UNIVERSITÉ
CAEN
NORMANDIE

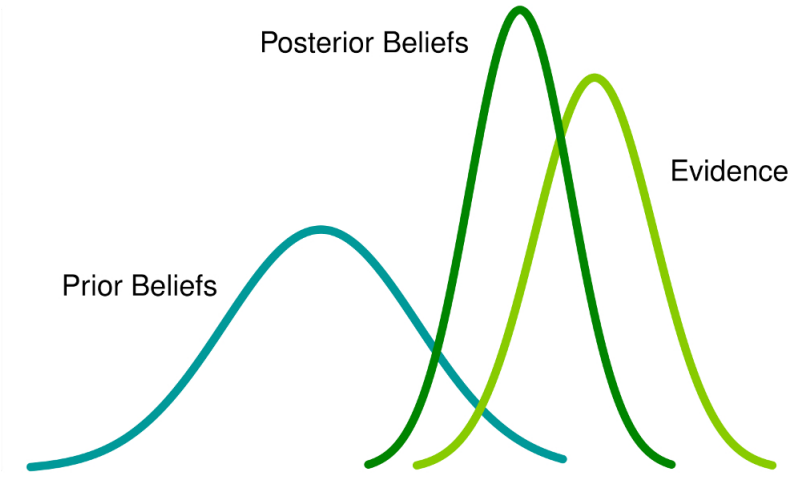
Normandie Université

Introduction to metrology

Uncertainties - Bayesian inference - Data analysis - Experimental designs

David Boilley[1]
Normandie université/Unicaen and GANIL

2018-2019



[1]GANIL, BP 55027, F-14076 Caen cedex 05, boilley@ganil.fr, tél : 02 31 45 47 81, http://tinyurl.com/davidboilley

If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts he shall end in certainties.

Francis Bacon, Book I, v, 8.

Foreword

"When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be."

William Thomson, alias Lord Kevin, in Lecture on "Electrical Units of Measurement" (3 May 1883), published in Popular Lectures Vol. I, p. 73.

Measurement is the cornerstone of any scientific knowledge. Measurement results that you produce will be used by others like you will use others' results. Thus, there should be no ambiguity in the reported results. In order to understand each other, one should share the same language, the same writing and use the same standards.

This started after the French revolution with the unification of the units. Nowadays most of the countries in the world use the same metric system. Standardisation process is still going on: a new definition of the base units is scheduled for November 2018. But beyond the units, one of the main standardisation processes in metrology of these last 30 years is related to the evaluation of uncertainties in measurements which requires a mathematical background. This will be the first part of this lecture.

Estimation of the uncertainty in a measurement can be more difficult than the measurement itself. But it leads to a deep understanding of the whole process that is necessary to improve the quality of the result.

Then measurements data need to be analysed. They are used to fit models in order to answer to specific questions. Data analysis consists in mining information in the data to be able to understand the world and do predictions. These predictions are tainted by uncertainties that have to be evaluated. This will be the second part of the lecture.

Finally, there is a way to optimise experiments in order to produce the maximum information with a minimal number of measurements. This is the goal of the so-called design of experiments that constitutes the last par of this lecture.

All this is a part of *metrology* [métrologie] which is the science of measurement and its application.

This is still a preliminary version of my lecture notes. It surely contains many misprints and mistakes. Any comment, remark, suggestion are welcome.

Contents

Part I

Producing data

Chapter 1

Printing and Using Symbols and Numbers in Scientific and Technical Documents

1.1 International System of Units

This section is mainly based on the presentation of the international system of units done by the *Bureau International des Poids et Mesures*, `http://www.bipm.org`. The system is presented in a brochure that can be downloaded in French or English [4].

1.1.1 Brief history

In the past, each village or province had its own units of measurement, hampering the exchanges and the development of the trade. Central authorities have tried to define a united system but, in Europe, this was only possible after the French revolution with the metric system. The convention of the metre in 1875 adopted the metric system as an international standard. The definition of base units have changed several times over the years.

The French revolution started in 1789 and the same year the first version of the Declaration of the Rights of Man of the Citizen was written with an ambition to be universal. One year later, the French national assembly decided to have a unified system of units. The pendulum was suggested for the second. And in 1791, the French Academy of Sciences adopted the decimal system, decided to define the metre from the length of the meridian and abandoned the pendulum for the second. In 1793, the king was killed and the second version of the Declaration of the Rights of Man of the Citizen is adopted. As you can see, the metric system was one of the highest priority of the new French authorities.

The French assembly declared on the 26th of March 1791:

> *"Considérant que, pour parvenir à établir l'uniformité des poids et mesures, il est nécessaire de fixer une unité de mesure naturelle et invariable et que le seul moyen d'étendre cette uniformité aux nations étrangères et de les engager à convenir d'un système de mesure est de choisir une unité qui ne renferme rien d'arbitraire ni de particulier à la situation d'aucun peuple sur le globe.... adopte la grandeur du quart du méridien terrestre pour base du nouveau système de mesures qui sera décimal ; les opérations nécessaires pour déterminer cette base, notamment la mesure d'un arc de méridien depuis Dunkerque jusqu'à Barcelone seront incessamment exécutées."*
> "Considering that, to achieve the establishment of a universality of weight and measures, it is necessary to fix a natural and unchanging unit and, to extend this uniformity to foreign nations... "

As for the Human rights, it was obvious from the beginning that the new standards should extended to the other nations.

The new metric system was adopted on the 7th of April 1795 (18 germinal III). The new unit is the metre for the length. At that time, the population was not educated and are was also defined for the surface area

and litre for the volume. Gram and bar were also adopted on that day. Multiple with Greek names were also defined: deca, hecto... and fractions with latin names: deci, centi, milli...

For practical reasons the standard metre was made of platinum ruler that can still be seen in front of the French Senat in Paris. The case of the gram is also interesting: the unit is the gram, but the actual standard is a multiple, the kilogram. Lavoisier, under the king Louis XVI first proposed to define the unit of weight (the grave) as the mass of one litre of water at freezing temperature. But after the French revolution, it was decided to adopt a smaller unit for practical reasons in trade.

The new system was adopted by law in France in 1799. The first international convention meeting was held in Paris in 1870. 24 countries joined, but it was stopped by the war. 17 countries finally established the so-called Conférence Générale des Poids et Mesures and Bureau International des Poids et Mesure in Paris in 1875. It still exists.

The BIPM (Bureau International des Poids et Mesures) presently have 56 member states and the process is not finish: even nowadays we all use inches to measure the size of a screen or barils for oil! However the SI is the only system of units that is universally recognised, so it has a distinct advantage for establishing international dialogue.

1.1.2 The new definition of the base units

The SI base units are a choice of seven well-defined units which by convention are regarded as dimensionally independent. See the SI brochure for details [4]. Note that a complete revision with new definitions is to be adopted in November 2018. It is based on fundamental constants as depicted in the new SI logo 1.1.

Figure 1.1: New SI logo.

Time

The second, symbol s, is the SI unit of time.

The first suggestion at the time of the French revolution was to use a pendulum. But it quickly turned out that this was not appropriate because it depends on g that is not uniform. The first international definition was based on the fraction 1/86 400 of the mean solar day, but the "mean solar day" was not well defined.

Since 1967, the second has been defined as the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom. This still the case with the new system of units.

Length

The metre, symbol m, is the SI unit of length.

First defined as 1/40 000 000 of the Earth's circumference, measurements showed that our planet is not a perfect sphere. This definition is not universal. In 1887, Michelson suggested to use interferencies to measure a wavelength and define the metre. In 1960, the wave length of a given transition of Krypton 86 was adopted to define it.

Since 1983, the metre has been defined as the length of the path travelled by light in vacuum during a time interval of 1/299 792 458 of a second. This still the case with the new system of units.

Mass

The kilogram, symbol kg, is the SI unit of mass.

Kilogram, called "grave" at the time of the French revolution, was first defined as the mass of one litre of pure water at the freezing temperature. But it turned out that the unit was too large for most of the daily trading on markets. A smaller unit, the gram, was chosen at the time of the French revolution, but the standard is still the kilogram.

Until November 2018, It was defined from the 1889 prototype, an artefact made of platinum-iridium, kept in Paris. The definition at that time was simply: this prototype shall henceforth be considered to be the unit of mass. But to avoid common confusion the public between "mass" and "weight", it was confirmed in 1901, that the kilogram is equal to the mass of the international prototype of the kilogram. However, due to the inevitable accumulation of contaminants on surfaces, the international prototype is subject to reversible surface contamination that approaches 1 μg per year in mass. For this reason, the CIPM declared that, pending further research, the reference mass of the international prototype is that immediately after cleaning and washing by a specified method.

In the new system, it is defined by taking the fixed numerical value of the Planck constant h to be $6.626\,070\,15 \times 10^{-34}$ when expressed in the unit J s, which is equal to kg m^2 s^{-1}, where the metre and the second are defined above. The value of the Planck constant is chosen to ensure that there will be no change in the SI kilogram at the time of redefinition.

Electric current

The ampere, symbol A, is the SI unit of electric current.

Until November 2018, the ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 metre apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per metre of length.

In the new system, it is defined by taking the fixed numerical value of the elementary charge e to be $1.602\,176\,634 \times 10^{-19}$ when expressed in the unit C, which is equal to A s.

The ampere and other electrical units, as practically realised at the highest metrological level, will become fully consistent with the definitions of these units. The transition from the 1990 convention to the revised SI will result in small changes to all disseminated electrical units. For the vast majority of measurement users, no action need be taken as the volt will change by about 0.1 parts per million and the ohm will change by even less. Practitioners working at the highest level of accuracy may need to adjust the values of their standards and review their measurement uncertainty budgets.

Thermodynamic temperature

The kelvin, symbol K, is the SI unit of thermodynamic temperature.

The former definition of the unit of thermodynamic temperature was given in 1954 and was based on the triple point of water, T_{TPW}, as a fundamental fixed point and assigned to it the temperature 273.16 K, thereby defining the kelvin. The 13th CGPM (1967-1968) adopted the name kelvin, symbol K, instead of "degree kelvin" for the unit. However, the practical difficulties in realising this definition, requiring a sample of pure water of well-defined isotopic composition and the development of new primary methods of thermometry, led to the adoption of a new definition of the kelvin.

It is now defined by taking the fixed numerical value of the Boltzmann constant k to be $1.380\,649 \times 10^{-23}$ when expressed in the unit J K^{-1}, which is equal to kg m^2 s^{-2} K^{-1}, where the kilogram, metre and second are defined above.

The kelvin is redefined with no immediate effect on temperature measurement practice or on the traceability of temperature measurements, and for most users, it will pass unnoticed. The redefinition lays the foundation for future improvements.

Amount of substance

The mole, symbol mol, is the SI unit of amount of substance.

The former definition was the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kilogram of carbon 12.

In the new system, one mole contains exactly $6.022\,140\,76 \times 10^{23}$ elementary entities. This number is the fixed numerical value of the Avogadro constant, N_A, when expressed in the unit mol^{-1} and is called the Avogadro number. The amount of substance, symbol n, of a system is a measure of the number of specified elementary entities. An elementary entity may be an atom, a molecule, an ion, an electron, any other particle or specified group of particles.

Atomic weights will be unaffected by this change in definition and the molar mass constant M_u will still be1 g/mol, although now with a measurement uncertainty. This uncertainty will be so small that the revised definition of the mole will not require any change to common practice.

Luminous intensity

The candela, symbol cd, is the SI unit of luminous intensity in a given direction.

The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of 1/683 watt per steradian. This still the case with the new system of units.

1.1.3 Realisation

As this fundamental constants are not easy to measure accurately, there are also practical realisations of the seven base units. See the BIPM for details.

1.1.4 Derived quantities

All other quantities may be called "derived quantities" and are measured using derived units, which can be written as products of powers of base units. Twenty-two derived units are given a special name, as listed in Table 1.1.

Although the hertz and the becquerel are both equal to the reciprocal second, hertz is used only for periodic phenomena, and becquerel is used only for stochastic processes in radioactive decay.

Some non-SI units are still widely used. A few, such as the minute, hour and day as units of time, will always be used because they are part of our culture. Others are used for historical reasons, to meet the needs of special interest groups, or because there is no convenient SI alternative. This is the case of electronvolt, litre, tonne... When non-SI units are used, the correspondence to the SI should always be quoted.

When symbols for units are named after an individual (for example: ampere, A; kelvin, K; hertz, Hz; or coulomb, C) they should begin with a capital letter. In all other cases, except the litre, they begin with a lower case letter (for example: metre, m; second, s; or mole, mol). The symbol for the litre is an exception; either a lower case letter 'l' or a capital 'L' may be used, the capital is allowed in this case to avoid confusion between the lower case letter l and the number one, 1.

1.1.5 Prefix

All the units and symbols that you can use are the ones of the International system of units. No other!

