)$] that is as close as possible to the uncorrupted signal $u(t)$ [or $U(f)$]. In other words, we will estimate the true signal by

$$\hat{U}(f) = \frac{C(f)\cdot\Phi(f)}{R(f)}. \tag{8.135}$$

What do we mean by being "close" to the true signal? We mean in a least-square sense, i.e.,

$$\min \int_{-\infty}^{\infty} [\hat{u}(t) - u(t)]^2 dt \leftrightarrow \min \int_{-\infty}^{\infty} \left[\hat{U}(f) - U(f)\right]^2 df, \tag{8.136}$$

where we have used Parseval's theorem to express the misfit in the frequency domain. With the transform of the noise given by $N(f)$, we substitute (8.134) and (8.135) into (8.136) to find

$$\min \int_{-\infty}^{\infty} \left\{ \frac{[S(f)+N(f)]\cdot\Phi(f)}{R(f)} - \frac{S(f)}{R(f)} \right\}^2 df.$$

Carrying out the square we obtain

$$\min \int_{-\infty}^{\infty} \frac{1}{R^2(f)} \left\{ [S^2(f) + 2S(f)N(f) + N^2(f)]\Phi^2(f) - 2S(f)[S(f)+N(f)]\Phi(f) + S^2(f) \right\} df,$$

which simplifies to

$$\min \int_{-\infty}^{\infty} \frac{1}{R^2(f)} \left\{ S^2(f)[1-\Phi(f)]^2 + N^2(f)\Phi(f) \right\} df. \tag{8.137}$$

This simplification is valid because we assume S and N are uncorrelated, hence their product integrated over all frequencies must equal zero.

Obviously, (8.137) is only minimized if the integrand is minimized for all frequencies. Thus, we can determine the best choice for $\Phi(f)$ by taking the derivative of the integrand with respect to Φ and setting it to zero:

$$-2S^2(f)\,[1-\Phi(f)]+2N^2(f)\Phi(f)=0.$$

Since this identity must hold for all frequencies we solve for the optimal filter as

$$\Phi(f)=\frac{S^2(f)}{S^2(f)+N^2(f)}=\frac{S^2(f)}{C^2(f)}. \tag{8.138}$$

This is the optimal filter known as the *Wiener* filter, named after *Norbert Wiener*. Note that it only involves the power

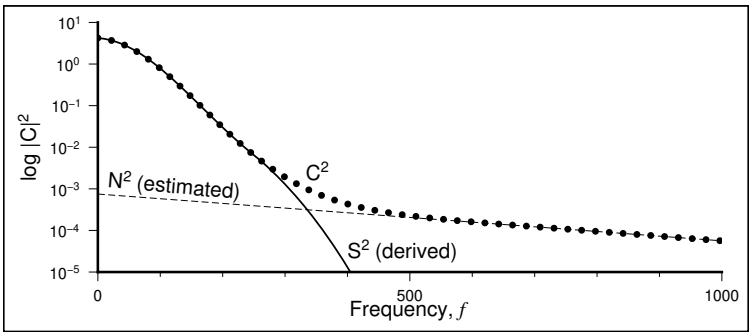


Figure 8.49: An optimal Wiener filter depends on the practitioner's ability to separate the spectrum (C^2) into the contributions from the signal (S^2) and the noise (N^2).

of S, the observed smeared signal, and the power of N, the noise. In particular, equation (8.138) does not contain U, the true signal. This simplifies life: We can determine $\Phi(f)$ independently of $R(f)$. However, we need to isolate S^2 and N^2 from C^2. There is no unique way to do that unless we have some extra information. Luckily, a way out is often presented by looking at the spectrum C^2. Often, it will be clear what shape the signal and noise spectra must have. Consider the situation sketched in Figure 8.49. It appears that the noise spectrum is slightly tilted. We simply extrapolate this trend for all frequencies and subtract it from C^2 to get S^2. We can now numerically form the filter $\Phi(f)$ from S^2 and N^2. As you can see, the filter will be unity where the noise is minimal and drop smoothly to zero where the noise dominates. Simple, but very powerful, and not very sensitive to errors in isolating S^2 and N^2. In fact, a crude separation by eye based on a power spectrum is usually adequate.

8.9.8 Time-domain filters

Many useful filters cannot be specified in the form of a convolution and thus cannot be made extremely efficient via the magic of the convolution theorem and the availability of FFTs. These time-domain (or space-domain) filters must be executed on the data directly and may take considerable computational power.

Median filter

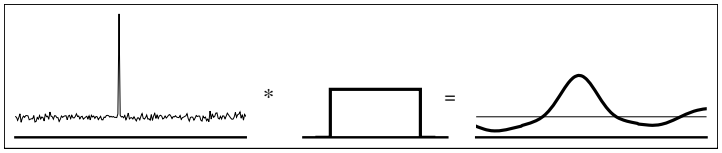


Figure 8.50: Convolution filters have a hard time when data sets contain outliers.

We have assumed for most of the time that all filtering can be described by a convolution, and in general this is the case. The advantage of being able to describe filtering as a convolution is very important: Thanks to the convolution

theorem we can simply transform both data and filter into the frequency domain, perform a multiplication and take the inverse transform. This approach can be vastly faster than computing the convolution in the time domain. Why would we give up that advantage? Consider again a discrete MA filter of a finite width, say

$$g = \left\{ \frac{1}{5} \frac{1}{5} \frac{1}{5} \frac{1}{5} \frac{1}{5} \right\}.$$

When convolved with data, it simply returns the mean value of the points inside the filter width. However, we know the mean is a least squares estimate of "average" value. What happens if we have the occasional bad data point? Because the filter uses an L_2 norm it returns bad values for the entire filter width (see Figure 8.50). Clearly, this is not a desirable result. However, we can design a more *robust* filter by using a more robust estimate of "average" value. A good estimator that is insensitive to the occasional outlier is the *median*, defined as

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}}, & n \text{ is odd} \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n+1}{2}}\right), & n \text{ is even} \end{cases},$$

where it is assumed the x_i have been sorted. Unfortunately, the median is not an analytic function and has no transform

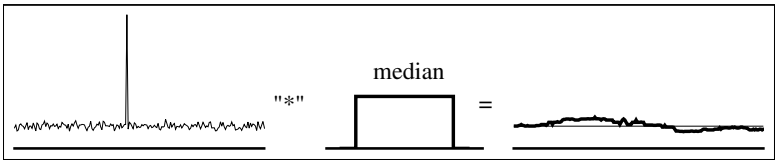


Figure 8.51: Median filters excel in eliminating single or narrow sets of outliers.

(ponder what its "impulse response" would be!). Hence, we are forced to calculate the result of median filtering in the time domain only, as convolution is not applicable. An n-point median filter will, for each lag, return the median value of the current n points. Thus, the filtering of the previous data example would result in the removal of the spike (Figure 8.51). In addition to being robust, i.e., insensitive to outliers, a median filter also preserves step-functions, provided the step length exceeds half the filter width. For data that consist of a noisy signal superimposed on a step-

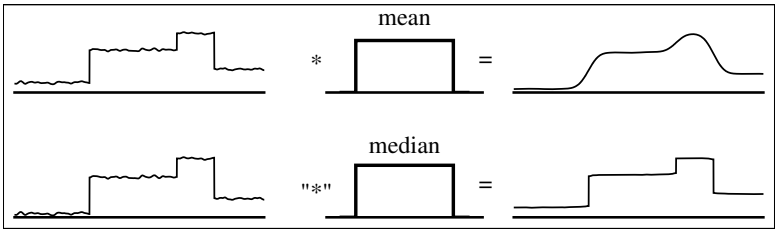


Figure 8.52: Convolution filters (top) will blur sudden steps while median filters (bottom) will pass the steps unchanged.

like background, median filters are very useful (see Figure 8.52). Consider the task of finding the "regional" depth from bathymetry data. The regional depth is "contaminated" by seamounts and faults and may also abruptly increase or decrease across fracture zones (which are approximately step functions due to jumps in crustal age). Finding the median of all points inside the filter width at each output point seems like an expensive operation since the data must be sorted for each output location. Fortunately, a simple iterative scheme may be used to speed up operations. Earlier, we found that the median satisfies the equality

$$\sum_{i=1}^{n} \frac{x_i - \tilde{x}}{|x_i - \tilde{x}|} = 0, \tag{8.139}$$

since this ratio can only take on the values $-1, 0, +1$. With $\tilde{x}$ being defined as having as many values above as below it (and any number of values may be equal to it) the sum must necessarily equal zero. We may rewrite (8.139) as

$$\sum_{i=1}^{n} \frac{x_i}{|x_i - \tilde{x}|} - \sum_{i=1}^{n} \frac{\tilde{x}}{|x_i - \tilde{x}|} = 0,$$

which can be used to determine the median via successive iterations, i.e.,

$$\tilde{x}_k = \frac{\sum_{i=1}^{n} \frac{x_i}{|x_i - \tilde{x}_{k-1}|}}{\sum_{i=1}^{n} \frac{1}{|x_i - \tilde{x}_{k-1}|}}, \tag{8.140}$$

where we obtain the k'th estimate of the median based on the previous value $\tilde{x}_{k-1}$. Initializing $\tilde{x}_0$ to the mid-point between the minimum and maximum data value, this scheme converges fairly quickly (e.g., Figure 8.53). This is especially true for filtering since we would expect the median output from the previous filter position to be similar to the median at the next output location. In fact, the previous filter output is usually a good starting point for the next output position.

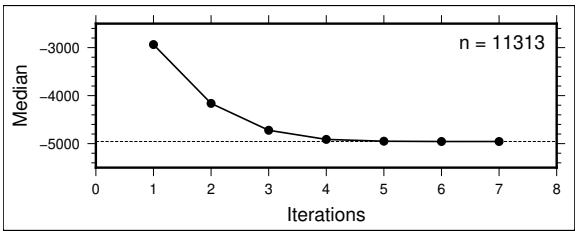


Figure 8.53: We iteratively apply (8.140), and after seven iterations find the median ($\tilde{x} = -4956$) for a bathymetric data set with 11,313 points.

Mode filter

Finally, one can design a filter based on the *mode* rather than the median. Such filters are called *maximum likelihood* filters and return the most frequently occurring value within the filter width. These filters are used to track representative levels through very noisy data. We can implement such a filter using the *Least Median of Squares* (LMS) approximation to the mode, as discussed in great detail in Chapter 6, or one can use a histogram (binning) approach and select the center of the bin with most points. Mode filtering has been used extensively in bathymetric studies to recover regional depths in the presence of seamounts and plateaus.

High-pass, low-pass, and band-pass filters

As their names imply, these are filters that seek to retain a limited part of the frequency content in the signal. Sometimes high-pass is called low-cut and low-pass is called high-cut, for obvious reasons. In the frequency domain, low-pass and high-pass filters are illustrated in Figure 8.54. It is straightforward to design a high-pass filter that is complementary

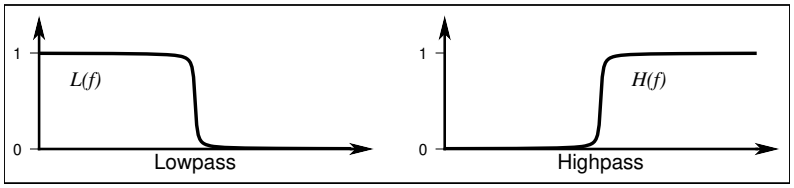


Figure 8.54: (left) A typical low-pass filter will leave the low-frequency components unchanged and attenuate high-frequency components. (right) A high-pass filter is complimentary and reverses what is being attenuated.

to the low-pass filter using

$$H(f) = 1 - L(f). \tag{8.141}$$

In the time-domain, we find the corresponding filter as the inverse transform of $H(f)$:

$$H(f) = 1 - L(f) \leftrightarrow h(t) = \delta(t) - l(t). \tag{8.142}$$

Thus, high-pass filtering the data $d(t)$ in the time domain is described by

$$y(t) = d(t) * h(t) = d(t) * [\delta(t) - l(t)] = d(t) * \delta(t) - d(t) * l(t) = d(t) - d(t) * l(t).$$

This operation simply says that we high-pass-filter a data set by first using a (complementary) low-pass filter and then subtract the output from the original signal. Band-pass filters are simply a linear combination of high-pass and low-pass filters. You can either design $B(f)$ directly or multiply data by $H(f)$ then by $L(f)$ (or two convolutions in the time domain) or let $B(f) = L(f) \cdot H(f)$. Note that (8.142) is not restricted to convolution filters only. Any filter, such as our spatial median and mode filters, that can be applied to data may be used to compute the complementary high-pass filter simply by subtracting the low-passed output from the original observations.

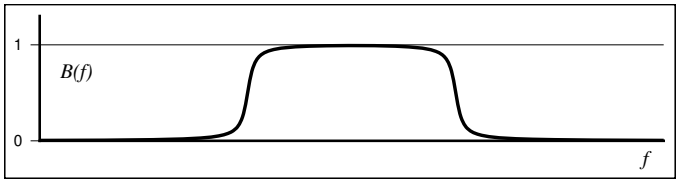


Figure 8.55: A band-pass filter is simply a combination of a low- and high-pass filter with different cutoff frequencies, which results in a certain band of frequencies being relatively unaffected.