Table 1.1: Derived units with special names in the SI

Derived quantity	**Name of unit**	**Symbol for unit**	**Expression in terms of other units**
plane angle	radian	rad	m/m
solid angle	steradian	sr	m^2/m^2
frequency	hertz	Hertz	s^{-1}
force	newton	N	kg m s^{-2}
pressure, stress	pascal	Pa	N/m^2 = kg m^{-1} s^{-2}
energy, work, amount of heat	joule	J	N m = kg m^2 s^{-2}
power, radiant flux	watt	W	J/s = kg m^2 s^{-3}
electric charge	coulomb	C	A s
electric potential difference	volt	V	W/A = kg m^2 s^{-3} A^{-1}
capacitance	farad	F	C/V = kg^{-1} m^{-2} s^4 A^2
electric resistance	ohm	Ω	V/A = kg m^2 s^{-3} A^{-2}
electric conductance	siemens	S	A/V = kg^{-1} m^{-2} s^3 A^2
magnetic flux	weber	Wb	V s = kg m^2 s^{-2} A^{-1}
magnetic flux density	tesla	T	Wb/m^2 = kg s^{-2} A^{-1}
inductance	henry	H	Wb/A = kg m^2 s^{-2} A^{-2}
Celsius temperature	degree Celsius	°C	K
luminous flux	lumen	lm	cd sr = cd
illuminance	lux	lx	lm/m^2 = cd m^{-2}
activity referred to a radionuclide	becquerel	Bq	s^{-1}
absorbed dose	gray	Gy	J/kg = m^2 s^{-2}
dose equivalent	sievert	Sv	J/kg = m^2 s^{-2}
catalytic activity	katal	kat	mol s^{-1}

Also avoid "billion", "trillion"... that have different meaning in US and UK. One billion is 10^9 in US and 10^{12} in UK or France. Rather use the powers of 10 or the prefix symbols of the International system. See table 1.2. The last four were introduced in 1991.

Prefix symbols are printed in roman (upright) type, as are unit symbols, regardless of the type used in the surrounding text, and are attached to unit symbols without a space between the prefix symbol and the unit symbol. With the exception of da (deca), h (hecto), and k (kilo), all multiple prefix symbols are capital (upper case) letters, and all submultiple prefix symbols are lower case letters. All prefix names are printed in lower case letters, except at the beginning of a sentence. Similarly prefix names are also inseparable from the unit names to which they are attached. Thus, for example, millimetre, micropascal, and meganewton are single words.

Note the grammatical rule: one writes 2 bar or two bars. In the first case, it is symbol and in the second, the name.

1.2 Vocabulary

As the results and their description should be unambiguous, the vocabulary is also fixed in the so-called "VIM" which stands for *Vocabulaire International de Métrologie*. The latest version is available online on the homepage of the BIPM: `http://www.bipm.org`.

Table 1.2: Prefix names and symbol of the International system.

Factor	Name	Symbol	Factor	Name	Symbol
10^{1}	deca	da	10^{-1}	deci	d
10^{2}	hecto	h	10^{-2}	centi	c
10^{3}	kilo	k	10^{-3}	milli	m
10^{6}	mega	M	10^{-6}	micro	μ
10^{9}	giga	G	10^{-9}	nano	n
10^{12}	tera	T	10^{-12}	pico	p
10^{15}	peta	P	10^{-15}	femto	f
10^{18}	exa	E	10^{-18}	atto	a
10^{21}	zetta	Z	10^{-21}	zepto	z
10^{24}	yotta	Y	10^{-24}	yocto	y

I will not copy and paste all the definitions because it would be as boring as reading a dictionary. But I do recommend you to give a look at the VIM in order to make sure that you correctly use the words. I am sure that you will learn something. It is in both French and English which is very convenient.

1.2.1 Few words

They are few words that are necessary for a good understanding of this lecture. I'll introduce them here. Many other words will be introduced later with their translation in French.

Measurement

The French word "mesure" has several meanings in everyday French language. For this reason, it is not used without further qualification. The French word "mesurage" has been introduced to describe the act of measurement. Nevertheless the word "mesure" occurs many times in the VIM, forming terms without ambiguity. Examples are: instrument de mesure, unité de mesure, méthode de mesure...

In English the word **measurement** is used without any ambiguity. The VIM definition reads: *"process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity".*

Measurand

The **measurand** [mesurande] is defined as the *"quantity intended to be measured"*. This sounds very simple, but in fact you have to be as specific as possible. If you say that you want to measure the temperature, it is not sufficient. Temperature of what? The air. Which air? In the room. Well, when? Average temperature. Ok, average over the day, the week, the year ? With which frequency of measurement? Where in the room? And so on. Always think to the users of your data. They'll need a lot of information to give a meaning to your result.

The VIM adds a first note: *"The specification of a measurand requires knowledge of the kind of quantity, description of the state of the phenomenon, body, or substance carrying the quantity, including any relevant component, and the chemical entities involved."*

You also have to note that *"the measurement, including the measuring system and the conditions under which the measurement is carried out, might change the phenomenon, body, or substance such that the quantity being measured may differ from the measurand as defined. In this case adequate correction is necessary".*

For example, the potential difference between the terminals of a battery may decrease if the internal impedance of the voltmeter is not large enough.

Measurement uncertainty

The **measurement uncertainty** [incertitude de mesure] is a *"non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used".*

As the estimation of the uncertainty is detailed in the next chapter, I won't be more specific here. But you have to keep in mind that the uncertainty is used to estimate the quality of a measurement and further of a laboratory.

1.3 Writing and rounding numbers

1.3.1 Writing

As usual, there should be no ambiguity when you write numbers. The decimal sign in the English text is the point, and the comma on the line is the decimal sign in the French text.

Because the comma is widely used as the decimal marker outside the United States, it should not be used to separate digits into groups of three. Instead, digits should be separated into groups of three, counting from the decimal marker towards the left and right, by the use of a thin, fixed space. However, this practice is not usually followed for numbers having only four digits on either side of the decimal marker except when uniformity in a table is desired[1]. Example: 1 234 567.891 012.

When powers of tens are used, the writing rule is to put a multiplication sign before 10: one should write 5×10^6 rather than 5.10^6 that is ambiguous and thus should be avoided.

Finally, the results should be printed in a form that is easily understandable by the readers. It may not be the same in a scientific publication and for the general audience.

1.3.2 Rounding

Numerical values should not be written with too many digits. For the uncertainty, two digits are generally enough. Then, the numerical value of the result should be rounded to the same increment as the uncertainty.

Rounding means replacing a number by its nearest integer multiple of some increment.

Example 1 Rounding to an increment 0.1, 12.235 becomes 12.2 and 12.251 becomes 12.3.

How to do when the number finishes by 50? A French standard (NF X 02-001) suggests to choose nearest even number in order to minimize the error with long series of numbers. Other rules are possible.

Example 2 Rounding to an increment 0.1, 12.25 becomes 12.2, and 12.35 becomes 12.4.
Rounding to an integer, 109.50 becomes 110 and 108.50 becomes 108.

If after rounding the last digit is 0, it should be kept: 60 and 6×10^1 do not mean the same thing. 60 has two significant digits whereas 6×10^1 only has one. 60 is equivalent to 6.0×10^1.

Rounding of values should be carried out only at the end of the calculation, to avoid rounding errors. For example, if 2.346 is rounded up to 2.35 at an early stage in a calculation, it could later be rounded up to 2.4. But if 2.346 is used throughout a calculation it would be correctly rounded to 2.3 at the final stage.

[1] See National Institute of Standards and Technology: Guide for the Use of the International System of Units (SI).

Chapter 2

Evaluation of the uncertainty according to the GUM

"In this world nothing can be said to be certain, except death and taxes"
Benjamin Franklin in a letter to Jean-Baptiste Leroy (1789)

Over the years, many different approaches to evaluating and expressing the uncertainty in measurement results have been used. Because of this lack of international agreement on the expression of uncertainty in measurement, in 1977, the International Committee for Weights and Measures (CIPM, Comité International des Poids et Measures), the world's highest authority in the field of metrology, asked the International Bureau of Weights and Measures (BIPM, Bureau International des Poids et Mesures), to address the problem in collaboration with the various national metrology institutes and to propose a specific recommendation for its solution. This led to the development of Recommendation INC-1 (1980) by the Working Group on the Statement of Uncertainties convened by the BIPM, a recommendation that the CIPM approved in 1981 and reaffirmed in 1986 via its own Recommendations 1 (CI-1981) and 1 (CI-1986).

This recommendation is a brief outline rather than a detailed prescription. Consequently, the CIPM asked the International Organization for Standardization (ISO) to develop a detailed guide based on the recommendation because ISO could more easily reflect the requirements stemming from the broad interests of industry and commerce.

The end result is the 100-page Guide to the Expression of Uncertainty in Measurement (or GUM as it is often called). It was published in 1993 (corrected and reprinted in 1995) by ISO in the name of the seven international organizations:

BIPM Bureau International des Poids et Mesures

IEC International Electrotechnical Commission

IFCC International Federation of Clinical Chemistry

ISO International Organization for Standardization

IUPAP International Union of Pure and Applied Physics

OIML International Organization of Legal Metrology.

ISO published the French translation of the GUM in 1995. I am sure that the GUM is now available in your own language. The GUM was was adopted as an European pre-standard in 1999 (ENV 13005:1999) and has the status of a French standard under the reference NF ENV 13005. The GUM has also been adopted by the American National Standards Institute (ANSI) as an American National Standard. Its official designation is ANSI/NCSL Z540-2-1997 and its full title is *American National Standard for Expressing Uncertainty–U.S. Guide to the Expression of Uncertainty in Measurement.*

This short historical overview is based on the presentation of the National Institute of Standards and Technology in the US [3]. The aim of this chapter is to present the GUM [5] that can be downloaded in French and English.

Note that there are several supplements to this Guide that are are also available on line. The Gum has also been adapted to chemistry: `http://www.measurementuncertainty.org/`

In addition to these officials texts, some textbooks were useful to prepare this lecture:

- *Estimer l'incertitude. Mesures - Essais*, Christophe Perruchet et Marc Priel, AFNOR 2000.
- *An introduction to error analysis, the study of uncertainties in physical environments* John R. Taylor, University Science Books, (1982)
- *An Introduction to Uncertainty in Measurement*, Les Kirkup and Bob Frenkel, Cambridge University Press (2006). One of the rare books on uncertainties that is based on the GUM and other standards.

There are many others.

2.1 When the measurement result is a random variable

2.1.1 Introduction

There is no perfect measurement. Two groups of students asked to measure precisely the same quantity might not find the same result. Consider an example as simple as the length of a table: the measurement done by the students is not always at the same place and/or they don't use the same instrument. Many other external factors can influence the result.

As a consequence, results of successive measurements vary in a unpredictable way. They are said to be random. Which value should be kept? If all the results are close to each other, we will have confidence in the choice. Not if they are widespread. An experimentalist has therefore two tasks: choose the best estimate for the result and quantify the dispersion of the results with universal mathematical tools.

Measurements are always done to answer to a question. Sometimes the best estimate and the dispersion of the results are not enough. Think to a pollutant in water for example: one needs to be sure that the measured level is below the acceptable threshold. In other words, one needs to estimate the probability that the contamination is larger than the threshold to take the decision to allow or not the sale of the product. This means that we might also need to estimate the **probability density function** [fonction de densité de probabilité] of the values that can be taken by the quantity.

Mathematically, measurement results are considered as a random variable and we will use probability theory to achieve our goal.

2.1.2 Mathematical tools

Let's consider the example of weighing: A first measurement gives $x_1 = 724.173$ g. After repeating the whole process, including calibration, one finds $x_2 = 724.190$ g. And so on...

We can then plot a histogram of the occurrences of the results, see fig. 2.1. Ordinate corresponds to n_{ab}, the number of values in the interval $[a, b]$, or better to the **frequency** [fréquence] n_{ab}/n, with n the total number of measurements. In an ideal situation, when the number of results goes to infinity, such a histogram tends to a continuous curve called **probability density function** [fonction densité de probabilité].

Mathematicians have defined tools to characterise such curves. Assuming that we have n results of measurement of a given quantity, we admit that when $n \to \infty$, the **average value** or **sample mean** [valeur moyenne] tends to a limit μ called **expected value** or **mathematical expectation** [espérance mathématique],

$$\mu = \mathrm{E}(x) = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{x_i}{n}. \tag{2.1}$$

The dispersion of the values is characterised by the **variance** [variance]

$$\sigma^2 = \mathrm{E}(x-\mu)^2 = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{n}. \tag{2.2}$$

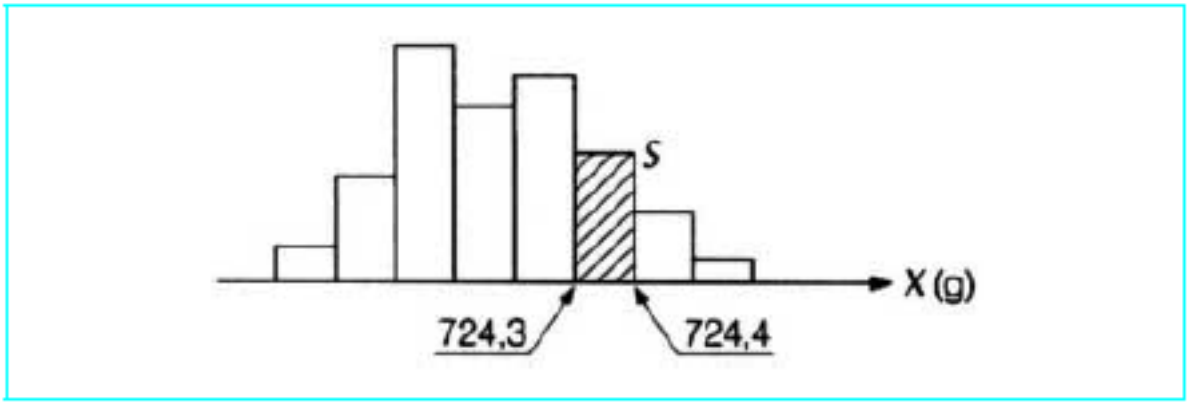


Figure 2.1: Histogram with the number of occurrences of the values of a series of measurements (figure from M. Neuilly)

The square root of the variance is called the **standard deviation** [écart-type]. It has the same dimension as x.