8.10 Problems for Chapter 8

Problem 8.1. The monthly flow of water in Cave Creek, Kentucky are given in the file *cavecreek.txt*, with the line number indicating months beginning with October, 1952. Plot the data and compute the power spectrum. What is the Nyquist frequency? Determine the most significant peak. What period does it correspond to? Is the peak significant at the 95% level of confidence?

Problem 8.2. The daily average temperatures in Billings, Montana (in °F) are given in file *Billings.txt*, with the first record corresponding to January 1, 1995. Plot the data and compute the power spectrum. What is the Nyquist frequency for this data set? Find the most significant peak in the spectrum. What period does it reflect, and is the peak significant at the 95% level of confidence?

Problem 8.3. The numbers of sunspots recorded per month for the period 1749–2015 are given in *sspots.txt*. Plot the data and compute the power spectrum. What is the Nyquist frequency for this data set? Find the most significant peak in the spectrum. What period does it reflect, and is the peak significant at the 95% level of confidence?

Problem 8.4. The daily discharge of the Columbia River (in m^3/day) are given in the file *columbia.txt*, covering most of the interval 2001–2008. Plot the data and compute the power spectrum. What is the Nyquist frequency? Determine the most significant peak. What period does it correspond to? Is the peak significant at the 95% level of confidence?

Problem 8.5. Revisit problem 6.7 and remove the linear trends from the data. Use `dft.m` (type `help dft`) to get the amplitudes a_j and b_j for each Fourier series. Make a plot of the raw power spectrum for each time-series using periods rather than frequency on the x-axis (skip the zero frequency which represents the infinite period). What is the Nyquist period? What are the dominant periods in the data? How well do the two series agree on this issue (a qualitative answer is fine).

Problem 8.6. Scientists from Scripps Institution of Oceanography have been measuring the CO_2 concentration on top of Mauna Loa since 1958 (`http://scrippsco2.ucsd.edu`). The table *CO2.txt* shows the record up to 2015 with decimal year versus CO_2 in ppm. The second column contains the raw data, while the third column contains a few interpolated values when raw data were missing (this makes the data continuous and suited to Fourier analysis).

a) Plot the data and determine the least-squares quadratic trend.

b) Remove this trend from the data and subject the residuals to spectral analysis. Determine the two strongest spectral peaks in the signal. What periods do these represent?

Problem 8.7. Revisit the noisy single-period data set *noisy.txt* discussed in Problem 5.7. Remove the least-squares regression line to detrend the data and compute the power spectrum of the residuals. Do you recover the single period you found earlier? Why might you get a different result in this analysis?

Problem 8.8. We will revisit problem 7.7 where we discussed the 3-km long Vostok ice core from Antarctica that resolves temperature variations relative to the present via oxygen isotopes. These data are given in table *vostok.txt*, which contains equidistant depths (in meter), the corresponding times (in year) and the relative change in temperature (in °C). Because of compaction, depth is not a good proxy for time, especially for the deeper (older) sections. To analyze the temporal periodicities we thus need an equidistant time-series. Use MATLAB's `spline` or other software to interpolate the data onto an equidistant interval in time ($\Delta t = 25$). What is the Nyquist frequency for this data set? Determine the most significant peak. What period does it correspond to? Is the peak significant at the 95% level of confidence?

Chapter 9

ANALYSIS OF DIRECTIONAL DATA

"Essentially, all models are wrong, but some are useful."

George E. P. Box, Statistician

Much environmental data contain information about orientations in the plane where the *orientations* and not the lengths of the features are the important attributes. Examples of such data types abound: Strikes and dips of bedding planes, fault surfaces, and joints are familiar from structural geology, augmented by glacial striations, sole marks, lineaments in satellite images, directions of winds and currents, and much more. First, we must distinguish between *directional* and *oriented* data. Directional data can take on unique values in the entire 0–360° range, like drumlins and wind directions, while oriented data consist of "two-headed" vectors: there is a 180° ambiguity inherent in the data. Examples of oriented data include fault traces and other lineaments on maps. Such data require special care in the analysis. However, environmental data involve not only directions in the plane but spatial directions as well, which introduces another degree of complexity. Because 3-D vector data are very common, particularly in the Earth and environmental sciences, we need to examine such data in more detail. We find that much of the measurements used in structural geology, such as strike and dip of fault planes, can be expressed as a normal vector to the fault plane. Other examples include vector measurements of the geomagnetic field, palaeomagnetic measurements, stress directions, wind and current directions, and determinations of crystallographic axes for petrofabric studies.

9.1 Circular Data

9.1.1 Displaying directional distributions

Directional data can be displayed in circular diagrams. One can either plot each direction as a unit vector, or we may count the number of vectors within a given angular sector and draw a polar histogram of the distribution (Figure 9.1):

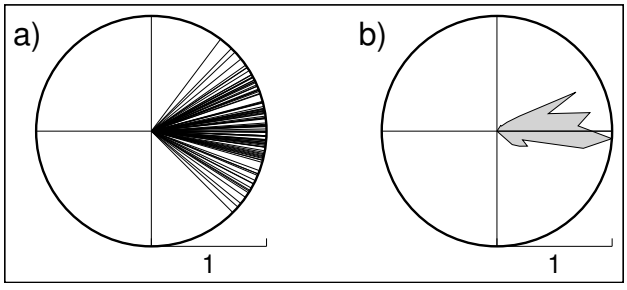


Figure 9.1: Two types of graphical presentations of the same directional data set. (a) Windrose diagram shows all individual directions, (b) Sector diagram shows distribution via a polar histogram (a so-called *rose* diagram).

Unfortunately, the sector diagram as described here is biased in the way it presents the data. Nevertheless, it is still

the most commonly used type of display for directional data. Where is this bias coming from? Consider the area of a sector of width $\Delta\alpha$, given by

$$A = \frac{\pi r^2 \Delta\alpha}{360} \propto r^2. \tag{9.1}$$

We see that the area of a sector is proportional to the radius squared, whereas for a conventional Cartesian histogram the column area is proportional to height, not height squared. Consequently, this leads to a visual distortion of sectors with high counts. Therefore, we should let the sector radius be proportional to the square root of the frequency (or count) to ensure that the final rose diagram will have an area proportional to frequency. If this is not done, the larger counts for some sectors will completely swamp smaller sectors due to the r^2 effect. Thus, small but significant trends may not be detected or may simply be considered noise.

Before we can statistically analyze orientation data, the 180° ambiguity must be accounted for. The simplest way to do this is to double all angles and analyze these doubled angles instead. For instance, if two fault traces are reported as having strikes of 45° and 225° (which means the two strikes have the same orientation), we double the angles and find 90° and $450^\circ - 360^\circ = 90^\circ$. Statistics derived from doubled angles can then be divided by two to recover their original meanings.

The dominant or *mean direction* can be found by computing the *vector resultant* (i.e., the vector sum) of the unit vectors that represent the various directions in the data. Since the coordinates of these vectors are given by

$$x_i = \cos\theta_i, \quad y_i = \sin\theta_i, \tag{9.2}$$

we find the coordinates of the resultant $\mathbf{R} = (x_r, y_r)$ to be

$$x_r = \sum_{i=1}^{n} \cos\theta_i, \quad y_r = \sum_{i=1}^{n} \sin\theta_i. \tag{9.3}$$

The mean direction $\bar{\theta}$ is then simply given by

$$\bar{\theta} = \tan^{-1}(y_r/x_r) = \tan^{-1}\left(\sum_{i=1}^{n} \sin\theta_i / \sum_{i=1}^{n} \cos\theta_i\right). \tag{9.4}$$

Computer implementations of (9.4) should take care to use the `atan2` function rather than `atan` so that the mean direction is found in the correct quadrant. The magnitude R of the resultant should be normalized by the number of vectors before it can be used further. This gives us the *mean resultant length*

$$\bar{R} = \frac{R}{n} = \frac{\sqrt{x_r^2 + y_r^2}}{n}, \tag{9.5}$$

which ranges from 0 to 1 and is a measure of (normalized) dispersion analogous to the variance. Since it increases for focused distribution we often convert it to a *circular variance* via

$$s_c^2 = 1 - \bar{R} = (n - R)/n. \tag{9.6}$$

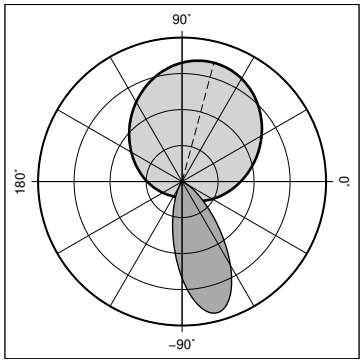


Figure 9.2: The circular von Mises distribution. One population is centered on $\mu = 75$ (dashed line), with very low directionality ($\kappa = 1$), and shows the effect of wrapping the wide distribution around the full circle. The other, at $\mu = -75$, is much more directional ($\kappa = 10$).)

To test various statistical hypotheses about circularly distributed data we need a probability density function that can be used to obtain critical values. A traditional distribution that has been used extensively in studies of directional data is the *von Mises* distribution, given by

$$p(\theta) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, \tag{9.7}$$

where κ is a measure of the *concentration* of the distribution about the mean direction μ, and I_0 is the modified Bessel function of the first kind and order zero (it normalizes the cumulative distribution of $p(\theta)$ over 360° to unity.) A large κ means we have a strongly preferred direction. The concentration parameter κ is obviously related to R (or s_c^2) and this relationship can be derived if we assume that the data represent a random sample from a population having a von Mises distribution. Statistical tables (e.g., Table A.16) report solutions to (A.2) that relate κ and $\bar{R}$.

9.1.2 Test for a random direction

The simplest test for circular data is to determine whether the directional observations are random or not, i.e., that there is no *preferred direction*. In that case we have a *uniform* distribution. In terms of the von Mises distribution (9.7), a uniform distribution means that κ must be zero. Hence, the hypothesis test becomes

$$H_0 : \kappa = 0 \quad \text{versus} \quad H_1 : \kappa > 0. \tag{9.8}$$

If the data came from a uniform distribution (i.e., κ = 0), we would expect a small but not necessarily zero value for $\bar{R}$. Lord Rayleigh devised this test, which gives critical values for $\bar{R}$ depending on the $\nu = n$ degrees of freedom and chosen level of significance, α. For $\nu \geq 10$ we can approximate the critical mean resultant via

$$\bar{R}_{\alpha,n} = \frac{1}{2n}\sqrt{1 + 4n(n+1) - (\log\alpha + 2n + 1)^2}, \tag{9.9}$$

while for smaller samples we need to consult exact tables (e.g., Table A.17). If $\bar{R}$ exceeds the critical $\bar{R}_{\nu,\alpha}$, then we must reject the null hypothesis.

Example 9–1. We have measured the strikes of fault planes in an area of the sea floor (called Area 1) imaged by a side-scan sonar device and found the following orientations:

$$\theta_i : 110^\circ, 300^\circ, 310^\circ, 135^\circ, 138^\circ, 320^\circ, 141^\circ, 145^\circ, 330^\circ, 335^\circ, 280^\circ, 160^\circ, 170^\circ \quad n = 13.$$

Since these are oriented features, we double the angles to remove any ambiguity in the orientations:

$$2\theta_i : 220^\circ, 240^\circ, 260^\circ, 270^\circ, 276^\circ, 280^\circ, 282^\circ, 290^\circ, 300^\circ, 310^\circ, 200^\circ, 320^\circ, 340^\circ.$$

Summing up the sines and cosines via (9.2), we find

$$x_r = 1.2972, \quad y_r = -10.349, \tag{9.10}$$

which gives us

$$2\bar{\theta}_1 = 277^\circ, \quad \bar{R}_1 = 0.802. \tag{9.11}$$

We convert back to original angles by dividing by 2 and find the mean orientation to be

$$\bar{\theta}_1 = 138.5^\circ. \tag{9.12}$$

The null hypothesis is not affected by doubling the angles since we are simply testing if $\kappa = 0$. At the $\alpha = 0.05$ level of significance we use Table A.17 to find critical $\bar{R}_{13,0.05} = 0.475$. Clearly, this value is greatly exceeded by the observed $\bar{R} = 0.802$ and we must conclude that the fault strikes are not randomly oriented.

9.1.3 Test for a specific direction

There are instances when we would like to test whether an observed trend equals a specified trend. The determination of critical values for such a scenario is more difficult and involves using complex equations or charts. As an alternative approach we may find the *confidence angle* $\Delta\theta$ around the mean direction of the sample. We can then investigate whether this angular interval is large enough to accommodate the specified trend we want to test against. The approximate standard error in $\bar{\theta}$ is given (in radians) by

$$s_e \approx 1/\sqrt{n\bar{R}k} = 1/\sqrt{Rk}, \tag{9.13}$$

where k is our sample estimate of κ, the population parameter. If we assume that the estimation errors are normally distributed then we may use critical z-values for a given confidence level to construct the angular interval as

$$\bar{\theta} \pm \left| z_{\frac{\alpha}{2}} \right| s_e, \tag{9.14}$$

which will contain the true mean (double angle) orientation, $2\mu, 100 \cdot \alpha$ % of the time.

Example 9–2. In our fault strike data case above, it is believed that the faults arose as tensional features in response to tectonic forces operating in the N30°E direction. Under such a stress regime we would theoretically expect the faults to trend at 90° to the direction of these tensile stresses, i.e., in the N120°E direction. From Table A.16 we find $\kappa = 2.897$, which yields

$$s_{e_2} = \frac{1}{\sqrt{13 \cdot 0.802 \cdot 2.897}} = 0.1820 = 10.4^\circ. \tag{9.15}$$

Note that this is the standard deviation about the *doubled* mean angle orientation of 277°. We divide this deviation by 2 to find the standard deviation associated with the original orientations, $s_{e_1} = 5.2^\circ$. The 95% confidence interval corresponds to $z_{0.95} = 1.96$, so the confidence interval around $\bar{\theta}_1$ becomes

$$\bar{\theta}_1 = 138.5^\circ \pm 10.2^\circ. \tag{9.16}$$

Since the predicted direction of 120° is outside this interval we conclude that the observed mean direction deviates from that predicted based on the orientation of current tectonic forces. It is possible that the orientation of stresses may have changed since the formation of the faults.

9.1.4 Test for equality of two mean directions

Another common situation where we may want to apply a statistical test is to determine if two sample mean directions are equal at some prescribed level of confidence. This situation, of course, is reminiscent of the two-sample mean test for scalars that we described in Section 4.2.2. For example, we may have obtained new side-scan sonar data for an area farther away and would like to test whether the mean direction of fault strikes at the second location is similar to what was found at the first location. We can carry out this test by comparing the vector resultants of the two groups to that produced by pooling the two data sets and recalculate a single grand resultant. The key principle here is that if the mean directions are different, then the pooled resultant should be *shorter* than the sum of the two individual resultants. This expectation can be examined using an F-test by comparing observed F to critical $F_{1,n-2,\alpha}$. Unfortunately, due to approximations needed to relate Cartesian and circular critical values, the procedure to evaluate the F-statistic changes with the dispersion parameter, κ, represented by our sample-based value, k. For large $k > 10$, we may compute

$$F = \frac{(n-2)(R_1 + R_2 - R_p)}{n - R_1 - R_2}, \tag{9.17}$$

with n being the combined number of observations, R_1 and R_2 are the full resultants (not *mean* resultants) for each data set, and R_p is the resultant from the combined data. Here, k is estimated from $\bar{R}_p$, the mean resultant. However, if $2 < k < 10$ then a more accurate F-statistic is obtained via the adjustment

$$F^* = \left(1 + \frac{3}{8k}\right) F, \tag{9.18}$$

while for even smaller values of k special tables must be used.

Example 9–3. We will use the equality test to determine whether the mean strike of the faults in our second area (Area 2) is the same as what was found for the first area ($\bar{\theta}_1 = 138.5^\circ$, double angle = 277°). The new data are:

$$\theta_i = 91^\circ, 280^\circ, 111^\circ, 115^\circ, 118^\circ, 300^\circ, 122^\circ, 126^\circ, 130^\circ, 80^\circ, 320^\circ, 149^\circ \quad (n = 12),$$

which we double to get

$$2\theta_i = 182^\circ, 200^\circ, 222^\circ, 230^\circ, 236^\circ, 240^\circ, 240^\circ, 244^\circ, 252^\circ, 260^\circ, 160^\circ, 280^\circ, 298^\circ.$$

Carrying out a similar analysis on our new data we now find $2\bar{\theta}_2$ = 234.5° and $\bar{R}_2$ = 0.805. The hypothesis becomes $H_0 : \bar{\theta}_1 = \bar{\theta}_2$ versus $H_1 : \bar{\theta}_1 \neq \bar{\theta}_2$, at $\alpha = 0.05$. For the combined data set (with $n = 25$), we obtain

$$2\bar{\theta}_p = 256.7^\circ, \quad \bar{R}_p = 0.749. \tag{9.19}$$

We find observed k from Table A.16 to be 2.36. Hence, we use (9.18) to obtain

$$F = \left(1 + \frac{3}{8 \cdot 2.36}\right) \frac{23\,(13 \cdot 0.802 + 12 \cdot 0.805 - 25 \cdot 0.749)}{25 - 13 \cdot 0.802 - 12 \cdot 0.805} = 7.42. \tag{9.20}$$

Table A.5 gives the critical $F_{1,n-2,\alpha}$ value for 1 and 23 degrees of freedom as 4.28. Therefore, we must reject the null hypothesis and conclude that the fracture directions in the two areas appear to have different orientations at the 95% confidence level.

9.1.5 Robust directions

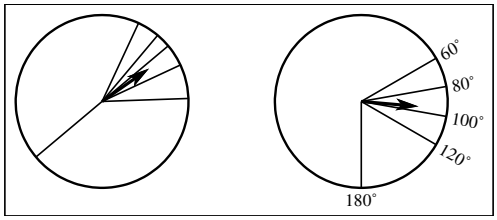


Figure 9.3: The effect of outliers on the mean direction is most severe when the outlier is orthogonal to the bulk of the data directions. Headed vectors represent the mean resultants.

The concept of robustness is also applicable to directional data since such data may also contain outliers. However, directional outliers cause less harm than their Cartesian counterparts discussed earlier since directional data are forced to be periodic. It is clear from Figure 9.3 that an outlier causes most damage to estimates of mean direction when it is oriented 90° away from the trend of the bulk of the data. We may find the circular analog of the median by finding the direction $\tilde{\theta}$ that minimizes the sum

$$\text{minimize} \sum_{i=1}^{n} d\left(\theta_i, \tilde{\theta}\right), \tag{9.21}$$

where $d\left(\theta_i, \tilde{\theta}\right)$ is the arc distance between θ_i and $\tilde{\theta}$ measured along the perimeter of the unit circle. Since this distance is proportional to $|\,\theta_i - \tilde{\theta}\,|$ it is the equivalent of minimizing the sum

$$\text{minimize} \sum_{i=1}^{n} |\,\theta_i - \tilde{\theta}\,|, \tag{9.22}$$

which we know gives the median of the θ_i data set. Similarly, a LMS mode estimate can be found by determining the midpoint of the shortest arc containing $n/2 + 1$ points. Again, this would involve sorting the directions first, possibly

after doubling orientations.

Example 9–4. Consider the wind directions

$$\theta_i : 75^\circ, 85^\circ, 90^\circ, 98^\circ, 170^\circ \quad \bar{\theta} = 100.7^\circ.$$

The shortest arc over $5/2+1 = 3$ points is located between 85° and 98°, giving the mode estimate $\hat{\theta} = 91.5^\circ$. This estimate suggests that 170° is an outlier with respect to the rest of the atmospheric data.

9.1.6 Data with length and direction

In the analysis so far we have only considered the direction (or orientation) of a feature and not its length. However in many cases, such as fracture data or fault traces, the features will have very different lengths. The analysis described above, when applied to such data, would give both a one km long and a 100 km long fault the same weight, which does not seem to be appropriate. We can account for this bias by weighing the directions by the respective *lengths* of the features. By keeping track of the length of the faults (and not their numbers) per sector we can obtain a rose diagram that reflects the proportions of the various fracture directions. The rationale employed here is that large and/or long faults may be more representative of the tectonic stresses than a few short fractures. The rose diagram may then be normalized by the total length of the fractures to give overall proportions in percent.

Staying with fault strike data, it is clear that in many regions the faults are not entirely straight lines but may actually curve or bend. From a directional analysis point of view, such fractures must first be approximated by shorter straight line segments. It then becomes obvious that we must weigh the pieces by their lengths, otherwise the directional frequencies would depend on the number of pieces used. Fault traces must therefore be digitized prior to analysis on a computer. A typical fault trace is displayed in Figure 9.4. Depending on the angular width of the polar histogram,

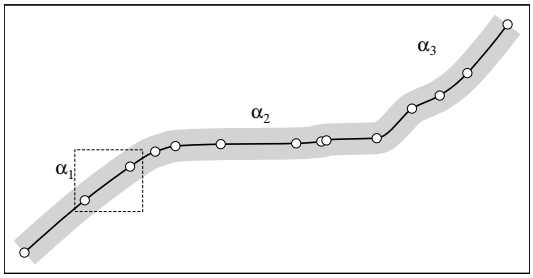


Figure 9.4: A digitized fault trace. Open circles indicate the digitized points. Digitizing too finely may lead to short-wavelength noise in the representation of the fault. Points within the dashed box are examined in Figure 9.5.

the digitized fault may end up in two bins corresponding to the orientations α_1 and α_2. We must also be aware of the fact that the digitizing process will introduce uncertainties in the digitized points. To see how this may affect the analysis, consider the line segment in Figure 9.5. We may assume that the exact position of each point is uncertain, here represented by the one-sigma uncertainty estimate s_r in radial position (gray circles) for each point. There are two points of interest here: First, the length of the segment, d, will have an uncertainty since it reflects a difference between two uncertain values. In Chapter 2, we found this to be

$$\Delta d = \sqrt{2 s_r}, \quad d = d_0 \pm \Delta d, \tag{9.23}$$

assuming the uncertainties are *independent*. Second, we see that the direction (or orientation) α may be in error by $\pm\Delta\alpha$, given by

$$\Delta\alpha = \tan^{-1}(2s_r/d), \quad \alpha = \alpha_0 \pm \Delta\alpha. \tag{9.24}$$

Consequently, one should not digitize lines so frequently that $\Delta\alpha$ exceeds the desired polar histogram interval. For instance, if this interval is 10° then you are best served by making the average digitizing interval $d > 10 s_r$. Alternatively, one can *filter* the digitized track so that short-wavelength noise is smoothed out before binning the segments.

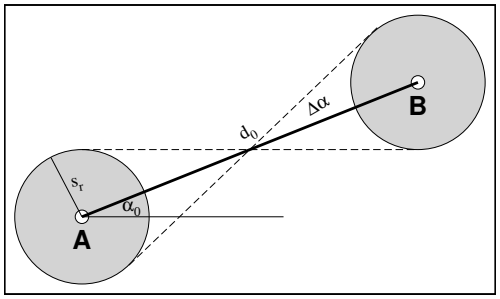


Figure 9.5: The uncertainty r_s in the locations of two digitized points (A and B) from Figure 9.4 introduce uncertainties Δd and $\Delta\alpha$ in the length (d_0) and orientation (α_0) of a line segment, respectively.

The length of all segments that have a direction within the width of a single bin is simply computed by adding up all the individual lengths. Note that since internal nodes are shared by adjacent line segments the uncertainty in the total length is independent of the number of line segments (e.g., Figure 2.2) and only depends on (9.23). However, the *sum* of the lengths of the individual segments will have an uncertainty that is cumulative since the segments are no longer connected. This method will provide a frequency distribution with error bars in which directions with many small segments will have higher uncertainty than directions with fewer and longer faults.

9.2 Spherical Data Distributions

The analysis of 3-D directions and orientations is an extension of the methods used for circular data. It is common to require that these 3-D vectors have unit lengths so that their endpoints all lie on the surface of a sphere with unit radius — hence the name spherical distributions. Similar to the 2-D case for circular data, there will be spherical data that only reflect orientations (i.e., *axes*) rather than directions. We need to use a 3-D Cartesian coordinate system to

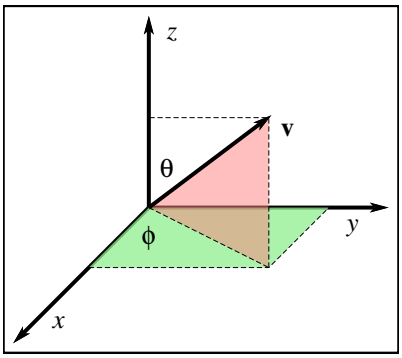


Figure 9.6: The relation between the Cartesian (x,y,z) and spherical (ϕ,θ,r) coordinate systems.

describe the unit vectors (Figure 9.6). Thus, any vector $\mathbf{v}$ is uniquely determined by the triplet (x,y,z). We could also use spherical angles θ (colatitude) and ϕ (longitude) to specify the vector direction, assuming the length $r = 1$. We relate the Cartesian coordinates and the spherical angles as follows:

$$\begin{aligned} x &= \sin\theta\cos\phi, \\ y &= \sin\theta\sin\phi, \\ z &= \cos\theta. \end{aligned} \tag{9.25}$$

However, geological measurements like *strike* and *dip* are more commonly used than Cartesian coordinates and spherical angles and these follow their own convention. We define a new local coordinate system in which x points toward north, y points east, and z points vertically down (i.e., in order to maintain a right-handed coordinate system). In such a system, fault plane dips are expressed as positive angles. For the fault plane in Figure 9.7 we find that the angle A is

the azimuth of the strike of the plane and D is the dip, measured positive down. The slip-vector OP is then given by its components

$$\begin{aligned} x &= -\sin A \cos D, \\ y &= \cos A \cos D, \\ z &= \sin D. \end{aligned} \tag{9.26}$$

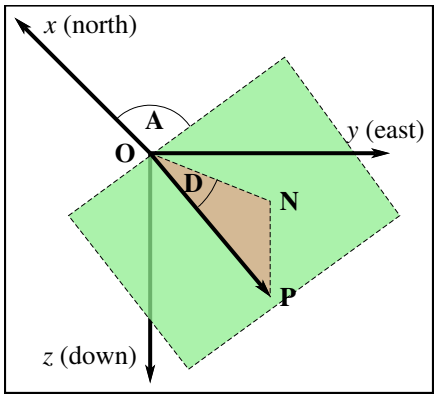


Figure 9.7: Local, right-handed coordinate system shows the convention used in structural geology. Here, A is the strike (measured from north over east) of the dipping plane (light green) and D is its dip, measured as the angle between the dip vector OP and its projection ON onto the horizontal $x-y$ plane. The red plane containing the dip vector is orthogonal to the light green plane.