The most common probability density function is the so-called Gaussian or normal distribution. This is a consequence of the central limit theorem. Then, the probability to measure a value x is equal to $f(x)\,dx$ with

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp - \left[\frac{(x-\mu)^2}{2\sigma^2}\right], \tag{2.3}$$

where μ et σ are constants. This function has a bell shape represented in Fig. 2.2. Its mean value is μ and its variance σ^2. The full width at half maximum is $FWHM = 2,36\sigma$.

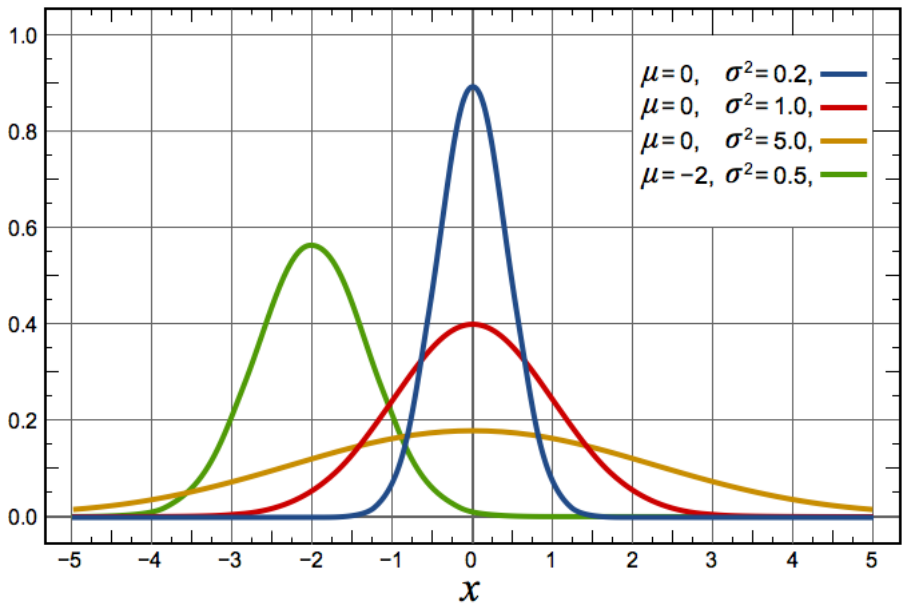


Figure 2.2: Gaussian distribution for various values of the parameters. (Figure reproduced from Wikipedia).

If the distribution of the values is Gaussian, it is very easy to estimate any probability. First of all, this distribution is normalized to one,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1. \tag{2.4}$$

It is also possible to count the number of values included in the interval $[\mu - \sigma, \mu + \sigma]$:

$$\int_{\mu-\sigma}^{\mu+\sigma} f(x)\,dx = 0,68. \tag{2.5}$$

This is represented in Fig. 2.3. The interval $[\mu - 2\sigma, \mu + 2\sigma]$ includes 95,44% of the values and $[\mu - 3\sigma, \mu + 3\sigma]$, 99,73%.

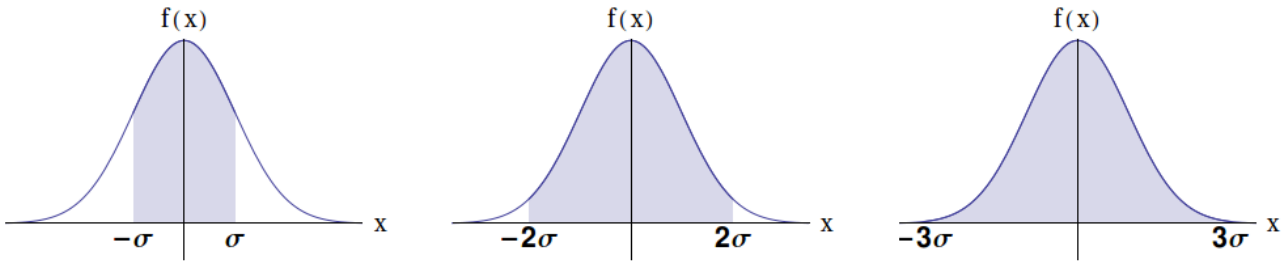


Figure 2.3: Gaussian distribution with the events included in a given interval for various values of the coverage factor k. The mean value is fixed to zero.

Note that these mathematical tools and properties are defined in an ideal situation of an infinity of results. In reality we can only deal with a limited number of results that are not accurate. Thus, one has to define estimators of the three quantities that are meaningful for an experimentalist: the mean value, the standard deviation and the probability density function.

2.2 Estimation of the uncertainty

Previous mathematical tools were defined for an idealistic case corresponding to very large numbers of observations. In the real world, one deals with finite numbers and we can only expect to estimate the expectation, the variance or the probability density function.

2.2.1 Uncertainty versus error

According to the GUM, *"the word "uncertainty" means doubt, and thus in its broadest sense "uncertainty of measurement" means doubt about the validity of the result of a measurement. Because of the lack of different words for this general concept of uncertainty and the specific quantities that provide quantitative measures of the concept, for example, the standard deviation, it is necessary to use the word "uncertainty" in these two different senses."*

In the Guide, *"great care is taken to distinguish between the terms "error" and "uncertainty". They are not synonyms, but represent completely different concepts; they should not be confused with one another or misused."* Error corresponds to measured quantity value minus the true value that is unknown. This differs from the uncertainty that characterises the dispersion of the quantity values being attributed to a measurand.

One can have a result with a large uncertainty that is close to the true value, without knowing it. On the other hand, it can happen that a measurement with a small uncertainty is wrong and the mean value is far from the true value, resulting in a large error.

To detect and correct errors, we should test the instruments and procedures. This can be done with measurement standards, other instruments... Accredited laboratories generally proceed to cross-checks with other laboratories. Pioneer measurements in research have to be validated by another laboratory.

Here, we will only focus on uncertainties.

2.2.2 Mean value

Assuming that we have n values x_i of a given measurand, we generally choose the mean value as the best estimate of the value of the quantity,

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}. \tag{2.6}$$

We will see later why. This mean value is an unbiaised estimator of the expectation.

2.2.3 Type A evaluation of standard uncertainty

The GUM defines the **type A evaluation of standard uncertainty** [évaluation de type A de l'incertitude] as *a method of evaluation of uncertainty by the statistical analysis of series of observations.*

Estimator of the variance and standard deviation

To evaluate if the distribution of values obtained by repeated observations are close to the average value or widespread, one uses the *variance* [variance]. For a limited number of measurements, the best estimate of the variance is given by

$$\boxed{s^2(x) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}.} \tag{2.7}$$

The denominator is $n-1$ when we have a finite number of values. It differs from the mathematical definition of the variance seen in the previous section. This correction is due to Friedrich Wilhelm Bessel and is called Bessel correction to have an unbiaised estimator of the variance.

$s = \sqrt{s^2}$ is called standard deviation [écart-type]. It has the same dimension as the measured quantity whereas the variance has the dimension of the square of the measured quantity. Be careful. The standard deviation is used to estimate the statistical contribution to the uncertainty.

For a widespread distribution resulting in large departures from the mean value, the variance and the standard deviation are large whereas for a narrow distribution, they are small.

Variance and standard deviation of the mean value

As an introductory example, consider the results obtained by all students to a national examination. They range between 0 and 20 in France or 0 and 100 in other countries. But if you consider the average result for a given high school or university. None of the mean values will be 0 or 20 for the French case. The dispersion of the mean values is narrower than the dispersion of the values.

Thus, one should distinguish the variance associated to a given value $\mathrm{Var}(x_i)$ and the variance of the mean value $\mathrm{Var}(\bar{x}) = \frac{1}{n}\mathrm{Var}(x_i)$. The larger n the smaller the variance of the mean value. This is why we choose the mean value as the best estimate of the value of a quantity. Eventually, the variance of the mean value for a series of results is estimated by

$$\boxed{s^2(\bar{x}) = \frac{1}{n}s^2(x_i) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2.} \tag{2.8}$$

With a series of values $\{x_1, \ldots, x_n\}$, we cannot say that one result is better than another one. If one picks up randomly one value as an estimate of the mesurand, associated uncertainty is $s(x_i)$. This is the case for example with te output of a factory. Items are sold one by one. But if one selects the mean value as the result of an experiment, the associated uncertainty $s(\bar{x})$ is smaller.

Proof: If a quantity X is the sum of two independent quantities A and B, $X = A + B$, then $\mathrm{Var}(X) = \mathrm{Var}(A) + \mathrm{Var}(B)$. This is Bienaymé formula. The demonstration is straightforward. We can deduce that

$$\mathrm{Var}(\bar{x}) = \mathrm{Var}(\sum_{i=1}^{n}\frac{x_i}{n}) = \sum_{i=1}^{n}\mathrm{Var}(\frac{x_i}{n}) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(x_i) = \frac{1}{n}\mathrm{Var}(x_i). \tag{2.9}$$

Variations of the estimator of the variance

Imagine that a group of students perform an experiment, repeating the observations a given number of times to get a mean value $\bar{x}_A$ and a variance $s^2(x_A)$. Another group of students do the same experiment and gets $\bar{x}_B$ and $s^2(x_B)$. There is no reason that these values are the same as the ones of group A. We consider other groups C, D... Then we can do a statistical analysis of the average values and variances.

Mean value of the mean values gives the same result as the mean of all values. Thus, the estimator of the expectation is said to be unbiaised. We can also calculate the mean value of the variances obtained by various groups of students. This is how Bessel obtained Bessel's correction to define an unbiaised estimator.

Let's consider now the variance of the variances obtained by the groups of students. This will give an indication of the quality of the estimator by simply repeating the exercise. The calculation is lengthy but not difficult (see Taylor's book [10]): relative uncertainty in $s^2(x_i)$ is equal to $1/\sqrt{2(n-1)}$. For $n = 10$ we get 0.24 or 24% which is very large. Even with $n = 100$, the relative uncertainty in the estimator of the variance is still 7%. n should be larger than 5 001 to have this relative uncertainty lower than 1%.

Annex E of the GUM gives a more complicated formula for the relative uncertainty in $s(\bar{x})$. I won't present the details, just the results in Table 2.1. Once again, we can see that when we are dealing with a small number of

number of observations	relative uncertainty on $s(\bar{x})$
2	76%
3	52%
4	42%
5	36%
10	24%
20	16%
30	13%
50	10%

Table 2.1: Relative uncertainty on $s(\bar{x})$ as a function of the number of observations.

observations, the estimation of the statistical contribution to the uncertainty is very bad. Below 10 observations, it is worthwhile to add few more observations. Whereas beyond 10 observations it can be expensive to improve the quality of the estimator of the variance.

Implementation

One should also check that $x_i - \bar{x}$ has two significative digits.

Formula (2.8) is not easy to use because one should first calculate the mean value before subtracting it to each value. Other equivalent formulas are available:

$$\boxed{s^2(\bar{x}) = \frac{1}{n(n-1)}\left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2\right] = \frac{1}{n(n-1)}\left[\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2\right].} \tag{2.10}$$

Their meaning is not obvious but they are easier to implement.

Imagine that the series of data reads $x_1 = 1.799\ 50, x_2 = 1.799\ 42, x_3 = 1.799\ 63, \ldots$ Il is simpler to use $y_i = x_i - 1.799\ 00$. We can easily show that

$$s^2(\bar{x}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} (y_i - \bar{y})^2. \tag{2.11}$$

Concluding remarks

Statistical analysis of the data does not contain all the uncertainty. Imagine the extreme case with all the x_i identical because of a inaccurate instrument. The standard deviation will be zero! This does not mean that

there is no uncertainty. One has to find a way to evaluate the inaccuracy of the instrument.

If the zeroing operation is not repeated, it has to be added. If it is repeated, it is already included in the statistical analysis. It is generally a difficult task to define the border between what is included in the statistical analysis and what is not included. A careful definition of the mesurand is necessary.

How to evaluate the other sources of uncertainty? How to add them? This will be addressed in the next sections.

2.2.4 Type B evaluation of standard uncertainty

The GUM defines the **type B evaluation of standard uncertainty** [évaluation de type B de l'incertitude] as *a method of evaluation of uncertainty by means other than the statistical analysis of series of observations.* It is generally written u.

Various cases

Sources of information are numerous:

- previous results;
- specifications from the maker of the instrument,
- reference values from tables or books...

If all these uncertainties are evaluated following the recommendations of the GUM, no problem. One just copies and paste the value, taking into account the coverage factor. If the evaluation is older, one might have to reconsider the value.

There are also cases that requires calculations.

Reading uncertainty

Let us consider the burette on the picture 2.4. The meniscus is between two successive values, a_- and a_+. As we have no reason to say that one value is more probable than an other, we suppose that the distribution is uniform in the interval $[a_-, a_+]$.

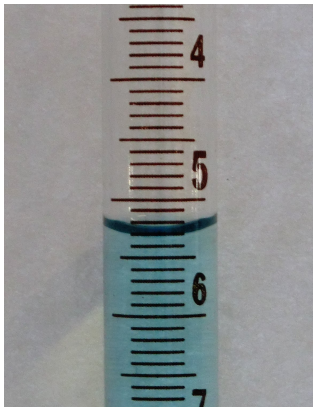


Figure 2.4: Meniscus in a burette.

With a uniform distribution in the interval $[a_-, a_+]$, the mean value is then $\mu = \frac{a_+ + a_-}{2}$ and the variance

$$u^2 = \frac{(a_+ - a_-)^2}{12} = \frac{a^2}{3}, \tag{2.12}$$

if $2a$ is the length of the interval.

With a trapezoidal distribution, with a base length of $2a$ and a summit length of $2\beta a$, the variance reads

$$u^2 = \frac{a^2(1+\beta^2)}{6}. \tag{2.13}$$

$\beta = 1$ corresponds to the previous rectangular distribution and $\beta = 0$ to a triangular distribution. The trapezoidal distribution is used when we are certain that the results is between a_- and a_+ but there are uncertainties in these limits. With class A Schellbach burette, international standards in chemistry recommend to use a triangular distribution in order to give more weight to the nominal value. Thus, the reading uncertainty is $u = a/\sqrt{6}$.

For any other distribution, one can evaluate the mean value and the variance by calculating

$$\mu = \frac{\int_{-\infty}^{\infty} xP(x)\,dx}{\int_{-\infty}^{\infty} P(x)\,dx} \qquad \text{et} \qquad \sigma^2 = \frac{\int_{-\infty}^{\infty} (x-\mu)^2 P(x)\,dx}{\int_{-\infty}^{\infty} P(x)\,dx}. \tag{2.14}$$

Supplement 1 to the GUM lists other cases with results and applications.

2.3 Propagation of the uncertainty

Most of the time we don't directly measure the quantity of interest. One uses mathematical operations to deduce the quantity from other observables. For example, an electric resistance is obtained by measuring V and I and then, calculating V/I. What is the uncertainty in the **output quantity** [grandeur de sortie] knowing the uncertainties in the **input quantities** [grandeurs d'entrée]? The GUM proposes a methodology.

2.3.1 Methodology

Let us assume that the output quantity Z depends on N input quantities $w_1, \ldots, w_N$. The relationship is given by a mathematical formula: $Z = f(w_1, \ldots, w_N)$. For the sake of simplicity, we will assume in the calculation that there are only two input quantities, w_1 and w_2.

Differencing Z, one gets,

$$dZ = \frac{\partial f}{\partial w_1} dw_1 + \frac{\partial f}{\partial w_2} dw_2. \tag{2.15}$$

Noting $\mu_z = f(\bar{w}_1, \bar{w}_2)$ the expected value, we have, to the first order,

$$z - \mu_z \simeq \frac{\partial f}{\partial w_1}(w_1 - \bar{w}_1) + \frac{\partial f}{\partial w_2}(w_2 - \bar{w}_2). \tag{2.16}$$

Taking the square, one gets,

$$(z-\mu_z)^2 = \left(\frac{\partial f}{\partial w_1}\right)^2 (w_1 - \bar{w}_1)^2 + \left(\frac{\partial f}{\partial w_2}\right)^2 (w_2 - \bar{w}_2)^2 + 2\frac{\partial f}{\partial w_1}\frac{\partial f}{\partial w_2}(w_1 - \bar{w}_1)(w_2 - \bar{w}_2). \tag{2.17}$$

Then, taking the mean value, one finally gets,

$$u^2(z) = \left(\frac{\partial f}{\partial w_1}\right)^2 u^2(w_1) + \left(\frac{\partial f}{\partial w_2}\right)^2 u^2(w_2) + 2\frac{\partial f}{\partial w_1}\frac{\partial f}{\partial w_2} u(w_1, w_2). \tag{2.18}$$

The last term corresponds to the possible correlation between w_1 and w_2. We will come back on this point later.

The generalisation to N input quantities $w_1, \ldots, w_N$ reads

$$\boxed{u^2(z) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial w_i}\right)^2 u^2(w_i) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \frac{\partial f}{\partial w_i}\frac{\partial f}{\partial w_j} u(w_i, w_j).} \tag{2.19}$$

This is the formula to use to evaluate the propagation of the uncertainty. By taking the square root, we can access to the **standard uncertainty** [incertitude type], $u(z)$, expressed as a standard deviation.

Let us first consider simple cases with independent input quantities.

2.3.2 Examples with independent input quantities

Example 1

Let us come back the example given in the introduction about the measurement of an electric resistance $R = V/I$. Assuming that we know both V and I with their respective uncertainties, what is the uncertainty in R? As V is measured with a voltmeter and I with an ampere-meter, there is no reason that these two input quantities are correlated.

Applying formula (2.19), we simply get,

$$u^2(R) = \frac{u^2(V)}{I^2} + \frac{V^2}{I^4}u^2(I). \tag{2.20}$$

Considering the relative uncertainty, the formula is even simpler,

$$\frac{u(R)}{R} = \sqrt{\frac{u^2(V)}{V^2} + \frac{u^2(I)}{I^2}}. \tag{2.21}$$

Considering that both V and I are known with a relative precision of 1%, one gets

$$\frac{u(R)}{R} = 1.4\%. \tag{2.22}$$

Example 2

Let us consider a class B burette like on the picture 2.4.

The figures written on the burette might not exact. When it is written 5 ml, it could be 4.99 ml or another value. Uncertainty in the written figures should be given by the manufacturer who might have done a statistical study on the output of its factory.

In addition to this problem, there is also the reading uncertainty due to the meniscus. If one reads a volume V_1 at the bottom of the meniscus, the real value is $V_1 + X_1$, X_1 being the reading error. The width of the meniscus corresponds to the smallest scale division of the burette. Assuming that X_1 is uniformly distributed on this smallest division d, its mean value is $d/2$ and its standard deviation $u(X) = d/\sqrt{12}$.

To determine the volume V of solution delivered into a vial, one needs two reading. Then,

$$V = (V_1 + X_1) - (V_2 + X_2), \tag{2.23}$$

where V_1 and V_2 correspond to the inficated values and the X's to the corrections due to the reading error. The mean values of the X's cancel out.

The uncertainty in V can easily be determined from

$$u^2(V) = u^2(V_1) + u^2(X_1) + u^2(V_2) + u^2(X_2). \tag{2.24}$$

Uncertainties $u(V_1)$ et $u(V_2)$ are given by the manufacturer of the burette and the reading uncertainties $u(X_i)$ are given above.

Example 3

Let us consider a simple calibration between two quantifies such as $Z = \alpha + \beta Q$, where α and β are two parameters with uncertainties. For each measurement, one has

$$z_k = \alpha + \beta q_k. \tag{2.25}$$

If we are interested in the mean value of Z,

$$\bar{z} = \frac{1}{n}\sum_{k=1}^{n} z_k = \frac{1}{n}\sum_{k=1}^{n}(\alpha + \beta q_k) = \alpha + \beta\bar{q}, \tag{2.26}$$

that is considered as the best estimation of z, we can directly estimate its uncertainty. Noticing that $\bar{z} = f(\alpha, \beta, q_k)$ and assuming that all the input values are independent, we can directly apply the propagation formula to get

$$u^2(\bar{z}) = u^2(\alpha) + \bar{q}^2 u^2(\beta) + \sum_{k=1}^{n} \left(\frac{\beta}{n}\right)^2 s^2(q_k) \tag{2.27}$$

$$= u^2(\alpha) + \bar{q}^2 u^2(\beta) + \beta^2 \frac{s^2(q_k)}{n}. \tag{2.28}$$

In this expression we can easily distinguish A and B type sources of the uncertainty. Here, $u^2(\alpha)$ and $u^2(\beta)$ were estimated by a previous calibration process or by the maker of an apparatus or come from another source. The third term with s^2 comes from a statistical analysis. The higher n the lower is the contribution to the variance. This is not the case for the first two terms. The GUM does not distinguish between A and B type of uncertainties that are both taken into account in the formula.

This example is also interesting if one considers the case where $\alpha = 0$ and $\beta = 1$ that leads to

$$s^2(\bar{q}) = \frac{1}{n} s^2(q_k). \tag{2.29}$$

One finds again that the variance of the mean value is not the variance of the value. This shows the interest of enlarging the number of observations.

2.3.3 Method of the logarithm in case of a product

Method

Be an output quantity y depending on two input quantities a and b that might be correlated. Suppose that $y = a^n b^m$. Derivation of a product is complicated, so we take the logarithm to get a sum:

$$\ln y = n \ln a + m \ln b \tag{2.30}$$

and differentiate it

$$\frac{dy}{y} = n\frac{da}{a} + m\frac{db}{b}. \tag{2.31}$$

The square leads to

$$\left[\frac{dy}{y}\right]^2 = n^2 \left[\frac{da}{a}\right]^2 + m^2 \left[\frac{db}{b}\right]^2 + 2nm\frac{dadb}{ab}. \tag{2.32}$$

Averaging over all possibilities, one finally gets that

$$\left[\frac{u(y)}{y}\right]^2 = n^2 \left[\frac{u(a)}{a}\right]^2 + m^2 \left[\frac{u(b)}{b}\right]^2 + 2nm\frac{u(a,d)}{ab}, \tag{2.33}$$

which is a simple expression. Of course, if a and b are independent, $u(a, b) = 0$.

Application to the first example

We had $R = V/I$. Thus $\ln R = \ln V - \ln I$. Differentiation leads to

$$\frac{dR}{R} = \frac{dV}{V} - \frac{dI}{I}, \tag{2.34}$$

and the square and average over all possibilities lead to

$$\left[\frac{u(R)}{R}\right]^2 = \left[\frac{u(V)}{V}\right]^2 + \left[\frac{u(I)}{I}\right]^2, \tag{2.35}$$

as the two input quantities are independent.

2.3.4 Inclusion of the correlations

Two input quantifies can be correlated if they are obtained,

- with the same instrument,
- with the use of the same standard, or
- the same reference value with a significative uncertainty.

The correlation between two quantities is estimated by the **covariance** [covariance], $u(w_i, w_j)$, or better, by the **correlation coefficient** [coefficient de corrélation],

$$r(w_i, w_j) = \frac{u(w_i, w_j)}{u(w_i)u(w_j)}, \tag{2.36}$$

that is always between -1 and 1. Naturally, if w_i et w_j are independent, $r(w_i, w_j) = 0$. On the contrary, if $r = \pm 1$, there is a linear relationship between w_i and w_j that reads $w_j = Aw_i + B$.

It is worth noticing that the covariance has the dimension of a variance but can be negative.

A deep analysis of the measurement process is necessary to know if two quantifies w_i and w_j are correlated or not. When there is a doubt on the correlation, we should always estimate the covariance and the correlation coefficient to decide whether we can neglect it or not.

There are several ways to evaluate the covariances or correlation coefficients.

Type A evaluation

Be two mean values $\bar{a}$ and $\bar{b}$ that are estimators of the expectations μ_a and μ_b of two quantities A and B. Let us assume that $\bar{a}$ and $\bar{b}$ are calculated from n couples of independent simultaneous observations of A and B done in similar conditions. Then, the covariance is estimated by

$$s(a_l, b_l) = \frac{1}{n-1} \sum_{k=1}^{n} (a_k - \bar{a})(b_k - \bar{b}). \tag{2.37}$$

The covariance of the mean values, $\bar{a}$ and $\bar{b}$, is obtained by dividing by n

$$s(\bar{a}, \bar{b}) = \frac{1}{n} s(a_l, b_l) = \frac{1}{n(n-1)} \sum_{k=1}^{n} (a_k - \bar{a})(b_k - \bar{b}). \tag{2.38}$$

The correlation coefficient can easily be obtained by dividing by the respective standard deviations.

Propagation

If the two quantities A and B are not directly measured but both depend on N uncorrelated input quantities w_i, the covariance is,

$$u(a_l, b_l) = \sum_{i=1}^{N} \frac{\partial A}{\partial w_i} \frac{\partial B}{\partial w_i} u^2(w_i). \tag{2.39}$$

To get the correlation coefficient, one should divide, as usual, by the standard deviations obtained from the following variances

$$u^2(a_l) = \sum_{i=1}^{N} \left(\frac{\partial A}{\partial w_i} \right)^2 u^2(w_i) \tag{2.40}$$

$$u^2(b_l) = \sum_{i=1}^{N} \left(\frac{\partial B}{\partial w_i} \right)^2 u^2(w_i). \tag{2.41}$$

If the input quantities w_i's are correlated, the expression is more complicated,

$$u(a_l, b_l) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial A}{\partial w_i} \frac{\partial B}{\partial w_j} u(w_i, w_j), \tag{2.42}$$

with $u(w_i, w_i) = u^2(w_i)$. Even if the formula does not look more complicated, it is a sum of N^2 terms whereas the previous formula, eq. (2.39), it is a sum of N terms.

Example 1: two faces of a parallelepiped

Be a rectangular parallelepiped with edges of length a, b and c. We need to know two surfaces $S_1 = ac$ and $S_2 = bc$.