Once we have converted our (A,D) data to (x,y,z) we can compute such quantities as mean direction and spherical variance, which are simple extensions of the 2-D or directional analogs. The length of the resultant vector is simply

$$R = \sqrt{\left(\sum x_i\right)^2 + \left(\sum y_i\right)^2 + \left(\sum z_i\right)^2}, \tag{9.27}$$

with the sum taken over all the n points. The resultant vector is usually normalized to give $\bar{R} = R/n$. The coordinates $\bar{x}, \bar{y}$ and $\bar{z}$ of the mean vector $\mathbf{m}$ are then obtained via

$$\bar{x} = \sum x_i/n, \quad \bar{y} = \sum y_i/n, \quad \bar{z} = \sum z_i/n, \tag{9.28}$$

so that the mean slip-vector is given by its two components

$$\begin{aligned} \bar{D} &= \sin^{-1}\bar{z}, \\ \bar{A} &= \tan^{-1}(-\bar{x}/\bar{y}). \end{aligned} \tag{9.29}$$

If all the vectors are closely clustered then the resultant R will approach n, but if the vectors are more randomly distributed then R will approach zero. As in 2-D, we can use R (or $\bar{R}$) as the basis for the *spherical variance*, s_s^2, give by

$$s_s^2 = (n-R)/n = 1-\bar{R}. \tag{9.30}$$

9.2.1 Test for a random direction

We can perform simple hypothesis tests in a manner analogous to those we carried out for 2-D directional data. Unlike the case for 2-D we now require a *spherical* probability density function from which we can derive critical values for comparison with our observed statistics. In response to the need for 3-D statistical analysis, in particular for palaeomagnetic studies that started to explore "continental drift" in the 1950s, the famous British statistician Ronald Fisher developed a suitable theoretical distribution for spherical data. His probability density function has since been called the *Fisher* distribution on a sphere and is given by

$$P(\mathbf{x}) = \frac{\kappa}{4\pi \sinh \kappa} e^{\kappa(\mathbf{x}\cdot\boldsymbol{\mu})}, \tag{9.31}$$

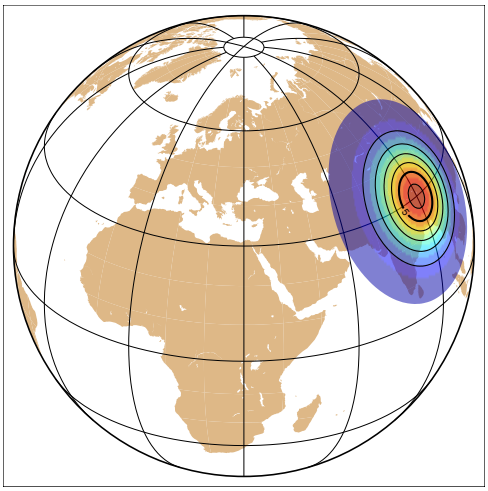


Figure 9.8: Well-focused Fisher distribution on a sphere, centered on a point in northern India, with $\kappa = 40$.

where the dot product between the mean direction $\boldsymbol{\mu}$ and any other direction $\mathbf{x}$ equals the cosine of the angle ψ between them, κ is the precision parameter similar to the one we encountered for the von Mises distribution on the circle, and sinh is the *hyperbolic sine* function (which is needed to ensure the cumulative distribution of $P(\mathbf{x})$ over the unit sphere equals one).

Fisher showed one can estimate κ from the sample, provided $n > 7$ and $\kappa > 3$. The estimate is then given by

$$k \approx \kappa = \frac{n-1}{n-R}, \tag{9.32}$$

but more accurate tables also exist (e.g., Table A.18) which solve (A.3).

Testing a spherical distribution for randomness follows the same approach used for directional data: We must first evaluate $\bar{R}$, the mean resultant. Then, we state the null hypothesis to be

$$H_0 : \kappa = 0 \quad H_1 : \kappa > 0. \tag{9.33}$$

As before, this test is executed by establishing critical values of $\bar{R}$ given the prescribed level of confidence, α. Table A.19 shows such critical $\bar{R}$ values for selected values of α and n.

Example 9–5. Let us examine a set of palaeomagnetic measurements. We have been given six measurements of *declination* and *inclination*, reported as

Dec	105°	130°	115°	120°	118°	145°
Inc	40°	49°	57°	32°	55°	45°

First, we note that a dip-vector specified by strike A and dip D is not using the same geometry as a magnetic field vector given by its magnetic declination D and inclination I. Because the projection of the dip-vector onto the horizontal plane is 90° away from the strike, we must use a slightly modified conversion suitable for magnetic vectors to obtain the Cartesian coordinates:

$$\begin{aligned} x &= \cos D \cos I, \\ y &= \sin D \cos I, \\ z &= \sin I. \end{aligned} \tag{9.34}$$

We convert the $n = 6$ observed declinations and inclinations to x, y, z and find

x	−0.20	−0.42	−0.23	−0.42	−0.27	−0.58	$\bar{x}$	−0.354
y	0.74	0.50	0.49	0.73	0.51	0.41	$\bar{y}$	0.564
z	0.64	0.75	0.84	0.53	0.82	0.71	$\bar{z}$	0.715

Computing the mean direction and resultant gives

$$\bar{I}_1 = \sin^{-1}\bar{z} = 45.7^\circ, \tag{9.35}$$

$$\bar{D}_1 = \tan^{-1}\bar{y}/\bar{x} = 122.1^\circ, \tag{9.36}$$

$$\bar{R}_1 = \frac{1}{6}\cdot 5.86 = 0.977, \tag{9.37}$$

$$k_1 = \frac{6-1}{6-5.86} = 35.7. \tag{9.38}$$

It is evident that our distribution has a clear preferred direction since k_1 is so large (the critical $\bar{R}$ for $\alpha = 0.05$ is only 0.642).

9.2.2 Test for a specific direction

Often, we will be interested in testing whether the observed mean direction equals a prescribed direction, given the uncertainties due to random errors. As for circular data, such tests are best performed by constructing the α confidence region around the mean direction. This statistic is based on the Fisher distribution and gives a spherical cap radius for a *cone of confidence* around the mean direction. As usual, this radius is a function of both the confidence level α and R. We find

$$\delta_{1-\alpha} = \cos^{-1}\left\{1 - \frac{n-R}{R}\left[\left(\frac{1}{\alpha}\right)^{\frac{1}{n-1}} - 1\right]\right\}. \tag{9.39}$$

This fairly complicated expression simplifies considerably if we assume (or actually know) that $k > 7$ and standardize our tests for $\alpha = 0.05$. Then,

$$\delta_{95\%} \approx 140^\circ/\sqrt{kn} \tag{9.40}$$

given in spherical degrees. We may now say that there is a 95% probability that the true mean direction lies within the cone of confidence specified by the angular radius $\delta_{95\%}$.

The application of the cone of confidence is more involved than for circular data since we cannot directly compare the two angular measures defining the mean direction (e.g., strike, dip) to the comparable quantities of a specific direction direction to be tested, here called $\hat{\mathbf{v}}$. Instead, we need to utilize their Cartesian vector representations. The procedure is still relatively straightforward:

1. State the null hypothesis that the unit vectors are the same, i.e., $H_0 : \hat{\mathbf{m}} = \hat{\mathbf{v}}$.
2. Take the dot-product of the mean unit vector with the test unit vector, i.e, $c = \hat{\mathbf{m}}\cdot\hat{\mathbf{v}}$.
3. Since a dot product gives the cosine of the spherical angle between two vectors we find this distance to be given by $\psi = \cos^{-1} c$.
4. If ψ exceeds $\delta_{95\%}$ then $\hat{\mathbf{v}}$ is outside the cone of confidence and we can reject the null hypothesis; otherwise we must reserve judgment.

For other levels of confidence we would substitute the general radius specified in (9.39).

Example 9–6. Given the data we just analyzed we want to test if the mean vector we found is compatible with an hypothesis that says the inclination should be 45°and the declination should be 120°. First, we find the unit vector pointing in the same direction as the mean resultant, $\mathbf{m}$. This vector is

$$\hat{\mathbf{m}} = \frac{\mathbf{m}}{|\mathbf{m}|} = \frac{(-0.354, 0.564, 0.715)}{\sqrt{0.354^2 + 0.564^2 + 0.715}} = (-0.3621, 0.5770, 0.7321). \tag{9.41}$$

We evaluate the test vector to be

$$\hat{\mathbf{v}} = (\cos 120^\circ\cdot\cos 45^\circ, \sin 120^\circ\cdot\cos 45^\circ, \sin 45^\circ) = (-0.3536, 0.6124, 0.7071). \tag{9.42}$$

Hence, the angle between the mean and the test vector is given by their dot product:

$$\psi = \cos^{-1}(0.3621 \cdot 0.3536 + 0.5770 \cdot 0.6124 + 0.7321 \cdot 0.7071) = \cos^{-1}(0.9990) = 2.53^\circ. \tag{9.43}$$

To determine the radius of the 95% confidence cone we use (9.40) and find

$$\delta_{95\%} \approx 140^\circ/\sqrt{35.7 \cdot 6} = 9.6^\circ. \tag{9.44}$$

This means that the true population mean direction probably (i.e., with 95% level confidence) lies within a spherical cone of radius 9.6° centered on the observed mean direction. Clearly, our test vector $\mathbf{v}$ also lies inside this cone. Therefore, we cannot reject the null hypothesis that they point in the same direction.

9.2.3 Test for equality of two mean directions

This test is equivalent to the one administered for circular data, relying on the same F-statistics and equations; see Section 9.1.4 for details.

Example 9–7. We also obtained palaeomagnetic measurements from a nearby second site:

Dec	65°	72°	51°	75°	50°	45°
Inc	25°	20°	30°	23°	18°	33°

We perform the conversion to Cartesian components and find

x	0.38	0.29	0.55	0.24	0.61	0.59	$\bar{x}$	0.443
y	0.82	0.89	0.67	0.89	0.73	0.59	$\bar{y}$	0.766
z	0.42	0.34	0.50	0.39	0.31	0.54	$\bar{z}$	0.418

which gives a resultant $R = 5.88$ and the mean resultant (and mean declination, inclination) as

$$\bar{R}_2 = 0.979, \quad \bar{I}_2 = 24.7^\circ, \quad \bar{D}_2 = 59.9^\circ, \quad n = 6. \tag{9.45}$$

We would like to compare these two mean directions to determine if they are statistically equivalent at the 95% confidence level, i.e.,

$$H_0 : \bar{I}_1 = \bar{I}_2, \quad \bar{D}_1 = \bar{D}_2. \tag{9.46}$$

The hypothesis test for this equality is exactly the same test we used for 2-D directional data. Again, we use an F-test to determine if the resultant from the pooled data is significantly different from the linear sum of the two individual resultants, i.e.,

$$F = \frac{(n-2)(R_1 + R_2 - R_p)}{n - R_1 - R_2}, \tag{9.47}$$

where again $n = n_1 + n_2$. We need to find the resultant for the combined data set of 12 points. It is found via (9.27) and gives $R_p = 10.50$. We may then normalize by n to find $\bar{R}_p = 0.875$. Our observed F-statistic thus becomes

$$F = \frac{10(5.86 + 5.88 - 10.50)}{12 - 5.86 - 5.88} = 47.4. \tag{9.48}$$

The observed F value by far exceeds the critical $F_{0.05,1,10} = 4.96$ (i.e., Table A.5) and we must reject the null hypothesis that the two directions are the same.

9.3 Problems for Chapter 9

Problem 9.1. Analyze the current directions in surface waters off Hog Neck, MA for a two-day period, as listed in *HogNeck.txt*. How do the current directions vary during a 24-hour period? Do the current speeds show a similar trend?

Problem 9.2. Paleocurrent measurements from ripples in sandstones contain information on past current directions.

a) Use the data in *ripple_A.txt* from sandstone formation A to test if there is a preferred direction in the data. A local sedimentologist bravely suggests that the ripples were formed by long-shore currents along a paleo-coastline trending 200°. Do the data support this hypothesis?

b) A second sandstone formation (B) was also examined for paleocurrents (*ripple_B.txt*). Do these data suggest there was any change in current direction between the formations of these two layers?

Problem 9.3. A geologist surveys a large section of an exposed batholith and measures the strike direction of numerous vertical joints. The file *joints.txt* contains the azimuths (measured from north toward east) of these orientations.

a) At the 95% level of confidence, is there a preferred orientation in the data? If so, what is the preferred orientation?

b) Regional tectonic considerations seem to favor a general extensional stress regime in the west-northwest – east-southeast orientation. Is this explanation for the joints consistent with the data at the 95% level of confidence?

c) 50 km further north another exposure of the batholith reveals additional joints (*morejoints.txt*). Are they randomly oriented?

d) Do these new joints deviate significantly from the preferred orientation you determined for the first site?

Problem 9.4. A structural geologist has measured the strikes and dips of faults in a Jurassic sedimentary section. The observations (in degrees from north toward east) are recorded below. Find the mean direction and the mean resultant length. Is this direction significant at the 95% level of confidence?

Strike	142°	157°	145°	150°	149°	156°	−13°	139°	155°	−9°	148°	142°
Dip	45°	50°	52°	38°	56°	44°	8°	49°	46°	11°	59°	43°

Problem 9.5. Principal fracturing directions for three oil fields near Odessa, TX are given in *odessa_north.txt*, *odessa_northwest.txt* and *odessa_west.txt*. Compare the mean directions of fractures from the three fields. Are they significantly different at the 95% level of confidence?

Problem 9.6. Measurements of magnetic remanence (given as declination and inclination) from a sandstone in Western Australia are listed in *Tumblagooda.txt*.

a) Compute the mean direction and the mean resultant length.

b) Estimate the angular radius of the 95% confidence cone.

c) Is this direction significant at the 95% level of confidence?

Problem 9.7. Scientific drilling in the Caribbean Sea on-board the *Glomar Challenger* resulted in a set of paleomagnetic data (reproduced in *glomar.txt*).

a) Determine if these data (declination clockwise from north, inclination downward from horizontal) have a preferred direction at the 95% level of confidence.

b) Compare this mean paleomagnetic vector to the present field at this site (declination = 354.3°and inclination = 46.4°). Do these two directions differ significantly at the 95% level of confidence? What does your analysis suggest for the plate tectonic motions for this region?

Appendix A

STATISTICAL TABLES

"The main purpose of a significance test is to inhibit the natural enthusiasm of the investigator."

Frederick Mosteller, Statistician

This appendix contains a set of standard statistical tables of critical values for a variety of tests or distributions at selected levels of confidence. When appropriate, both one-sided and two-sided critical values are provided. A two-sided test is used when the null hypothesis is asserting an *equality*, while a one-sided test is appropriate when the null hypothesis is stating an *inequality*. Reject the null hypothesis if the calculated statistic *exceeds* the critical value (except for the U-test where rejection of H_0 should result when the calculated U is *less* than the critical value.)

Test or Distribution	Table
Normal distribution	A.1
Student's t	A.2
$\chi^2_{\alpha,\nu}$	A.3
$F_{0.9,\nu_1,\nu_2}$	A.4
$F_{0.95,\nu_1,\nu_2}$	A.5
$F_{0.975,\nu_1,\nu_2}$	A.6
$F_{0.99,\nu_1,\nu_2}$	A.7
$U_{0.05,n_1,n_2}$ (one-sided)	A.8
$U_{0.01,n_1,n_2}$ (one-sided)	A.9
$U_{0.05,n_1,n_2}$ (two-sided)	A.10
$U_{0.01,n_1,n_2}$ (two-sided)	A.11
Kolmogorov-Smirnov (one sample, given μ,σ)	A.12
Kolmogorov-Smirnov (one sample, estimate μ,σ) (Normally called Lilliefors test for normality)	A.13
Kolmogorov-Smirnov (two sample)	A.14
Spearman's rank correlation	A.15
Determine κ from $\bar{R}$ (two-dimensional)	A.16
Lord Rayleigh test for $\bar{R}$ (two-dimensional)	A.17
Determine κ from $\bar{R}$ (three-dimensional)	A.18
Fisher test for $\bar{R}$ (three-dimensional)	A.19

A.1 Cumulative Probabilities for the Normal Distribution

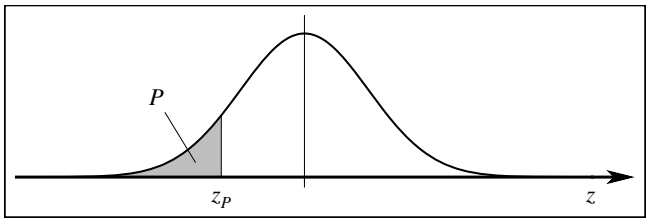


Figure A.1: Given a chosen z_P-value, the probability P (gray area under the curve from $-\infty$ to z_P) can be read from this table. Here, z_P is given in the format *-a.bc*, where *-a.b* and *0.0c* correspond to a unique row and column combination.

z_P	**0.09**	**0.08**	**0.07**	**0.06**	**0.05**	**0.04**	**0.03**	**0.02**	**0.01**	**0.00**
-3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
-3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004	.0005	.0005	.0005
-3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006	.0006	.0007	.0007
-3.1	.0007	.0007	.0008	.0008	.0008	.0008	.0009	.0009	.0009	.0010
-3.0	.0010	.0010	.0011	.0011	.0011	.0012	.0012	.0013	.0013	.0013
-2.9	.0014	.0014	.0015	.0015	.0016	.0016	.0017	.0018	.0018	.0019
-2.8	.0019	.0020	.0021	.0021	.0022	.0023	.0023	.0024	.0025	.0026
-2.7	.0026	.0027	.0028	.0029	.0030	.0031	.0032	.0033	.0034	.0035
-2.6	.0036	.0037	.0038	.0039	.0040	.0041	.0043	.0044	.0045	.0047
-2.5	.0048	.0049	.0051	.0052	.0054	.0055	.0057	.0059	.0060	.0062
-2.4	.0064	.0066	.0068	.0069	.0071	.0073	.0075	.0078	.0080	.0082
-2.3	.0084	.0087	.0089	.0091	.0094	.0096	.0099	.0102	.0104	.0107
-2.2	.0110	.0113	.0116	.0119	.0122	.0125	.0129	.0132	.0136	.0139
-2.1	.0143	.0146	.0150	.0154	.0158	.0162	.0166	.0170	.0174	.0179
-2.0	.0183	.0188	.0192	.0197	.0202	.0207	.0212	.0217	.0222	.0228
-1.9	.0233	.0239	.0244	.0250	.0256	.0262	.0268	.0274	.0281	.0287
-1.8	.0294	.0301	.0307	.0314	.0322	.0329	.0336	.0344	.0351	.0359
-1.7	.0367	.0375	.0384	.0392	.0401	.0409	.0418	.0427	.0436	.0446
-1.6	.0455	.0465	.0475	.0485	.0495	.0505	.0516	.0526	.0537	.0548
-1.5	.0559	.0571	.0582	.0594	.0606	.0618	.0630	.0643	.0655	.0668
-1.4	.0681	.0694	.0708	.0721	.0735	.0749	.0764	.0778	.0793	.0808
-1.3	.0823	.0838	.0853	.0869	.0885	.0901	.0918	.0934	.0951	.0968
-1.2	.0985	.1003	.1020	.1038	.1056	.1075	.1093	.1112	.1131	.1151
-1.1	.1170	.1190	.1210	.1230	.1251	.1271	.1292	.1314	.1335	.1357
-1.0	.1379	.1401	.1423	.1446	.1469	.1492	.1515	.1539	.1562	.1587
-0.9	.1611	.1635	.1660	.1685	.1711	.1736	.1762	.1788	.1814	.1841
-0.8	.1867	.1894	.1922	.1949	.1977	.2005	.2033	.2061	.2090	.2119
-0.7	.2148	.2177	.2206	.2236	.2266	.2296	.2327	.2358	.2389	.2420
-0.6	.2451	.2483	.2514	.2546	.2578	.2611	.2643	.2676	.2709	.2743
-0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981	.3015	.3050	.3085
-0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336	.3372	.3409	.3446
-0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707	.3745	.3783	.3821
-0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090	.4129	.4168	.4207
-0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483	.4522	.4562	.4602
-0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880	.4920	.4960	.5000

Table A.1: Normal cumulative distribution function. For $z > 0$ use $P(z) = 1 - P(-z)$.

For critical z_α or $z_{\alpha/2}$ values, see the last entry ($\nu = \infty$) in the table of critical values for the Student-t (Table A.2).

A.2 Critical Values for the Student's t Distribution

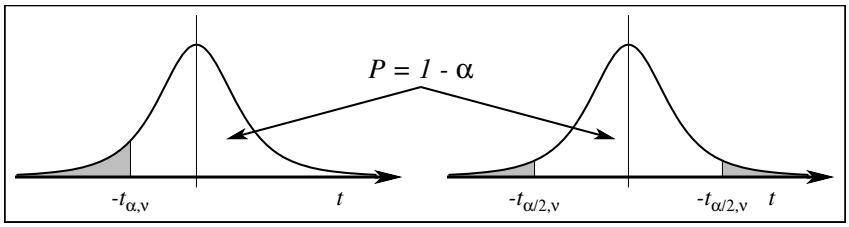


Figure A.2: Given a chosen confidence level, α (area of the tail(s)), and sample size, n, then the degrees of freedom, ν, is $n-1$. Find the corresponding row and column entries and read the critical $t_{\alpha,\nu}$ value (one-sided test, e.g., for $H_0 : t < 0$) or $t_{\alpha/2,\nu}$ value (two-sided test for $H_0 : t = 0$). The sign of t depends on which tail you are considering.

One tail, α:	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
Two tails, α:	**0.20**	**0.1**	**0.05**	**0.02**	**0.01**
ν					
1	3.078	6.314	12.71	31.82	63.66
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
∞	1.282	1.645	1.960	2.326	2.576

Table A.2: Critical values for the Student's t distribution.

A.3 Critical Values for the χ^2 Distribution

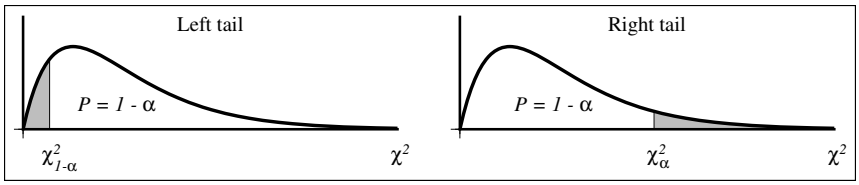


Figure A.3: Given α and degrees of freedom, $\nu = n - 1$, read critical $\chi^2_{1-\alpha,\nu}$ or $\chi^2_{\alpha,\nu}$ values.

Degrees of	**Left tail** $(1-\alpha)$					**Right tail** (α)				
freedom, ν	**0.995**	**0.99**	**0.975**	**0.95**	**0.90**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.60
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09	16.75
6	0.676	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81	18.55
7	0.989	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48	20.28
8	1.344	1.646	2.180	2.733	3.490	13.36	15.51	17.53	20.09	21.95
9	1.735	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67	23.59
10	2.156	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21	25.19
11	2.603	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72	26.76
12	3.074	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22	28.30
13	3.565	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69	29.82
14	4.075	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14	31.32
15	4.601	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58	32.80
16	5.142	5.812	6.908	7.962	9.312	23.54	26.30	28.85	32.00	34.27
17	5.697	6.408	7.564	8.672	10.09	24.77	27.59	30.19	33.41	35.72
18	6.265	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.81	37.16
19	6.844	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.434	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.034	8.897	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	8.643	9.542	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	9.260	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	9.886	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	57.15	60.39	64.28	96.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	118.5	124.3	129.6	135.8	140.2

Table A.3: Critical values for the χ^2 distribution.

A.4 Critical Values for the F Distribution

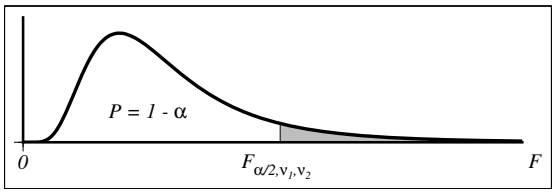


Figure A.4: Given chosen confidence level α (area of the tail) and the degrees of freedom ($\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$), find the corresponding row and column entries and read the critical F_{α,ν_1,ν_2} value (two-sided test, e.g., $H_0 : F = 1$).

A.4.1 F table for 90% confidence level

.