We shall assume that a, b and c are independent. As both output quantities S_1 and S_2 depend on a common input quantity c, they are correlated. The covariance is calculated with the help of eq. (2.39) that leads to

$$u(S_1, S_2) = \frac{\partial S_1}{\partial a} \frac{\partial S_2}{\partial a} u^2(a) + \frac{\partial S_1}{\partial b} \frac{\partial S_2}{\partial b} u^2(b) + \frac{\partial S_1}{\partial c} \frac{\partial S_2}{\partial c} u^2(c). \tag{2.43}$$

As S_1 does not depend on b, the related partial derivative is zero. Similarly for S_2 and a. Therefore,

$$u(S_1, S_2) = ab\, u^2(c), \tag{2.44}$$

where the only common input quantity to S_1 and S_2 contributes, as expected.

To calculate the correlation coefficient, $r(S_1, S_2)$, we need $u(S_1)$ et $u(S_2)$ that can be simply obtained from

$$\frac{u^2(S_1)}{S_1^2} = \frac{u^2(a)}{a^2} + \frac{u^2(c)}{c^2} \qquad \text{and} \qquad \frac{u^2(S_2)}{S_2^2} = \frac{u^2(b)}{b^2} + \frac{u^2(c)}{c^2}. \tag{2.45}$$

Finally,

$$r(S_1, S_2) = \frac{u^2(c)}{c^2 \cdot \sqrt{\frac{u^2(a)}{a^2} + \frac{u^2(c)}{c^2}} \cdot \sqrt{\frac{u^2(b)}{b^2} + \frac{u^2(c)}{c^2}}}. \tag{2.46}$$

Assuming that

$$\frac{u(a)}{a} = \frac{u(b)}{b} = \frac{u(c)}{c} = 1\%, \tag{2.47}$$

one gets a very simple result, $r(S_1, S_2) = \frac{1}{2}$.

But if the uncertainty in c dominates, the correlation coefficient is larger. Let us assume, for example, that

$$\frac{u(a)}{a} = \frac{u(b)}{b} = 1\% \qquad \text{and} \qquad \frac{(c)}{c} = 5\%, \tag{2.48}$$

then $r(S_1, S_2) = \frac{25}{26} \simeq 1$.

On the contrary, with

$$\frac{u(a)}{a} = \frac{u(b)}{b} = 1\% \qquad \text{and} \qquad \frac{u(c)}{c} = 0.2\%, \tag{2.49}$$

$r(S_1, S_2) = \frac{0.04}{1.04}$, that is very small.

Example 2: mass of a sum

This example is adapted from the French standard FD X 07-021.

Two masses m_1 and m_2 of 50 g each are used have a mass m_a of 100 g. These two masses were compared to a standard m_s of 50 g. What is the uncertainty in m_a?

Here, one simply has $m_a = m_1 + m_2$, but m_1 and m_2 are correlated because they are both compared to the same standard. Therefore,

$$u^2(m_a) = u^2(m_1) + u^2(m_2) + 2u(m_1, m_2). \tag{2.50}$$

To evaluate each term of this equation, one has to write the measurement process in mathematical language:

$$m_1 = m_s + x_1 \qquad \text{and} \qquad m_2 = m_s + x_2. \tag{2.51}$$

where x_1 and x_2 are the errors due to each weighing. Then, it is just an application of the formulas that we have seen so far. One gets

$$u^2(m_1) = u^2(m_s)+u^2(x_1), \qquad u^2(m_2) = u^2(m_s)+u^2(x_2) \qquad \text{and} \qquad u(m_1, m_2) = \frac{\partial m_1}{\partial m_s}\frac{\partial m_2}{\partial m_s}u^2(m_s) = u^2(m_s). \tag{2.52}$$

Finally,

$$u^2(m_a) = u^2(x_1) + u^2(x_2) + 4u^2(m_s). \tag{2.53}$$

Note the factor 4 in front of $u^2(m_s)$ due to the correlation.

We could have proceeded in another way by directly expressing m_a with the measured quantities,

$$m_a = m_s + x_1 + m_s + x_2, \tag{2.54}$$

to immediately get

$$u^2(m_a) = u^2(x_1) + u^2(x_2) + 4u^2(m_s). \tag{2.55}$$

We get the same result.

It is not alway possible to directly express the output quantity as a function of the measured quantities like we did in this very simple example, and one has to proceed step by step with correlations that might propagate.

Example 3: Covariance between input and output quantities

To prepare a solution with a given concentration c, a technician desolves $m \pm u(m)$ moles of a substance into a volume $V \pm u(V)$ of water. The concentration is then $c = m/V$ and the uncertainty is easily determined from:

$$\frac{u^2(c)}{c^2} = \frac{u^2(m)}{m^2} + \frac{u^2(V)}{V^2}. \tag{2.56}$$

Someone else receives the solution with only indications the concentration $c \pm u(c)$ and the volume $V \pm u(V)$ and wants to determine the number of moles, $M = cV$ and its uncertainty. Naively, that person might write that

$$\frac{u^2(M)}{M^2} = \frac{u^2(c)}{c^2} + \frac{u^2(V)}{V^2}, \tag{2.57}$$

ignoring the correlation between c and V. If a new determination of the volume was done, this is correct. But if the volume is the one indicated by the technician, this is not correct. Injecting equation (2.56) into equation (2.57), one gets

$$\frac{u^2(M)}{M^2} = \frac{u^2(m)}{m^2} + 2\frac{u^2(V)}{V^2}, \tag{2.58}$$

that counts twice the uncertainty in the volume.

As a matter of fact, after preparation of the solution, the output quantity c is correlated to the input quantity V. The corresponding covariance is

$$u(c, V) = \frac{\partial c}{\partial V}\frac{\partial V}{\partial V}u^2(V) = \frac{\partial c}{\partial V}u^2(V) = -\frac{m}{V^2}u^2(V). \tag{2.59}$$

Then, equation (2.57) should be replaced by

$$\frac{u^2(M)}{M^2} = \frac{u^2(c)}{c^2} + \frac{u^2(V)}{V^2} + 2\frac{u(c,V)}{cV} \tag{2.60}$$

$$= \frac{u^2(c)}{c^2} + \frac{u^2(V)}{V^2} - 2\frac{u^2(V)}{V^2} \tag{2.61}$$

$$= \frac{u^2(c)}{c^2} - \frac{u^2(V)}{V^2}. \tag{2.62}$$

Injecting equation (2.56) into this result, the user recovers the original uncertainty in the mass: $u(M) = u(m)$.

As a conclusion, be careful when evaluating uncertainties. There might be unsuspected covariances that should be taken into account. In such a situation, the technician of the solution should have indicated $u(c, V)$.

Example 4: Mean of correlated data

We shall consider the case of repeated readings of the same quantity with the same instrument. Thus the values of the sequence $(x_1, \ldots, x_n)$ are correlated. Let us suppose that these values have a common mutual correlation coefficient r, such as $u(x_i, x_j) = ru(x_i)u(x_j) \, \forall i \neq j$.

We now calculate the mean,

$$\bar{x} = \frac{x_1 + \ldots + x_n}{n}.$$

As $\partial \bar{x}/\partial x_i = 1/n$, the propagation formula gives

$$u^2(\bar{x}) = \frac{1}{n^2} \sum_{i=1}^{n} u^2(x_i) + 2\frac{1}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} u(x_i, x_j).$$

With all the $u(x_1) = \ldots = u(x_n) = u(x)$, we also have $u(x_i, x_j) = ru^2(x)$ if $i \neq j$ and then

$$u^2(\bar{x}) = \frac{1}{n^2} n u^2(x) + \frac{1}{n^2}(n^2 - n) r u^2(x) = \left[\frac{1}{n} + r\frac{n-1}{n}\right] u^2(x). \tag{2.63}$$

Let us now consider two extreme cases. If $r = 0$, one recovers that $u^2(\bar{x}) = u^2(x)/n$. If $r = 1$, one gets that $u^2(\bar{x}) = u^2(x)$. Thus, in the latter case, the uncertainty in the mean value remains the same as the uncertainty in x. There is no need to increase the number of repetitions.

2.4 Level of confidence and degrees of freedom

2.4.1 Problem

As already mentioned in the introduction, standard deviation is sometimes not enough and one also needs the probability density function to interpret the results. It might be requested for example that 95% or 99% of the values are included in a given interval. In simpler cases it is sufficient to multiply the standard deviation by a **coverage factor** [facteur d'élargissement] to get an **expanded uncertainty** [incertitude élargie] corresponding to a **specified level of confidence** [niveau de confiance spécifié].

When the number of observations is large enough to have the probability density function close to a Gaussian distribution, which is an ideal case, the problem is represented in figure 2.3. Table 2.2 gives the correspondence between the level of confidence and the coverage factor for a Gaussian distribution.

Level of confidence p (percentage)	Coverage factor k_p
68.27	1
90	1.645
95	1.960
95.45	2
99	2.576
99.73	3

Table 2.2: *Value of the coverage factor k_p that gives the interval corresponding to a given level of confidence p for a Gaussian distribution.*

But how to do when we are not sure that the distribution is Gaussian or when there is a large uncertainty in the standard deviation itself? We already saw that when there are 30 values of repeated observations, a relative uncertainty of 13% remains in the estimator of the standard deviation. Even if we sure that the distribution is Gaussian, we cannot make the difference between 94, 95 or 96%. The central limit theorem shows that the probability density functions tends to a normal distribution for a large number of degrees of freedom (generally more than 20), but such conditions are not always satisfied and one has to use other probability density functions.

2.4.2 Student table or law

The Student law gives the relationship between the coverage factor associated with a specified level of confidence for type A evaluation of the uncertainty when the number of observations is to low to do a Gaussian approximation. As the Student law itself is very difficult to use, one rather uses the Student table that directly gives numerical values. Student [37] is the pseudonym of William Sealy Gosset, who worked at the Guinness Brewery in Dublin.

Note that Gosset published 14 papers in Biometrika (and 7 others elsewhere) over a 25-year period. All but one appeared under the pseudonym Student. The true identity was likely first revealed in the Journal of the American Statistical Association by Harold Hotelling in 1930: *"This concern years ago adopted a rule forbidding its chemists to publish their findings. Gosset pleaded that his mathematical and philosophical conclusions were of no possible practical use to competing brewers, and finally was allowed to publish them, but under a pseudonym, to avoid difficulties with the rest of the staff."*

Be z a random variable with a mathematical expectation μ_z and a standard deviation σ that satisfy a Gaussian distribution. If $\bar{z}$ is the mean value of n independent observations z_k and $s(\bar{z})$ the standard deviation, then the t variable

$$t = \frac{\bar{z} - \mu_z}{s(\bar{z})}, \tag{2.64}$$

follows Student law:

$$p(t,\nu) = \frac{1}{\sqrt{\pi\nu}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{t^2}{\nu}\right]^{-(\nu+1)/2}, \tag{2.65}$$

where Γ is the Gamma function and $\nu = n - 1$ the number of degrees of freedom.

Practically, one uses a table called Student table that gives the coverage factor $k_p(\nu)$ for a specified level of confidence p as a function of the number of degrees of freedom ν,

$$Pr[\bar{z} - k_p(\nu)s(\bar{z}) \leq \mu_z \leq \bar{z} + k_p(\nu)s(\bar{z})] = p. \tag{2.66}$$

A selection of values of the Student table is given in Table 2.3. A more extended table is available in the GUM.

ν	$p = 68.27\%$	$p = 95\%$	$p = 99\%$	$p = 99.73\%$
1	1.84	12.71	63.66	235.80
2	1.32	4.30	9.92	19.21
3	1.20	3.18	5.84	9.22
4	1.14	2.78	4.60	6.62
5	1.11	2.57	4.03	5.51
10	1.05	2.23	3.17	3.96
20	1.03	2.09	2.85	3.42
50	1.01	2.01	2.68	3.16
100	1.005	1.984	2.626	3.077
∞	1.000	1.960	2.576	3.000

Table 2.3: Coverage factor $k_p(\nu)$ as a function of the number of degrees of freedom ν for a specified level of confidence p.

Although Student distribution and table are only valid for type-A evaluation of the uncertainty, the GUM extends their application to all kinds of uncertainties. Then, the main difficulty is then to estimate the number of degrees of freedom.

2.4.3 Number of degrees of freedom

Type A

The **number of degrees of freedom** [nombre de degrés de liberté] is equal to $n-1$ for a single quantity estimated by the mean value of n independent observations. If one determines both the slope and the intercept by the least square method, $\nu = n-2$. If one determines m parameters from n values by the least square method, $\nu = n-m$.

Type B

For a type B evaluation of the uncertainty, the number of degrees of freedom is given by

$$\nu_i \simeq \frac{1}{2}\left[\frac{\Delta u(x_i)}{u(x_i)}\right]^{-2}, \tag{2.67}$$

where $\Delta u(x_i)/u(x_i)$ is the relative uncertainty in $u(x_i)$. It is estimated by expert judgment and is quite subjective. For example, if a relative uncertainty is estimated to be reliable at 10%, this gives an estimated number of degrees of freedom equal to $\nu_i = 50$. If the probability is uniform in a given interval (rectangular law), one generally considers that it includes all the values and it is realistic to choose $\nu \to \infty$.

Propagation

How to do when the quantity of interest is not directly measured but obtained from a calculation with N input quantities? For the uncertainty, one uses the propagation formula, eq. (2.19) and for the number of degrees of freedom, the GUM recommends to use Welch-Satterthwaite formula [38, 40] to estimate the effective number of degrees of freedom ν_{eff},

$$\frac{u^4(y)}{\nu_{eff}} = \sum_{i=1}^{N} \frac{\left[\frac{\partial f}{\partial x_i}u(x_i)\right]^4}{\nu_i}. \tag{2.68}$$

Be careful, this law is written in another way in the GUM. Practically, one rounds ν_{eff} to the closest lower integer.

Then, ν_{eff} is used to determine the coverage factor with Student table for both type A and B uncertainties although Welch-Satterthwaite formula was only derived for type A uncertainties. Some scientific articles shows cases where it cannot be used.

Example reproduced from the GUM: Let us assume that $Y = bX_1X_2X_3$ and that input quantities, x_1, x_2, x_3, are respectively estimated with $n_1 = 10$, $n_2 = 5$ and $n_3 = 15$ independent observations. Relative standard uncertainties are $s(x_1)/x_1 = 0,25\%$, $s(x_2)/x_2 = 0,57\%$ and $s(x_3)/x_3 = 0,82\%$. Relative uncertainty on y is straightforward,

$$\left[\frac{u_c(y)}{y}\right]^2 = \left[\frac{s(x_1)}{x_1}\right]^2 + \left[\frac{s(x_2)}{x_2}\right]^2 + \left[\frac{s(x_3)}{x_3}\right]^2 = (1,03\%)^2. \tag{2.69}$$

Similarly, one can also directly use relative uncertainties to evaluate ν_{eff} in this particular example, but it is not always possible. This single example can be misleading.

$$\nu_{eff} = \frac{(u_c(y)/y)^4}{\frac{(s(x_1)/x_1)^4}{\nu_1} + \frac{(s(x_2)/x_2)^4}{\nu_2} + \frac{(s(x_3)/x_3)^4}{\nu_3}} = \frac{1.03^4}{\frac{0.25^4}{10-1} + \frac{0.57^4}{5-1} + \frac{0.82^4}{15-1}} = 19.0. \tag{2.70}$$

In Student table, $k_{95\%}(19) = 2.09$. Expanded relative uncertainty for a level of confidence of 95% is $U_{95} = 2.09 \times 1.03\% = 2.2\%$.

2.4.4 Important notice

The conditions for the output probability density function being a Student's t distribution are quite strong and not likely to be met in many cases. This latter part corresponds to the weakest part of the GUM [49].

One can always expand the uncertainty with a coverage factor but the link to a specified degree of confidence is not always correct. In specific cases, remedies can rely on a Bayesian approach or on a Monte-Carlo simulation, as we shall see later.

2.5 Recommandations

This is the original recommendation on the expression of experimental uncertainties:

Recommendation INC-1 (1980)
Expression of experimental uncertainties

1. The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated.

   Type A. Those which are evaluated by statistical methods

   Type B. Those which are evaluated by other means

   There is not always a simple correspondence between the classification into categories A or B and the previously used classification into "random" and "systematic" uncertainties. The term "systematic uncertainty" can be misleading and should be avoided.

   Any detailed report of uncertainty should consist of a complete list of the components, specifying for each the method used to obtain its numerical value.

2. The components in category A are characterised by the estimated variances s_i^2 (or the estimated "standard deviations" s_i) and the number of degrees of freedom ν_i. Where appropriate the covariances should be given.

3. The components in category B should be characterised by quantities u_j^2, which may be considered approximations to the corresponding variances, the existence of which is assumed. The quantities u_j^2 may be treated like variances and the quantities u_j like standard deviations. Where appropriate, the covariances should be treated in a similar way.

4. The combined uncertainty should be characterised by the numerical value obtained by applying the usual method for the combination of variances. The combined uncertainty and its components should be expressed in the form of "standard deviations."

5. If for particular applications, it is necessary to multiply the combined uncertainty by an overall uncertainty, the multiplying factor must always be stated.

2.6 Propagation of distributions using a Monte Carlo method

This section is based on the Supplement 1 to the GUM: Propagation of distributions using a Monte Carlo method [7].

The propagation formula, equation (2.19), is based on a linearisation of the model using a Taylor expansion. If relative uncertainties are not small or if the relationship between one or several input quantities and the output quantity is highly non-linear, such an expansion is not correct anymore. Moreover, the coverage factor was estimated assuming a Gaussian or a Student distribution. If the dominating input quantity has a rectangular distribution, tables for the coverage factor are not correct anymore. For such cases, a numerical approach based on a Monte Carlo method is a practical alternative to the GUM uncertainty framework. A numerical approach is also necessary when the input quantities and the output quantity are linked by a numerical code.

The Supplement 1 to the GUM [7] provides a general numerical approach, consistent with the broad principles of the GUM, for carrying out the calculations required as part of an evaluation of measurement uncertainty. It also also provides guidance in situations where the conditions for the GUM uncertainty framework are not fulfilled.

The Supplement explains: *"In practice, only for simple cases can the propagation of distributions be implemented without making approximations. The GUM uncertainty framework implements one approximate method, and MCM another. For a small but important subset of problems, the GUM uncertainty framework is exact. MCM is never exact, but is more valid than the GUM uncertainty framework for a large class of problems."*

It is worth noticing that beyond measurements, theoretical predictions require a careful assessment of the uncertainties although it is frequent that the numerical results are presented without uncertainty estimates. In an editorial available in the appendix, the *Physical Review A* called for an uncertainty analysis for all theoretical predictions to a measurable quantity. Actually, the theoretical uncertainties can arise from several sources:

- uncertainties in the input parameters;
- numerical approximations;
- physical effects not included in the calculation from the beginning.

The latter is rather an error. Only the first source related to uncertainties due to parameters and its propagation is addressed by the GUM.

2.6.1 Method and mathematical procedure

Method

As before, the goal is to determine the uncertainty in an output quantity Y that depends on N input quantities, X_1, X_2, X_3..., but the method is based on PDF rather than standard deviations or variances. In practice, a value is randomly drawn for each input quantity and the output quantity is calculated. The procedure is repeated a large number of times, typically one million times. Then, a statistic treatment is performed with the set of results of the output quantity in order to estimate the mean value, the standard deviation, the confidence interval and the PDF. This is similar to repeated experiments, but generally with a larger statistics. See Fig. 2.5 for a schematic representation.

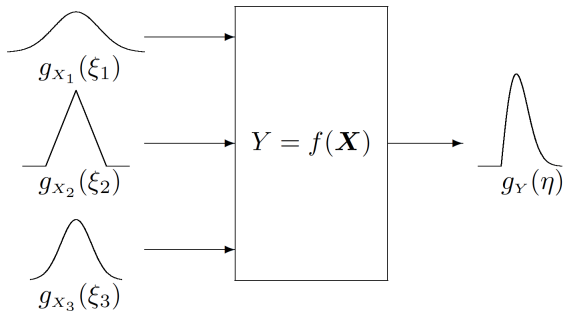


Figure 2.5: Illustration of the propagation of distributions for 3 independent input quantities.

The main difficulty lies in the random sampling of input quantities according to their PDF for which the supplement to the GUM gives recommendations.

Conventions and notation

Let us assume that a single output quantity Y can be expressed as a function of N input quantities

$$[X] = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix} = [X_1, \ldots, X_N]^T. \tag{2.71}$$

Here Y is a scalar and $[X]$ is a vertical vector. In these lecture notes, vectors are written with square brackets. Bold characters are used in the GUM.

The functional relationship between these quantities is expressed as

$$Y = f([X]). \tag{2.72}$$

In the supplement, the symbol f is reserved for the model. The relationship could also have an implicit form: $h(Y,[X])=0$.

Each X_i is regarded as a random variable [variable aléatoire] with possible values ξ_i and expectation x_i. Consequently, Y is also a random variable with possible values η and expectation y.

For a single input quantity X, the probability density function, PDF, [fonction densité de probabilité] is written $g_X(\xi)$. The probability that $X \in [\xi, \xi+d\xi[$ is $g_X(\xi)\,d\xi$. Therefore, expectation value and variance are given by

$$E(X) = \int_{-\infty}^{+\infty} \xi\, g_X(\xi)\, d\xi \tag{2.73}$$

$$V(X) = \int_{-\infty}^{+\infty} (\xi - E(X))^2\, g_X(\xi)\, d\xi \tag{2.74}$$

respectively.

With two input quantities, X_1 et X_2, there is a single PDF for $[X]$ denoted $g_{[X]}([\xi])$. With a Gaussian distribution, it reads

$$g_{[X]}([\xi]) = \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2\sigma_2^2-\sigma_{12}^2}} \exp\left(-\frac{\sigma_2^2(\xi_1-x_1)^2+\sigma_1^2(\xi_2-x_2)^2-2\sigma_{12}(\xi_1-x_1)(\xi_2-x_2)}{2(\sigma_1^2\sigma_2^2-\sigma_{12}^2)}\right) \tag{2.75}$$

$$= \frac{1}{2\pi}\frac{1}{\sqrt{\det[\sigma^{\mathbf{2}}]}} \exp\left(-\frac{1}{2}[\xi-x]^T\cdot[\sigma^2]^{-1}\cdot[\xi-x]\right). \tag{2.76}$$

If the two input quantities are independent, then $\sigma_{12}=0$ and thus, we simply have

$$g_{[X]}([\xi]) = g_{X_1}(\xi_1)\cdot g_{X_2}(\xi_2). \tag{2.77}$$

If the two input quantities are correlated, we cannot define a PDF for each of them. The expectation reads

$$x_i = E(X_i) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \xi_i\, g_{[X]}([\xi])\, d\xi_1\, d\xi_2, \tag{2.78}$$

and for the variance we have a 2×2 matrix,

$$[V([X])] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [X-E([X])]\cdot[X-E([X])]^T\, g_{[X]}([\xi])\, d\xi_1\, d\xi_2. \tag{2.79}$$

Diagonal terms correspond to the variances and off-diagonal ones to covariances:

$$Cov(X_1,X_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\xi_1-E(X_1))(\xi_2-E(X_2))\, g_{[X]}([\xi])\, d\xi_1\, d\xi_2. \tag{2.80}$$

This can be easily generalised to N input quantities.

Main stages of uncertainty evaluation

The main stages of uncertainty evaluation constitute formulation, propagation, and summarising.

The formulation consist in

- defining the output quantity Y,
- determining the input quantities $[X]=[X_1,...,X_N]^T$ upon which Y depends,
- developing a model relating Y and $[X]$, and
- assigning PDF's to the X_i and a joint PDF to those X_i that are not independent.

Guidance on the assignment of PDFs is given in the next subsection.

The propagation is based on a repeated random sampling from the PDFs for the X_i to obtain numerical results for the outpu quantity Y. For that, one randomly draws a set of values ξ_i for each of the input quantities, calculate a value η of the output quantity and repeat the process $M-1$ times. One gets M values for Y from which one can build the PDF for Y. This is illustrated in Fig. 2.5 for three independent input quantities. M is typically of the order of 10^6.

Summarising consist in using the PDF for Y to obtain

1. the expectation of Y, taken as an estimate y of the quantity,
2. the standard deviation of Y, taken as the standard uncertainty $u(y)$ associated with y, and
3. a coverage interval containing Y with a specified probability (the coverage probability).

The numerical methods to get these values are detailed in the supplement1 to the GUM. They make use of the distribution function

$$G_Y(\eta) = \int_{-\infty}^{\eta} g_Y(z)\, dz \tag{2.81}$$

that encodes all the information known about Y. Thus, any property of Y such as expectation, variance and coverage intervals can be approximated using G. The quality of these calculated results improves as the number of times the PDFs are sampled increases. Note that the method is a bit tricky when the PDF for the output quantity is asymmetric. See the Supplement 1 for details.

Finally, y, $u(y)$ and the endpoints of a $100p\%$ coverage interval for Y be reported would typically be reported. Example: $y = 1.024$ V; $u(y) = 0.028$ V; shortest 95% coverage interval $= [0,983,\ 1,088]$.

Implementation

Practically, MCM can be stated as a step-by-step procedure:

1. select the number M of Monte Carlo trials to be made;
2. generate M vectors, by sampling from the assigned PDFs, as realisations of the set of N input quantities X_i;
3. for each such vector, form the corresponding model value of Y, yielding M model values y_r $(r = 1, \ldots, M)$;
4. sort these M model values into strictly increasing order to provide G;
5. use G to form an estimate y of Y and the standard uncertainty $u(y)$ associated with y;
6. use G to form an appropriate coverage interval for Y, for a stipulated coverage probability p.

2.6.2 Probability density functions for the input quantities

One usually does not know the PDF for a given input quantity. The assignment of PDFs to the input quantities X_i depends on the available information. Such an assignment makes use of the principle of maximum entropy that will be explained in chapter 3.

Here are two examples. Many others are given in the Supplement 1 to the GUM [7].

Rectangular distributions

If the only available information regarding a quantity X is a lower limit a and an upper limit b with $a < b$, then, according to the principle of maximum entropy, a rectangular distribution $R(a, b)$ over the interval $[a, b]$ would be assigned to X.

The PDF for X is,

$$g_X(\xi) = \begin{cases} 1/(b-a), & \text{si } \xi \in [a,b] \\ 0, & \text{otherwise} \end{cases} . \tag{2.82}$$

Then, X has expectation and variance

$$E(X) = \frac{a+b}{2} \qquad \text{et} \qquad V(X) = \frac{(b-a)^2}{12} \tag{2.83}$$

as we already know.