$\downarrow \nu_2 \vert \nu_1 \rightarrow$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**	**20**	**25**	**50**
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	61.22	61.74	62.05	62.69
2	8.526	9.000	9.162	9.243	9.293	9.326	9.349	9.367	9.381	9.392	9.425	9.441	9.451	9.471
3	5.538	5.462	5.391	5.343	5.309	5.285	5.266	5.252	5.240	5.230	5.200	5.184	5.175	5.155
4	4.545	4.325	4.191	4.107	4.051	4.010	3.979	3.955	3.936	3.920	3.870	3.844	3.828	3.795
5	4.060	3.780	3.619	3.520	3.453	3.405	3.368	3.339	3.316	3.297	3.238	3.207	3.187	3.147
6	3.776	3.463	3.289	3.181	3.108	3.055	3.014	2.983	2.958	2.937	2.871	2.836	2.815	2.770
7	3.589	3.257	3.074	2.961	2.883	2.827	2.785	2.752	2.725	2.703	2.632	2.595	2.571	2.523
8	3.458	3.113	2.924	2.806	2.726	2.668	2.624	2.589	2.561	2.538	2.464	2.425	2.400	2.348
9	3.360	3.006	2.813	2.693	2.611	2.551	2.505	2.469	2.440	2.416	2.340	2.298	2.272	2.218
10	3.285	2.924	2.728	2.605	2.522	2.461	2.414	2.377	2.347	2.323	2.244	2.201	2.174	2.117
11	3.225	2.860	2.660	2.536	2.451	2.389	2.342	2.304	2.274	2.248	2.167	2.123	2.095	2.036
12	3.177	2.807	2.606	2.480	2.394	2.331	2.283	2.245	2.214	2.188	2.105	2.060	2.031	1.970
13	3.136	2.763	2.560	2.434	2.347	2.283	2.234	2.195	2.164	2.138	2.053	2.007	1.978	1.915
14	3.102	2.726	2.522	2.395	2.307	2.243	2.193	2.154	2.122	2.095	2.010	1.962	1.933	1.869
15	3.073	2.695	2.490	2.361	2.273	2.208	2.158	2.119	2.086	2.059	1.972	1.924	1.894	1.828
16	3.048	2.668	2.462	2.333	2.244	2.178	2.128	2.088	2.055	2.028	1.940	1.891	1.860	1.793
17	3.026	2.645	2.437	2.308	2.218	2.152	2.102	2.061	2.028	2.001	1.912	1.862	1.831	1.763
18	3.007	2.624	2.416	2.286	2.196	2.130	2.079	2.038	2.005	1.977	1.887	1.837	1.805	1.736
19	2.990	2.606	2.397	2.266	2.176	2.109	2.058	2.017	1.984	1.956	1.865	1.814	1.782	1.711
20	2.975	2.589	2.380	2.249	2.158	2.091	2.040	1.999	1.965	1.937	1.845	1.794	1.761	1.690
21	2.961	2.575	2.365	2.233	2.142	2.075	2.023	1.982	1.948	1.920	1.827	1.776	1.742	1.670
22	2.949	2.561	2.351	2.219	2.128	2.060	2.008	1.967	1.933	1.904	1.811	1.759	1.726	1.652
23	2.937	2.549	2.339	2.207	2.115	2.047	1.995	1.953	1.919	1.890	1.796	1.744	1.710	1.636
24	2.927	2.538	2.327	2.195	2.103	2.035	1.983	1.941	1.906	1.877	1.783	1.730	1.696	1.621
25	2.918	2.528	2.317	2.184	2.092	2.024	1.971	1.929	1.895	1.866	1.771	1.718	1.683	1.607
26	2.909	2.519	2.307	2.174	2.082	2.014	1.961	1.919	1.884	1.855	1.760	1.706	1.671	1.594
27	2.901	2.511	2.299	2.165	2.073	2.005	1.952	1.909	1.874	1.845	1.749	1.695	1.660	1.583
28	2.894	2.503	2.291	2.157	2.064	1.996	1.943	1.900	1.865	1.836	1.740	1.685	1.650	1.572
29	2.887	2.495	2.283	2.149	2.057	1.988	1.935	1.892	1.857	1.827	1.731	1.676	1.640	1.562
30	2.881	2.489	2.276	2.142	2.049	1.980	1.927	1.884	1.849	1.819	1.722	1.667	1.632	1.552
40	2.835	2.440	2.226	2.091	1.997	1.927	1.873	1.829	1.793	1.763	1.662	1.605	1.568	1.483
50	2.809	2.412	2.197	2.061	1.966	1.895	1.840	1.796	1.760	1.729	1.627	1.568	1.529	1.441
60	2.791	2.393	2.177	2.041	1.946	1.875	1.819	1.775	1.738	1.707	1.603	1.543	1.504	1.413
70	2.779	2.380	2.164	2.027	1.931	1.860	1.804	1.760	1.723	1.691	1.587	1.526	1.486	1.392
80	2.769	2.370	2.154	2.016	1.921	1.849	1.793	1.748	1.711	1.680	1.574	1.513	1.472	1.377
90	2.762	2.363	2.146	2.008	1.912	1.841	1.785	1.739	1.702	1.670	1.564	1.503	1.461	1.365
100	2.756	2.356	2.139	2.002	1.906	1.834	1.778	1.732	1.695	1.663	1.557	1.494	1.453	1.355
110	2.752	2.351	2.134	1.997	1.900	1.828	1.772	1.727	1.689	1.657	1.550	1.488	1.446	1.347
120	2.748	2.347	2.130	1.992	1.896	1.824	1.767	1.722	1.684	1.652	1.545	1.482	1.440	1.340
130	2.745	2.344	2.126	1.989	1.892	1.820	1.764	1.718	1.680	1.648	1.541	1.477	1.435	1.334
140	2.742	2.341	2.123	1.985	1.889	1.817	1.760	1.714	1.677	1.645	1.537	1.473	1.431	1.329
150	2.739	2.338	2.121	1.983	1.886	1.814	1.757	1.712	1.674	1.642	1.533	1.470	1.427	1.325

Table A.4: Critical values for the F distribution, $\alpha = 0.1$. Note: ν_1, ν_2 are the degrees of freedom for the numerator and denominator, respectively.

A.4.2 F table for 95% confidence level

.

$\downarrow \nu_2 \vert \nu_1 \rightarrow$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**	**20**	**25**	**50**
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	245.9	248.0	249.3	251.8
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.43	19.45	19.46	19.48
3	10.13	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786	8.703	8.660	8.634	8.581
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.858	5.803	5.769	5.699
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.619	4.558	4.521	4.444
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	3.938	3.874	3.835	3.754
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.511	3.445	3.404	3.319
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347	3.218	3.150	3.108	3.020
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.006	2.936	2.893	2.803
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.845	2.774	2.730	2.637
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854	2.719	2.646	2.601	2.507
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.617	2.544	2.498	2.401
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.533	2.459	2.412	2.314
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602	2.463	2.388	2.341	2.241
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.403	2.328	2.280	2.178
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494	2.352	2.276	2.227	2.124
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450	2.308	2.230	2.181	2.077
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412	2.269	2.191	2.141	2.035
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378	2.234	2.155	2.106	1.999
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.203	2.124	2.074	1.966
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321	2.176	2.096	2.045	1.936
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297	2.151	2.071	2.020	1.909
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275	2.128	2.048	1.996	1.885
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255	2.108	2.027	1.975	1.863
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236	2.089	2.007	1.955	1.842
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220	2.072	1.990	1.938	1.823
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204	2.056	1.974	1.921	1.806
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190	2.041	1.959	1.906	1.790
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177	2.027	1.945	1.891	1.775
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165	2.015	1.932	1.878	1.761
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077	1.924	1.839	1.783	1.660
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026	1.871	1.784	1.727	1.599
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993	1.836	1.748	1.690	1.559
70	3.978	3.128	2.736	2.503	2.346	2.231	2.143	2.074	2.017	1.969	1.812	1.722	1.664	1.530
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951	1.793	1.703	1.644	1.508
90	3.947	3.098	2.706	2.473	2.316	2.201	2.113	2.043	1.986	1.938	1.779	1.688	1.629	1.491
100	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927	1.768	1.676	1.616	1.477
110	3.927	3.079	2.687	2.454	2.297	2.182	2.094	2.024	1.966	1.918	1.758	1.667	1.606	1.466
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910	1.750	1.659	1.598	1.457
130	3.914	3.066	2.674	2.441	2.284	2.169	2.081	2.010	1.953	1.904	1.744	1.652	1.591	1.449
140	3.909	3.061	2.669	2.436	2.279	2.164	2.076	2.005	1.947	1.899	1.738	1.646	1.585	1.442
150	3.904	3.056	2.665	2.432	2.274	2.160	2.071	2.001	1.943	1.894	1.734	1.641	1.580	1.436

Table A.5: Critical values for the F distribution, $\alpha = 0.05$. Note: ν_1, ν_2 are the degrees of freedom for the numerator and denominator, respectively.

A.4.3 F table for 97.5% confidence level

.

$\downarrow \nu_2 \vert \nu_1 \rightarrow$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**	**20**	**25**	**50**
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	984.9	993.1	998.1	1008
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.43	39.45	39.46	39.48
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.25	14.17	14.12	14.01
4	12.22	10.65	9.979	9.605	9.364	9.197	9.074	8.980	8.905	8.844	8.657	8.560	8.501	8.381
5	10.01	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681	6.619	6.428	6.329	6.268	6.144
6	8.813	7.260	6.599	6.227	5.988	5.820	5.695	5.600	5.523	5.461	5.269	5.168	5.107	4.980
7	8.073	6.542	5.890	5.523	5.285	5.119	4.995	4.899	4.823	4.761	4.568	4.467	4.405	4.276
8	7.571	6.059	5.416	5.053	4.817	4.652	4.529	4.433	4.357	4.295	4.101	3.999	3.937	3.807
9	7.209	5.715	5.078	4.718	4.484	4.320	4.197	4.102	4.026	3.964	3.769	3.667	3.604	3.472
10	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717	3.522	3.419	3.355	3.221
11	6.724	5.256	4.630	4.275	4.044	3.881	3.759	3.664	3.588	3.526	3.330	3.226	3.162	3.027
12	6.554	5.096	4.474	4.121	3.891	3.728	3.607	3.512	3.436	3.374	3.177	3.073	3.008	2.871
13	6.414	4.965	4.347	3.996	3.767	3.604	3.483	3.388	3.312	3.250	3.053	2.948	2.882	2.744
14	6.298	4.857	4.242	3.892	3.663	3.501	3.380	3.285	3.209	3.147	2.949	2.844	2.778	2.638
15	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060	2.862	2.756	2.689	2.549
16	6.115	4.687	4.077	3.729	3.502	3.341	3.219	3.125	3.049	2.986	2.788	2.681	2.614	2.472
17	6.042	4.619	4.011	3.665	3.438	3.277	3.156	3.061	2.985	2.922	2.723	2.616	2.548	2.405
18	5.978	4.560	3.954	3.608	3.382	3.221	3.100	3.005	2.929	2.866	2.667	2.559	2.491	2.347
19	5.922	4.508	3.903	3.559	3.333	3.172	3.051	2.956	2.880	2.817	2.617	2.509	2.441	2.295
20	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837	2.774	2.573	2.464	2.396	2.249
21	5.827	4.420	3.819	3.475	3.250	3.090	2.969	2.874	2.798	2.735	2.534	2.425	2.356	2.208
22	5.786	4.383	3.783	3.440	3.215	3.055	2.934	2.839	2.763	2.700	2.498	2.389	2.320	2.171
23	5.750	4.349	3.750	3.408	3.183	3.023	2.902	2.808	2.731	2.668	2.466	2.357	2.287	2.137
24	5.717	4.319	3.721	3.379	3.155	2.995	2.874	2.779	2.703	2.640	2.437	2.327	2.257	2.107
25	5.686	4.291	3.694	3.353	3.129	2.969	2.848	2.753	2.677	2.613	2.411	2.300	2.230	2.079
26	5.659	4.265	3.670	3.329	3.105	2.945	2.824	2.729	2.653	2.590	2.387	2.276	2.205	2.053
27	5.633	4.242	3.647	3.307	3.083	2.923	2.802	2.707	2.631	2.568	2.364	2.253	2.183	2.029
28	5.610	4.221	3.626	3.286	3.063	2.903	2.782	2.687	2.611	2.547	2.344	2.232	2.161	2.007
29	5.588	4.201	3.607	3.267	3.044	2.884	2.763	2.669	2.592	2.529	2.325	2.213	2.142	1.987
30	5.568	4.182	3.589	3.250	3.026	2.867	2.746	2.651	2.575	2.511	2.307	2.195	2.124	1.968
40	5.424	4.051	3.463	3.126	2.904	2.744	2.624	2.529	2.452	2.388	2.182	2.068	1.994	1.832
50	5.340	3.975	3.390	3.054	2.833	2.674	2.553	2.458	2.381	2.317	2.109	1.993	1.919	1.752
60	5.286	3.925	3.343	3.008	2.786	2.627	2.507	2.412	2.334	2.270	2.061	1.944	1.869	1.699
70	5.247	3.890	3.309	2.975	2.754	2.595	2.474	2.379	2.302	2.237	2.028	1.910	1.833	1.660
80	5.218	3.864	3.284	2.950	2.730	2.571	2.450	2.355	2.277	2.213	2.003	1.884	1.807	1.632
90	5.196	3.844	3.265	2.932	2.711	2.552	2.432	2.336	2.259	2.194	1.983	1.864	1.787	1.610
100	5.179	3.828	3.250	2.917	2.696	2.537	2.417	2.321	2.244	2.179	1.968	1.849	1.770	1.592
110	5.164	3.815	3.237	2.904	2.684	2.525	2.405	2.309	2.232	2.167	1.955	1.836	1.757	1.577
120	5.152	3.805	3.227	2.894	2.674	2.515	2.395	2.299	2.222	2.157	1.945	1.825	1.746	1.565
130	5.142	3.796	3.218	2.886	2.666	2.507	2.386	2.291	2.213	2.148	1.936	1.816	1.737	1.555
140	5.134	3.788	3.211	2.879	2.658	2.500	2.379	2.284	2.206	2.141	1.929	1.808	1.729	1.546
150	5.126	3.781	3.204	2.872	2.652	2.494	2.373	2.278	2.200	2.135	1.922	1.801	1.722	1.538

Table A.6: Critical values for the F distribution, $\alpha = 0.025$. Note: ν_1, ν_2 are the degrees of freedom for the numerator and denominator, respectively.

A.4.4 F table for 99% confidence level

.