To sample from $R(a,b)$, make a draw r from the standard rectangular distribution $R(0,1)$, and form

$$\xi = a + (b-a)r. \tag{2.84}$$

There are prescriptions about the random number generator: the enhanced Wichmann-Hill generator for producing pseudo-random numbers is recommended.

Simple Gaussian distributions

If a best estimate x and associated standard uncertainty u(x) are the only information available regarding a quantity X, then, according to the principle of maximum entropy, a Gaussian probability distribution $N(x, u^2(x))$ would be assigned to X. This will be demonstrated in chapter 3.

The PDF for X is

$$g_X(\xi) = \frac{1}{\sqrt{2\pi}u(x)} \exp\left(-\frac{(\xi-x)^2}{2u^2(x)}\right). \tag{2.85}$$

As expected, expectation and variance are $E(X) = x$ et $V(X) = u^2(x)$.

To sample from $N(x, u^2(x))$, make a draw z from the standard Gaussian distribution $N(0,1)$, and form

$$\xi = x + u(x)z. \tag{2.86}$$

The Box-Muller method is recommended to draw z.

Multivariate Gaussian distributions

For an N-dimensional quantity $[X] = (X_1, \ldots, X_N)^T$ for which the only information available is a best estimate $[x] = (x_1, \ldots, x_N)^T$ and the associated (strictly) positive definite uncertainty matrix

$$[U] = \begin{bmatrix} u^2(x_1) & u(x_1,x_2) & \cdots & u(x_1,x_N) \\ u(x_2,x_1) & u^2(x_2) & \cdots & u(x_2,x_N) \\ \vdots & \vdots & \ddots & \vdots \\ u(x_N,x_1) & u(x_N,x_2) & \cdots & u^2(x_N) \end{bmatrix}, \tag{2.87}$$

a multivariate Gaussian distribution would be assigned to X:

$$g_{[X]}([\xi]) = \frac{1}{\sqrt{(2\pi)^N \det[U]}} \exp\left(-\frac{1}{2}[\xi - x]^T \cdot [U]^{-1} \cdot [\xi - x]\right). \tag{2.88}$$

Obviously, expectation and covariance matrix are $E([X] = [x]$ and $V(X) = [U]$.

To sample from $N(x, [U]_x)$, one should use the Cholesky decomposition which consist in determining a upper triangular matrix such as $[U] = [R]^T \cdot [R]$. This is quite easy to compute. Then, one makes N draws z_i, $i = 1, \ldots, N$, independently from the standard Gaussian distribution $N(0,1)$, and form

$$[\xi] = [x] + [R]^T \cdot [z], \tag{2.89}$$

where $[z] = (z_1, \ldots, z_N)^T$.

The Cholesky decomposition can easily be understood with an simple example. Let's consider a symmetric matrix to be decomposed into

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 13 & 6 \\ -3 & 6 & 50 \end{pmatrix} = \begin{pmatrix} r_{11} & 0 & 0 \\ r_{12} & r_{22} & 0 \\ r_{13} & r_{23} & r_{33} \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix}. \tag{2.90}$$

Calculating the product of the matrixes of left hand side of the previous equation, on gets

$$\begin{array}{rcll} 1 &=& r_{11}^2 & \Rightarrow \quad r_{11} = 1 \\ 2 &=& r_{11}r_{12} & \Rightarrow \quad r_{12} = 2 \\ -3 &=& r_{11}r_{13} & \Rightarrow \quad r_{13} = -3 \\ 13 &=& r_{12}^2 + r_{22}^2 & \Rightarrow \quad r_{22} = 3 \\ 6 &=& r_{12}r_{13} + r_{22}r_{23} & \Rightarrow \quad r_{23} = 4 \\ 50 &=& r_{13}^2 + r_{23}^2 + r_{33}^2 & \Rightarrow \quad r_{33} = 5 \end{array} \tag{2.91}$$

that can be easily solved step by step as indicated. Finally,

$$[R] = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 3 & 4 \\ 0 & 0 & 5 \end{pmatrix}. \tag{2.92}$$

For larger matrix, this is very easy to compute. There are also subroutines available in libraries.

2.6.3 Number of Monte Carlo trials

The value of M, the number of Monte Carlo trials, directly influences the quality of the numerical results provided by MCM. How to select it?

A value of $M = 10^6$ can often be expected to deliver a 95% coverage interval for the output quantity such that this length is correct to one or two significant decimal digits. In general, the choice of a value of M that is large compared with $1/(1-p)$, e.g. M at least 10^4 times greater than $1/(1-p)$, should be made.

Because there is no guarantee that this or any specific pre-assigned number will suffice, a procedure that selects M adaptively, i.e. as the trials progress, can be used. Some guidance in this regard is available in the supplement 1 to the GUM [7].

2.7 Extension to any number of output quantities

This section is based on the Supplement 2 to the GUM: Extension to any number of output quantities [8].

The "Guide to the expression of uncertainty in measurement" (GUM) is mainly concerned with univariate measurement models, namely models having a single scalar output quantity. However, models with more than one output quantity arise across metrology. Such quantities are generally mutually correlated because they depend on common input quantities. The Supplement 2 provides a generalisation of the GUM uncertainty framework in order to provide estimates of the output quantities, the standard uncertainties associated with the estimates, and covariances associated with pairs of estimates.

2.7.1 Conventions and notation

As before, the N input quantities are denoted X_i and collectively presented by a vertical vector $[X] = [X_1, \ldots, X_N]^T$. There are characterised by an assigned PDF, $g_{[X]}(\xi_1, \ldots, \xi_N)$. Likewise the m output quantities are generically denoted by Y_k and collectively by a vertical vector $[Y] = [Y_1, \ldots, Y_m]^T$.

In the simplest case, there are m functional relationships between the output and input quantities:

$$Y_1 = f_1(X_1, \ldots, X_N), \qquad \ldots, \qquad Y_m = f_m(X_1, \ldots, X_N), \tag{2.93}$$

which can be written in a more compact form, $[Y] = [f]([X])$.

The relationship can take an implicite form with a system of m equations,

$$h_1(Y_1, \ldots, Y_m, X_1, \ldots, X_N) = 0, \qquad \ldots, \qquad h_m(Y_1, \ldots, Y_m, X_1, \ldots, X_N) = 0, \tag{2.94}$$

that can be written in a more compact form, $[h]([Y], [X]) = 0$. In the following, we shall only consider the simplest case.

An estimate of $[X]$ is denoted $[x] = [x_1, \ldots, x_N]^T$ and the covariance matrix associated with $[x]$ is denoted by $[U]_{[x]}$, a matrix of dimension $N \times N$.

Likewise, an estimate of $[Y]$ is denoted by $[y] = [y_1, \ldots, y_N]^T$ and the covariance matrix associated with $[y]$ is denoted by $[U]_{[y]}$, a matrix of dimension $m \times m$.

$[U]_{[y]}$ is the counterpart, for m output quantities, of the variance $u^2(y)$ associated with y in the context of the univariate measurement models of the GUM and GUM Supplement 1. When estimates of the output quantities are to be used individually, each of these quantities may be considered as the output quantity in the corresponding univariate (scalar) measurement model. When the output quantities are to be considered together, for instance used in a subsequent calculation, any correlations associated with pairs of estimates of the output quantities need to be taken into account.

What we want to determine is the PDF for $[Y]$, denoted by $g_{[Y]}([\eta])$. Within the GUM uncertainty framework based on analytical formulas, $[y]$ and $[U]_{[y]}$ are sufficient.

2.7.2 GUM uncertainty framework

Framework

Given an estimate $[x]$ of $[X]$, an estimate of $[Y]$ is

$$[y] = [f]([x]). \tag{2.95}$$

The covariance matrix of dimension $m \times m$ associated with $[y]$ is,

$$[U]_{[y]} = \begin{bmatrix} u^2(y_1) & u(y_1, y_2) & \cdots & u(y_1, y_m) \\ u(y_2, y_1) & u^2(y_2) & \cdots & u(y_2, y_m) \\ \vdots & \vdots & \ddots & \vdots \\ u(y_m, y_1) & u(y_m, y_2) & \cdots & u^2(y_m) \end{bmatrix}, \tag{2.96}$$

is given by

$$\boxed{[U]_{[y]} = [C]_{[x]} \cdot [U]_{[x]} \cdot [C]_{[x]}^T,} \tag{2.97}$$

where $[C]_{[x]}$ is the sensitivity matrix of dimension $m \times N$ given by evaluating

$$[C]_{[x]} = \begin{bmatrix} \frac{\partial f_1}{\partial X_1} & \frac{\partial f_1}{\partial X_2} & \cdots & \frac{\partial f_1}{\partial X_N} \\ \frac{\partial f_2}{\partial X_1} & \frac{\partial f_2}{\partial X_2} & \cdots & \frac{\partial f_2}{\partial X_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial X_1} & \frac{\partial f_m}{\partial X_2} & \cdots & \frac{\partial f_m}{\partial X_N} \end{bmatrix}, \tag{2.98}$$

at $[X] = [x]$. This is the new propagation formula when there are several output quantities.

The case of an implicit multivariate model specifies a relationship between output quantities $[Y]$ and input quantities $[X]$ with the function h_k is also treated in the supplement.

Example: two faces of a parallelepiped

Let us consider again a rectangular parallelepiped with edges of length a, b and c. We need to know two surfaces $S_1 = ac$ and $S_2 = bc$. As before, we shall assume that a, b and c are independent.

Input quantities are $[X] = [a, b, c]^T$ and output ones $[Y] = [S_1, S_2]^T$. As the three input quantities are independent, the associated covariance matrix is diagonal:

$$[U]_{[x]} = \begin{bmatrix} u^2(a) & 0 & 0 \\ 0 & u^2(b) & 0 \\ 0 & 0 & u^2(c) \end{bmatrix}. \tag{2.99}$$

Then, the sensitivity matrix of dimension 2×3 reads,

$$[C]_{[x]} = \begin{bmatrix} \frac{\partial S_1}{\partial a} & \frac{\partial S_1}{\partial b} & \frac{\partial S_1}{\partial c} \\ \frac{\partial S_2}{\partial a} & \frac{\partial S_2}{\partial b} & \frac{\partial S_2}{\partial c} \end{bmatrix} = \begin{bmatrix} c & 0 & a \\ 0 & c & b \end{bmatrix}, \tag{2.100}$$

to get the covariance matrix associated with $[y]$

$$\begin{aligned} [U]_{[y]} &= [C]_{[x]} \cdot [U]_{[x]} \cdot [C]_{[x]}^T = \begin{bmatrix} c & 0 & a \\ 0 & c & b \end{bmatrix} \cdot \begin{bmatrix} u^2(a) & 0 & 0 \\ 0 & u^2(b) & 0 \\ 0 & 0 & u^2(c) \end{bmatrix} \cdot \begin{bmatrix} c & 0 \\ 0 & c \\ a & b \end{bmatrix} \\ &= \begin{bmatrix} c^2u^2(a) + a^2u^2(c) & ab\,u^2(c) \\ ab\,u^2(c) & c^2u^2(b) + b^2u^2(c) \end{bmatrix}. \end{aligned} \tag{2.101}$$

Eventually, this leads to

$$u^2(S_1) = c^2u^2(a) + a^2u^2(c) \tag{2.102}$$

$$u^2(S_2) = c^2u^2(b) + b^2u^2(c) \tag{2.103}$$

$$u(S_1, S_2) = u(S_2, S_1) = ab\,u^2(c), \tag{2.104}$$

which is similar to what we found before.

When there are many quantities, the matrix formalism is more compact.

2.7.3 Coverage region for a vector output quantity

With a single output quantity, once the expectation and variance obtained, we could evaluate the confidence interval associated to a given level of confidence p. With several output quantities, given an estimate $[y]$ of the output quantity and its associated covariance matrix $[U]_{[y]}$, we do not deal with an interval anymore but a region in m-dimensional space that contains $[Y]$ related to the a coverage probability p. Thus, we want to determine a *coverage region* for a specied *coverage probability*.

If the output quantities are correlated, we cannot study them separately. If $[y]$ and $[U]_{[y]}$ constitute the only available information about $[Y]$, according to the principle of maximum entropy the Gaussian PDF is used to describe the state of knowledge of $[Y]$. Thus, we shall first consider a Gaussian PDF. For other PDFs, a numerical method such as a Monte Carlo method is helpful in obtaining approximate solutions that are acceptable for practical purposes. This goes beyond the ambition of this lecture. Details are available in the supplement 2 to the GUM [8].

We shall start with he determination of a coverage region for a bivariate quantity (two output quantities) before generalising to the more general multivariate case.

Bivariate case

Let us consider the case of two quantities Y_1 and Y_2. The available information constitutes estimates y_1 and y_2 of the coordinates, standard uncertainties $u(y_1)$ and $u(y_2)$ associated with the estimates, as well as the covariance $u(y_1, y_2)$. The associated Gaussian PDF simply reads

$$g_{[y]}([\eta]) = \frac{1}{2\pi} \frac{1}{\sqrt{\det[U_y]}} \exp\left(-\frac{1}{2}[\eta - y]^T \cdot [U_{[y]}]^{-1} \cdot [\eta - y]\right) \tag{2.105}$$

$$= \frac{1}{2\pi} \frac{1}{\sqrt{u_1^2u_2^2 - u_{12}^2}} \exp\left(-\frac{u_2^2(\eta_1 - y_1)^2 + u_1^2(\eta_2 - y_2)^2 - 2u_{12}(\eta_1 - y_1)(\eta_2 - y_2)}{2(u_1^2u_2^2 - u_{12}^2)}\right) \tag{2.106}$$

$$\propto \exp\left(-\frac{1}{2}\chi^2\right). \tag{2.107}$$

Isoprobability lines correspond to a fixed chi-squared, χ^2. There are ellipses centred at $[y]$.