$\downarrow \nu_2 \vert \nu_1 \rightarrow$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**	**20**	**25**	**50**
1	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6157	6209	6240	6303
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.43	99.45	99.46	99.48
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	26.87	26.69	26.58	26.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.20	14.02	13.91	13.69
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.722	9.553	9.449	9.238
6	13.75	10.92	9.780	9.148	8.746	8.466	8.260	8.102	7.976	7.874	7.559	7.396	7.296	7.091
7	12.25	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719	6.620	6.314	6.155	6.058	5.858
8	11.26	8.649	7.591	7.006	6.632	6.371	6.178	6.029	5.911	5.814	5.515	5.359	5.263	5.065
9	10.56	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351	5.257	4.962	4.808	4.713	4.517
10	10.04	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942	4.849	4.558	4.405	4.311	4.115
11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632	4.539	4.251	4.099	4.005	3.810
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388	4.296	4.010	3.858	3.765	3.569
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191	4.100	3.815	3.665	3.571	3.375
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030	3.939	3.656	3.505	3.412	3.215
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895	3.805	3.522	3.372	3.278	3.081
16	8.531	6.226	5.292	4.773	4.437	4.202	4.026	3.890	3.780	3.691	3.409	3.259	3.165	2.967
17	8.400	6.112	5.185	4.669	4.336	4.102	3.927	3.791	3.682	3.593	3.312	3.162	3.068	2.869
18	8.285	6.013	5.092	4.579	4.248	4.015	3.841	3.705	3.597	3.508	3.227	3.077	2.983	2.784
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.631	3.523	3.434	3.153	3.003	2.909	2.709
20	8.096	5.849	4.938	4.431	4.103	3.871	3.699	3.564	3.457	3.368	3.088	2.938	2.843	2.643
21	8.017	5.780	4.874	4.369	4.042	3.812	3.640	3.506	3.398	3.310	3.030	2.880	2.785	2.584
22	7.945	5.719	4.817	4.313	3.988	3.758	3.587	3.453	3.346	3.258	2.978	2.827	2.733	2.531
23	7.881	5.664	4.765	4.264	3.939	3.710	3.539	3.406	3.299	3.211	2.931	2.781	2.686	2.483
24	7.823	5.614	4.718	4.218	3.895	3.667	3.496	3.363	3.256	3.168	2.889	2.738	2.643	2.440
25	7.770	5.568	4.675	4.177	3.855	3.627	3.457	3.324	3.217	3.129	2.850	2.699	2.604	2.400
26	7.721	5.526	4.637	4.140	3.818	3.591	3.421	3.288	3.182	3.094	2.815	2.664	2.569	2.364
27	7.677	5.488	4.601	4.106	3.785	3.558	3.388	3.256	3.149	3.062	2.783	2.632	2.536	2.330
28	7.636	5.453	4.568	4.074	3.754	3.528	3.358	3.226	3.120	3.032	2.753	2.602	2.506	2.300
29	7.598	5.420	4.538	4.045	3.725	3.499	3.330	3.198	3.092	3.005	2.726	2.574	2.478	2.271
30	7.562	5.390	4.510	4.018	3.699	3.473	3.304	3.173	3.067	2.979	2.700	2.549	2.453	2.245
40	7.314	5.179	4.313	3.828	3.514	3.291	3.124	2.993	2.888	2.801	2.522	2.369	2.271	2.058
50	7.171	5.057	4.199	3.720	3.408	3.186	3.020	2.890	2.785	2.698	2.419	2.265	2.167	1.949
60	7.077	4.977	4.126	3.649	3.339	3.119	2.953	2.823	2.718	2.632	2.352	2.198	2.098	1.877
70	7.011	4.922	4.074	3.600	3.291	3.071	2.906	2.777	2.672	2.585	2.306	2.150	2.050	1.826
80	6.963	4.881	4.036	3.563	3.255	3.036	2.871	2.742	2.637	2.551	2.271	2.115	2.015	1.788
90	6.925	4.849	4.007	3.535	3.228	3.009	2.845	2.715	2.611	2.524	2.244	2.088	1.987	1.759
100	6.895	4.824	3.984	3.513	3.206	2.988	2.823	2.694	2.590	2.503	2.223	2.067	1.965	1.735
110	6.871	4.803	3.965	3.495	3.188	2.970	2.806	2.677	2.573	2.486	2.206	2.049	1.947	1.716
120	6.851	4.787	3.949	3.480	3.174	2.956	2.792	2.663	2.559	2.472	2.192	2.035	1.932	1.700
130	6.834	4.772	3.936	3.467	3.161	2.944	2.780	2.651	2.547	2.460	2.179	2.022	1.920	1.686
140	6.819	4.760	3.925	3.456	3.151	2.933	2.769	2.641	2.536	2.450	2.169	2.012	1.909	1.675
150	6.807	4.749	3.915	3.447	3.142	2.924	2.761	2.632	2.528	2.441	2.160	2.003	1.900	1.665

Table A.7: Critical values for the F distribution, $\alpha = 0.01$. Note: ν_1, ν_2 are the degrees of freedom for the numerator and denominator, respectively.

A.5 Critical Values for the Mann-Whitney (U-Test)

These critical values are used when comparing two samples based on their *ranks* instead of their values. These tables are symmetrical so it does not matter which sample you label "1" and which one is sample "2", i.e., $U_{\alpha,n_1,n_2} \equiv U_{\alpha,n_2,n_1}$. For sample sizes larger than 20 you may use the normal distribution approximation (4.61).

A.5.1 One-sided tests

This section concerns itself with *one-sided* comparisons, i.e., $H_0 : \mu_1 > \mu_2$. Depending on your level of confidence, use either Table A.8 (for $\alpha = 0.05$) or Table A.9 (for $\alpha = 0.01$).

$\downarrow n_1 \vert n_2 \rightarrow$	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
2	0	0	0	0	0	0	1	1	1	1	2	2	3	3	3	3	4	4	4
3	0	0	0	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11
4	0	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
5	0	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
6	0	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
7	0	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
8	1	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
9	1	4	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
10	1	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
11	1	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
12	2	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
13	2	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
14	3	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

Table A.8: Critical values for the one-sided U-test for $\alpha = 0.05$. Note the use of n rather than ν.

$\downarrow n_1 \vert n_2 \rightarrow$	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
3	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	0	0	0	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	0	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	0	0	1	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7	0	0	1	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
8	0	0	2	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
9	0	1	3	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
10	0	1	3	6	8	11	13	16	19	22	24	27	30	33	36	38	41	44	47
11	0	1	4	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
12	0	2	5	8	11	14	17	21	24	28	31	35	38	42	46	49	53	56	60
13	0	2	5	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
14	0	2	6	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56	61	66	70	75	80
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Table A.9: Critical values for the one-sided U-test for $\alpha = 0.01$. Note the use of n rather than ν.

A.5.2 Two-sided tests

This section concerns itself with *two-sided* comparisons, i.e., $H_0 : \mu_1 = \mu_2$. Depending on your level of confidence, use either Table A.10 (for $\alpha = 0.05$) or Table A.11 (for $\alpha = 0.01$).

$\downarrow n_1 \vert n_2 \rightarrow$	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
2	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	2	2	2	2
3	0	0	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	0	0	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
5	0	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	0	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	0	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	0	3	5	8	11	14	17	20	23	26	29	30	36	39	42	45	48	52	55
11	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	1	4	8	12	16	20	24	28	30	37	41	45	50	54	59	63	67	72	76
14	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	2	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

Table A.10: Critical values for the two-sided U-test for $\alpha = 0.05$. Note the use of n rather than ν.

$\downarrow n_1 \vert n_2 \rightarrow$	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	0	0	0	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	0	0	0	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	0	0	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	0	0	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	0	0	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	0	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	0	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	0	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	0	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	0	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	57	60
14	0	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	0	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	0	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	0	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	0	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	0	3	7	12	17	22	28	33	39	45	51	57	63	69	74	81	87	93	99
20	0	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

Table A.11: Critical values for the two-sided U-test for $\alpha = 0.01$. Note the use of n rather than ν.

A.6 Critical Values for the Kolmogorov-Smirnov Distribution

A.6.1 One-sample comparison with a known distribution

Use table A.12 when comparing a single sample's cumulative distribution to that of a *known* distribution, i.e., its parameters (μ, σ) are prescribed by the null hypothesis.

One tail, α:	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
Two tails, α:	**0.20**	**0.1**	**0.05**	**0.02**	**0.01**
n=**1**	0.900	0.950	0.975	0.990	0.995
2	0.684	0.776	0.842	0.900	0.929
3	0.565	0.636	0.708	0.785	0.829
4	0.493	0.565	0.624	0.689	0.734
5	0.447	0.509	0.563	0.627	0.669
6	0.410	0.468	0.519	0.577	0.617
7	0.381	0.436	0.483	0.538	0.576
8	0.358	0.410	0.454	0.507	0.542
9	0.339	0.387	0.430	0.480	0.513
10	0.323	0.369	0.409	0.457	0.489
11	0.308	0.352	0.391	0.437	0.468
12	0.296	0.338	0.375	0.419	0.449
13	0.285	0.325	0.361	0.404	0.432
14	0.275	0.314	0.349	0.390	0.418
15	0.266	0.304	0.338	0.377	0.404
16	0.258	0.295	0.327	0.366	0.392
17	0.250	0.286	0.318	0.355	0.381
18	0.244	0.279	0.309	0.346	0.371
19	0.237	0.271	0.301	0.337	0.361
20	0.232	0.265	0.294	0.329	0.352
21	0.226	0.259	0.287	0.321	0.344
22	0.221	0.253	0.281	0.314	0.337
23	0.216	0.247	0.275	0.307	0.330
24	0.212	0.242	0.269	0.301	0.323
25	0.208	0.238	0.264	0.295	0.317
26	0.204	0.233	0.259	0.290	0.311
27	0.200	0.229	0.254	0.284	0.305
28	0.197	0.225	0.250	0.279	0.300
29	0.193	0.221	0.246	0.275	0.295
30	0.190	0.218	0.242	0.270	0.290
31	0.187	0.214	0.238	0.266	0.285
32	0.184	0.211	0.234	0.262	0.281
33	0.182	0.208	0.231	0.258	0.277
34	0.179	0.205	0.227	0.254	0.273
35	0.177	0.202	0.224	0.251	0.269
36	0.174	0.199	0.221	0.247	0.265
37	0.172	0.196	0.218	0.244	0.262
38	0.170	0.194	0.215	0.241	0.258
39	0.168	0.191	0.213	0.238	0.255
40	0.165	0.189	0.210	0.235	0.252
$n > 40$	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Table A.12: Critical values for the one-sample Kolmogorov-Smirnov test. Note the use of n rather than ν.

A.6.2 One-sample comparison with an unknown distribution (Lilliefors)

Use table A.13 when comparing a single sample's cumulative distribution to that of an *unknown* normal distribution. Equate the unknown distribution's parameters (μ, σ) to those from the sample $(\bar{x}, s)$ in order to compute its cumulative distribution. The parameter estimation reduces the degrees of freedom by two and the standard Kolmogorov-Smirnov critical values are not appropriate. Lilliefors developed a numerical solution that gives the correct critical values. This is a two-sided test.

α:	**0.10**	**0.05**	**0.01**	**0.001**
n=**4**	0.344	0.375	0.414	0.432
5	0.320	0.344	0.398	0.427
6	0.298	0.323	0.369	0.421
7	0.281	0.305	0.351	0.399
8	0.266	0.289	0.334	0.383
9	0.252	0.273	0.316	0.366
10	0.240	0.261	0.305	0.350
11	0.231	0.251	0.291	0.331
12	0.223	0.242	0.281	0.327
14	0.208	0.226	0.262	0.302
16	0.195	0.213	0.249	0.291
18	0.185	0.201	0.234	0.272
20	0.176	0.192	0.223	0.266
25	0.159	0.173	0.202	0.236
30	0.146	0.159	0.186	0.219
40	0.127	0.139	0.161	0.190
50	0.114	0.125	0.145	0.173
60	0.105	0.114	0.133	0.159
75	0.094	0.102	0.119	0.138
100	0.082	0.089	0.104	0.121
$n > 100$	$\frac{0.816}{\sqrt{n}}$	$\frac{0.888}{\sqrt{n}}$	$\frac{1.038}{\sqrt{n}}$	$\frac{1.212}{\sqrt{n}}$

Table A.13: Critical values for the Lilliefors test. Note the use of n rather than ν.

A.6.3 Two-sample comparison

Table A.14 gives critical D-values for $\alpha = 0.05$ (upper row) and $\alpha = 0.01$ (lower row) for various sample sizes, n_1 and n_2. The asterisk (*) means you cannot reject H_0 regardless of observed D. For blank entries simply reverse n_1 and n_2.

$n_2 \vert n_1$:	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	*	*	*	*	*	*	*	*	*	*
	*	*	*	*	*	*	*	*	*	*
2	*	*	*	*	*	16/16	18/18	20/20	22/22	24/24
	*	*	*	*	*	*	*	*	*	*
3	*	*	15/15	18/18	21/21	21/24	24/27	27/30	30/33	30/36
	*	*	*	*	*	24/24	27/27	30/30	33/33	36/36
4		16/16	20/20	20/24	24/28	28/32	28/36	30/40	33/44	36/48
		*	*	24/24	28/28	32/32	32/36	36/40	40/44	44/48
5			*	24/30	30/35	30/40	35/45	40/50	39/55	43/60
			*	30/30	35/35	35/40	40/45	45/50	45/55	50/60
6				30/36	30/42	34/48	39/54	40/60	43/66	48/72
				36/36	36/42	40/48	45/54	48/60	54/66	60/72
7					42/49	40/56	42/63	46/70	48/77	53/84
					42/49	48/56	49/63	53/70	59/77	60/84
8						48/64	46/72	48/80	53/88	60/96
						56/64	55/72	60/80	64/88	68/96
9							54/81	53/90	59/99	63/108
							63/81	70/90	70/99	75/108
10								70/100	60/110	66/120
								80/100	77/110	80/120
11									77/121	72/132
									88/121	86/132
12										96/144
										84/144

Table A.14: Critical values for the two-sample Kolmogorov-Smirnov test. Note the use of n rather than ν.

For larger sample sizes, the approximate critical value D_α is given by the equation

$$D_\alpha = c(\alpha)\sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}, \tag{A.1}$$

where the coefficient $c(\alpha)$ is given via the table below:

α	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**
$c(\alpha)$	1.22	1.36	1.48	1.63	1.73	1.95

Examples:

1. For $\alpha = 0.05$ and sample sizes 5 and 8, $D_\alpha = 30/40 = 0.75$.
2. For $\alpha = 0.01$ and sample sizes 15 and 28, $D_\alpha = 1.63\sqrt{\frac{15+28}{15 \cdot 28}} = 0.522$.

A.7 Critical Values for Spearman's Rank Correlation

This is either a one-sided (e.g., $H_0 : \rho > 0$) or two-sided ($H_0 : \rho = 0$) test for the nonparametric rank correlation. Given the number of pairs, n, in the sample and the chosen confidence level, α, Table A.15 shows the critical correlation that must be exceeded for H_0 to be rejected. The asterisk identify situations when H_0 cannot be rejected.

One tail, α:	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
Two tails, α:	**0.20**	**0.1**	**0.05**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
$n \leq$ **3**	*	*	*	*	*	*	*	*
4	1.000	1.000	*	*	*	*	*	*
5	0.800	0.900	1.000	1.000	*	*	*	*
6	0.657	0.829	0.886	0.943	1.000	1.000	*	*
7	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.455	0.564	0.648	0.745	0.794	0.830	0.879	0.903
11	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.406	0.503	0.587	0.678	0.727	0.769	0.818	0.846
13	0.385	0.484	0.560	0.648	0.703	0.747	0.791	0.824
14	0.367	0.464	0.538	0.626	0.679	0.723	0.771	0.802
15	0.354	0.446	0.521	0.604	0.654	0.700	0.750	0.779
16	0.141	0.429	0.503	0.582	0.635	0.679	0.729	0.762
17	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.299	0.380	0.447	0.520	0.570	0.612	0.662	0.696
21	0.292	0.170	0.435	0.508	0.556	0.599	0.648	0.681
22	0.284	0.161	0.425	0.496	0.544	0.586	0.634	0.667
23	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.265	0.137	0.398	0.466	0.511	0.551	0.598	0.630
26	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.255	0.324	0.182	0.448	0.491	0.531	0.577	0.608
28	0.250	0.317	0.175	0.440	0.483	0.522	0.567	0.598
29	0.245	0.312	0.168	0.433	0.475	0.513	0.558	0.589
30	0.240	0.306	0.162	0.425	0.467	0.504	0.549	0.580
31	0.236	0.101	0.356	0.418	0.459	0.496	0.541	0.571
32	0.232	0.296	0.150	0.412	0.452	0.489	0.533	0.563
33	0.229	0.291	0.345	0.405	0.446	0.412	0.525	0.554
34	0.225	0.287	0.340	0.399	0.439	0.475	0.517	0.547
35	0.222	0.283	0.335	0.394	0.433	0.468	0.510	0.539
36	0.219	0.279	0.330	0.388	0.427	0.462	0.504	0.533
37	0.216	0.215	0.125	0.183	0.421	0.456	0.497	0.526
38	0.212	0.271	0.321	0.378	0.415	0.450	0.491	0.519
39	0.210	0.267	0.317	0.173	0.410	0.444	0.485	0.513
40	0.207	0.264	0.113	0.168	0.405	0.439	0.479	0.507
41	0.204	0.261	0.309	0.364	0.400	0.433	0.473	0.501
42	0.202	0.257	0.305	0.159	0.395	0.428	0.468	0.495
43	0.199	0.254	0.101	0.355	0.391	0.423	0.463	0.490
44	0.197	0.251	0.298	0.351	0.386	0.419	0.458	0.484
45	0.194	0.248	0.294	0.147	0.382	0.414	0.453	0.479
46	0.192	0.246	0.291	0.343	0.378	0.410	0.448	0.474
47	0.190	0.243	0.288	0.140	0.174	0.405	0.443	0.469
48	0.188	0.240	0.285	0.336	0.370	0.401	0.439	0.465
49	0.186	0.238	0.282	0.333	0.366	0.197	0.104	0.460
50	0.184	0.235	0.279	0.329	0.163	0.393	0.430	0.456

Table A.15: Critical values for Spearman's rank correlation. Note the use of n rather than ν.

A.8 Relationship Between κ and $\bar{R}$ for 2-D Directional Data

Given a mean resultant ($\bar{R}$) we use Table A.16 to obtain the corresponding concentration parameter (κ) for directional data in the plane. Alternatively, one can solve the implicit equation for κ given by

$$\bar{R} = \frac{I_1(\kappa)}{I_0(\kappa)}. \tag{A.2}$$

$\bar{R}$	κ	$\bar{R}$	κ	$\bar{R}$	κ
0.00	0.00000	0.34	0.72356	0.68	1.89637
0.01	0.02000	0.35	0.74783	0.69	1.95357
0.02	0.04001	0.36	0.77241	0.70	2.01363
0.03	0.06003	0.37	0.79730	0.71	2.07685
0.04	0.08006	0.38	0.82253	0.72	2.14359
0.05	0.10013	0.39	0.84812	0.73	2.21425
0.06	0.12022	0.40	0.87408	0.74	2.28930
0.07	0.14034	0.41	0.90043	0.75	2.36930
0.08	0.16051	0.42	0.92720	0.76	2.45490
0.09	0.18073	0.43	0.95440	0.77	2.54686
0.10	0.20101	0.44	0.98207	0.78	2.64613
0.11	0.22134	0.45	1.01022	0.79	2.75382
0.12	0.24175	0.46	1.03889	0.80	2.87129
0.13	0.26223	0.47	1.06810	0.81	3.00020
0.14	0.28279	0.48	1.09788	0.82	3.14262
0.15	0.30344	0.49	1.12828	0.83	3.30114
0.16	0.32419	0.50	1.15932	0.84	3.47901
0.17	0.34503	0.51	1.19105	0.85	3.68041
0.18	0.36599	0.52	1.22350	0.86	3.91072
0.19	0.38707	0.53	1.25672	0.87	4.17703
0.20	0.40828	0.54	1.29077	0.88	4.48876
0.21	0.42962	0.55	1.32570	0.89	4.85871
0.22	0.45110	0.56	1.36156	0.90	5.30469
0.23	0.47273	0.57	1.39842	0.91	5.85223
0.24	0.49453	0.58	1.43635	0.92	6.53939
0.25	0.51649	0.59	1.47543	0.93	7.42572
0.26	0.53863	0.60	1.51574	0.94	8.61035
0.27	0.56097	0.61	1.55738	0.95	10.2717
0.28	0.58350	0.62	1.60044	0.96	12.7668
0.29	0.60625	0.63	1.64506	0.97	16.9289
0.30	0.62922	0.64	1.69134	0.98	25.2581
0.31	0.65242	0.65	1.73945	0.99	50.2506
0.32	0.67587	0.66	1.78953	1.00	∞
0.33	0.69958	0.67	1.84177		

Table A.16: Relationship between κ and $\bar{R}$ in 2-D.

A.9 Critical Values of $\bar{R}$ for 2-D Directional Data

Given the level of significance we determine the critical value for the mean resultant length under the null hypothesis of no preferred direction in the plane ($H_0 : \bar{R} = 0$), with the alternative hypothesis being that the data can be described via the von Mises distribution (9.7) with a preferred trend ($H_1 : \bar{R} \neq 0$).

α:	**0.10**	**0.05**	**0.025**	**0.01**
$n =$ **4**	0.768	0.847	0.905	0.960
5	0.677	0.754	0.816	0.879
6	0.618	0.690	0.753	0.825
7	0.572	0.642	0.702	0.771
8	0.535	0.602	0.660	0.725
9	0.504	0.569	0.624	0.687
10	0.478	0.540	0.594	0.655
11	0.456	0.516	0.567	0.627
12	0.437	0.494	0.544	0.602
13	0.420	0.475	0.524	0.580
14	0.405	0.458	0.505	0.560
15	0.391	0.443	0.489	0.542
16	0.379	0.429	0.474	0.525
17	0.367	0.417	0.460	0.510
18	0.357	0.405	0.447	0.496
19	0.348	0.394	0.436	0.484
20	0.339	0.385	0.425	0.472
21	0.331	0.375	0.415	0.461
22	0.323	0.367	0.405	0.451
23	0.316	0.359	0.397	0.441
24	0.309	0.351	0.389	0.432
25	0.303	0.344	0.381	0.423
30	0.277	0.315	0.348	0.387
35	0.256	0.292	0.323	0.359
40	0.240	0.273	0.302	0.336
45	0.226	0.257	0.285	0.318
50	0.214	0.244	0.270	0.301

Table A.17: Critical values for $\bar{R}$ in the plane. Note the use of n rather than ν.

A.10 Relationship Between κ and $\bar{R}$ for 3-D Directional Data

Given a mean resultant ($\bar{R}$) we use Table A.18 to obtain the corresponding concentration parameter (κ) for directional data in space. Alternatively, one can solve the implicit equation for κ given by

$$\coth \kappa - 1/\kappa = \bar{R}. \tag{A.3}$$

$\bar{R}$	κ	$\bar{R}$	κ	$\bar{R}$	κ
0.00	0.00000	0.34	1.09951	0.68	3.08456
0.01	0.03000	0.35	1.13739	0.69	3.19091
0.02	0.06001	0.36	1.17584	0.70	3.30354
0.03	0.09005	0.37	1.21490	0.71	3.42314
0.04	0.12012	0.38	1.25459	0.72	3.55051
0.05	0.15023	0.39	1.29497	0.73	3.68655
0.06	0.18039	0.40	1.33605	0.74	3.83232
0.07	0.21062	0.41	1.37789	0.75	3.98905
0.08	0.24093	0.42	1.42053	0.76	4.15819
0.09	0.27132	0.43	1.46401	0.77	4.34143
0.10	0.30182	0.44	1.50839	0.78	4.54076
0.11	0.33242	0.45	1.55372	0.79	4.75857
0.12	0.36315	0.46	1.60005	0.80	4.99772
0.13	0.39402	0.47	1.64745	0.81	5.26167
0.14	0.42503	0.48	1.69599	0.82	5.55463
0.15	0.45621	0.49	1.74573	0.83	5.88181
0.16	0.48756	0.50	1.79676	0.84	6.24971
0.17	0.51909	0.51	1.84915	0.85	6.66652
0.18	0.55083	0.52	1.90300	0.86	7.14279
0.19	0.58278	0.53	1.95842	0.87	7.69228
0.20	0.61497	0.54	2.01550	0.88	8.33333
0.21	0.64740	0.55	2.07437	0.89	9.09091
0.22	0.68009	0.56	2.13515	0.90	10.0000
0.23	0.71306	0.57	2.19799	0.91	11.1111
0.24	0.74632	0.58	2.26304	0.92	12.5000
0.25	0.77990	0.59	2.33049	0.93	14.2857
0.26	0.81381	0.60	2.40050	0.94	16.6667
0.27	0.84806	0.61	2.47331	0.95	20.0000
0.28	0.88269	0.62	2.54914	0.96	25.0000
0.29	0.91771	0.63	2.62825	0.97	33.3333
0.30	0.95315	0.64	2.71093	0.98	50.0000
0.31	0.98902	0.65	2.79751	0.99	100.000
0.32	1.02536	0.66	2.88836	1.00	∞
0.33	1.06218	0.67	2.98389		

Table A.18: Relationship between κ and $\bar{R}$ in 3-D.

A.11 Critical Values of $\bar{R}$ for 3-D Directional Data

Given the level of significance we determine the critical value for the mean resultant length under the null hypothesis of no preferred direction in space ($H_0 : \bar{R} = 0$), with the alternative hypothesis being that the data can be described via the Fisher distribution (9.31) with a preferred trend ($H_1 : \bar{R} \neq 0$).

α:	**0.10**	**0.05**	**0.025**	**0.01**
n = **5**	0.637	0.700	0.765	0.805
6	0.583	0.642	0.707	0.747
7	0.541	0.597	0.659	0.698
8	0.506	0.560	0.619	0.658
9	0.478	0.529	0.586	0.624
10	0.454	0.503	0.558	0.594
11	0.433	0.480	0.533	0.568
12	0.415	0.460	0.512	0.546
13	0.398	0.442	0.492	0.526
14	0.384	0.427	0.475	0.507
15	0.371	0.413	0.460	0.491
16	0.359	0.400	0.446	0.476
17	0.349	0.388	0.443	0.463
18	0.339	0.377	0.421	0.450
19	0.330	0.367	0.410	0.438
20	0.322	0.358	0.399	0.428
21	0.314	0.350	0.390	0.418
22	0.307	0.342	0.382	0.408
23	0.300	0.334	0.374	0.400
24	0.294	0.328	0.366	0.392
25	0.288	0.321	0.359	0.384
30	0.26	0.29	0.33	0.36
35	0.24	0.27	0.31	0.33
40	0.23	0.26	0.29	0.31
45	0.22	0.24	0.27	0.29
50	0.20	0.23	0.26	0.28
100	0.14	0.16	0.18	0.19

Table A.19: Critical values for $\bar{R}$ in 3-D space. Note the use of n rather than ν.

Index