Thus, the coverage region corresponding to a coverage probability p is the area in the elliptical region defined such that $\chi^2 = k_p^2$. With two quantities and the coverage probability $p = 0.95$, $k_p = 2.45$.

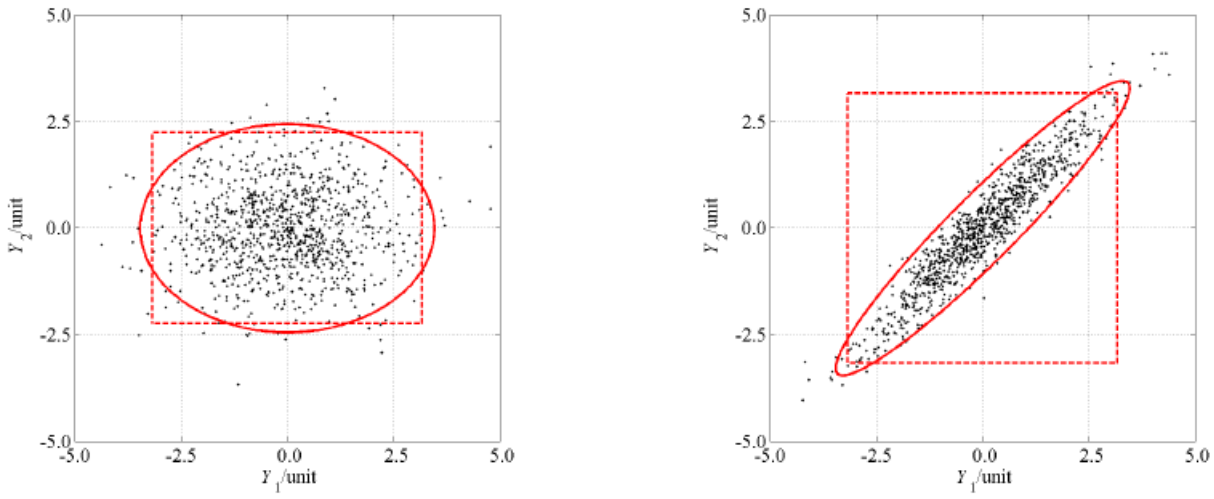


Figure 2.6: Elliptical and rectangular coverage regions for a bivariate quantity characterised by a Gaussian probability distribution and a coverage probability equal to 0.95. In the left panel the component quantities are mutually independent. In the right panel there are highly correlated. See text for details. Figure reproduced from the supplement 2 to the GUM [8].

Let us consider two examples with a Gaussian PDF. The first one assumes that

$$[y] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad [U]_{[y]} = \begin{bmatrix} 2,0 & 0 \\ 0 & 1,0 \end{bmatrix}, \tag{2.108}$$

which corresponds to two independent quantities. The second one assumes a strong correlation between the two quantities:

$$[y] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad [U]_{[y]} = \begin{bmatrix} 2,0 & 1,9 \\ 1,9 & 2,0 \end{bmatrix}. \tag{2.109}$$

The coverage region corresponding to a coverage probability $p = 0.95$ for these two examples is shown in Fig. 2.6. For both cases, the coverage region is centred on the expectation. At left, if the two uncertainties were equal, the coverage region would be a circle. At right, the strong correlation between the two quantities is obvious: if η_1 is far away from y_1, there is a high probability that η_2 is also far away from y_2.

What would we obtain if we consider the two quantities separately? The PDF for Y_1 is given by marginalisation:

$$g_{Y_1} = \int_{-\infty}^{\infty} g_{Y_1,Y_2}(\eta_1, \eta_2)\, d\eta_2. \tag{2.110}$$

irrespective of the state of knowledge of Y_2. The coverage factor k_q used for the coverage interval $y_1 \pm k_q u(y_1)$ is determined for coverage probability

$$q = 1 - (1 - p)/2 = (1 + p)/2. \tag{2.111}$$

For the coverage probability $p = 0.95$ for both quantities, $q = 0.975$ for each of them taken separately. When Y_1 and Y_2 are independent, q in the above expression can be replaced by $q = \sqrt{p}$. The coverage region is then rectangular as shown in Fig. 2.6.

When the two quantities Y_1 et Y_2 are independent (Figure 2.6, left panel), the 95% elliptical and rectangular coverage regions are close to each other. However, when the two quantities are highly correlated (Figure 2.6, right panel), elliptical and rectangular coverage regions differ a lot. As an illustration, 1 000 points representing random draws from this probability distribution are also shown. When the two quantities Y_1 et Y_2 are independent, 950 of the 1 000 points are contained within the elliptical coverage region, which has area

26.6 unit^2, and 953 within the rectangular region, which has area 28.4 unit^2. When the component quantities Y_1 and Y_2 are positively correlated, 957 of the 1 000 points are contained within the elliptical coverage region, which has area 11.8 unit^2, and 972 within the rectangular region, which has area 40.1 unit^2, indicating that the coverage probability for the rectangular region exceeds 0.95. The rectangular region, whose shape does not reflect the correlated component quantities and the distribution of the randomly drawn points, might be considered inappropriate as a coverage region for $[Y]$. A rectangle with sides parallel to the axes of the ellipse would have smaller area and might be considered more appropriate, but might be inconvenient since it would be expressed in terms of variables that would be artificial in terms of the application.

Multivariate case

For more than two quantities, coverage regions are no longer easily visualised, but can be constructed in an analogous manner to the bivariate case. The coverage region is a hyper-ellipsoid (or multivariate ellipse) in m dimensions.

Table 2.4 gives values for the coverage factor k_p for the coverage probability $p = 0.95$ and m jointly evaluated Gaussian quantities for a selection of values of m. If $m = 1$, k_p corresponds to the coverage factor of a simple Gaussian: $k_p = 1.96$.

m	k_p	m	k_p	m	k_p	m	k_p
1	1.96	6	3.55	11	4.44	20	5.60
2	2.45	7	3.75	12	4.59	25	6.14
3	2.80	8	3.94	13	4.73	30	6.62
4	3.08	9	4.11	14	4.87	40	7.47
5	3.33	10	4.28	15	5.00	50	8.82

Table 2.4: Coverage factors for hyper-ellipsoidal coverage regions corresponding to coverage probability $p = 0.95$.

2.7.4 Monte Carlo method

The method is similar to the method used with a single output quantity: making M repeated draws from the PDFs for the X_i (or joint PDF for $[X]$) and evaluating the vector output quantity in each case. The set of output quantity values are used to build up the joint probability distribution for $[Y]$.

Let $[y]_r$, for $r = 1, \ldots, M$, represent the output quantity values. The average and covariance matrix,

$$[\tilde{y}] = \frac{1}{M}\sum_{r=1}^{M}[y]_r \qquad \text{and} \qquad [U]_{[\tilde{y}]} = \frac{1}{M-1}\sum_{r=1}^{M}([y]_r - [\tilde{y}])([y]_r - [\tilde{y}])^T, \tag{2.112}$$

are taken, respectively, as an estimate $[y]$ of $[Y]$ and the covariance matrix $[U]_{[y]}$ associated with $[y]$.

The main difficulty arises from the evaluation of the coverage region. See details in the supplement 2 to the GUM [8]. One of the possible methods consists in subdividing the hyper-rectangular region that contains all output values into a mesh of small rectangles and counting output quantity values assigned to each rectangle. Then the rectangles are sorted in terms of decreasing probability to form the cumulative sum of probabilities for these listed rectangles, stopping when this sum is no smaller than p. The fineness of the mesh influences the approximate coverage region produced. The quality of the approximation generally improves with M. To achieve a sufficiently good approximation, particularly for a number of output quantities larger than about two or three, might necessitate a very large value of M, otherwise the region so obtained might be disjointed. This is illustrated in Fig. 2.7.

The supplement 2 propose an adaptive Monte Carlo procedure to determine the optimum value for M. See [8] for details.

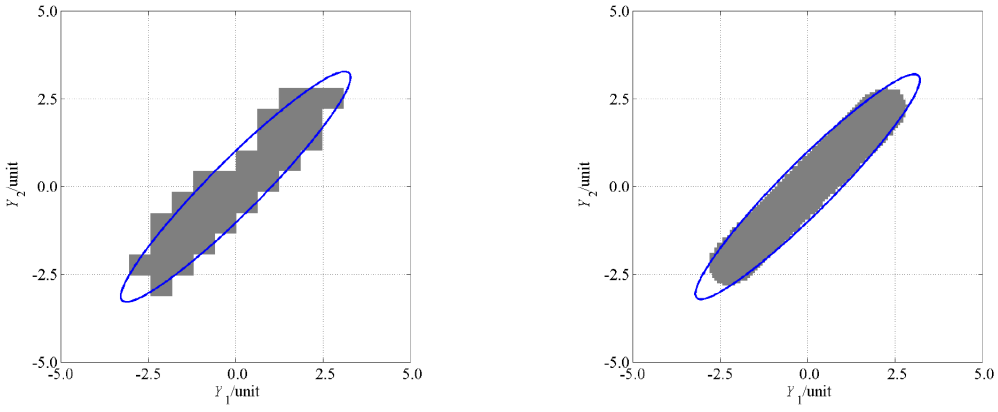


Figure 2.7: Elliptical coverage regions based on (2.109) and approximations to the smallest coverage region for a 10×10 mesh and 1 000 points drawn randomly, and (right) for a 100×100 mesh and 1 000 000 points. Figure reproduced from the supplement 2 to the GUM [8].

2.8 Conclusion

This chapter is just an introduction to the estimation of the uncertainty in measurements. Note that there are several supplements to the GUM already published by the BIPM. Especially the first supplements present the numerical methods of evaluation of the uncertainty when the relation between the input quantities and the output ones is not a formula but a numerical code. These two supplements are also useful when the uncertainty is large and the propagation formula, based on a first order expansion, cannot be used. Other supplements are announced.

As a conclusion, we quote section 3.4.8 of the GUM:

> *Although this* Guide *provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value.*

Chapter 3

Bayesian inference to evaluate uncertainties

"The fundamental problem of scientific progress, and a fundamental one in everyday life, is that of learning from experience."

Harold Jeffreys, in *Theory of Probability*

At the beginning of this lecture, we started from a repetition of observations in order to estimate both expectation and variance. The larger the number of observations the better the estimation. This is the so-called frequentist approach. Unfortunately, it is not always feasible. When you do a blood test for example, it is not possible to take large samples.

When we cannot repeat the process a large number of times, one has to take into account other considerations that are better established than the information given by the measurement. For example, a given quantity such as temperature, activity... is always positive. How to include such information in the data analysis? This is the aim of the Bayesian inference.

The type B evaluation is also an estimate based on the available information. In such a case, probability represents the likelihood of a value, not its frequency. As a matter of facts, the GUM is not consistent: it uses a frequentist approach for the type A evaluation and a so-called plausibilist approach for the type B. In the scientific literature and in some recent standards there is a tendency to treat all probabilities as likelihood. An experiment is then seen as a learning process which is treated by Bayes theorem. The goal of this chapter is to introduce the method. There are several books on the subject. H. Jeffreys was the first to reintroduced Bayesian inference [13] and E.T. Jaynes also contributed to its developments [14]. For a pedagogical introduction, I recommend the one of D.S. Sivia [15] that is very good.

This is a change of paradigm. We mentioned the probability density function of the values taken by a quantity, or an estimation of it. But actually, one should admit that we only deal with the information we have on the values taken by a given quantity or its likelihood. Therefore, we are going to use concepts from the theory of information, inference...

Actually, both definitions of probabilities has always existed as explained by Ian Hacking in his excellent book on the emergence of probability [16]. Pascal wager is one of the first known link between the two concepts: *"God is, or He is not." But to which side shall we incline? Reason can decide nothing here. There is an infinite chaos which separated us. A game is being played at the extremity of this infinite distance where heads or tails will turn up. What will you wager?*

I can also recommend two enlightening articles on the subject: *Measurement as Inference: Fundamental Ideas* by W. Tyler Estler [28] and *Where do we stand on maximum entropy?* by E.T. Jaynes [27]. The second one is available online. One of the first articles suggesting a Bayesian approach in the standardisation of the uncertainty estimation is from 1992 [45].

This chapter will start by some reminders on probability theory and Bayesian inferences before addressing the specific applications to uncertainty analysis.

3.1 Probabilities

"La théorie des probabilités n'est que le bon sens réduit au calcul[1]*."*
Pierre Simon, Marquis de Laplace

3.1.1 Quiz

Let us start by a quiz to understand the two definitions of probabilities.

Imagine that there are 7 black balls and 3 white balls in an urn. Colour is the only difference. If a ball is drawn from the urn, we all agree that it has 70% chances to be black and 30% chances to be white. Suppose that the ball is hidden without knowing the colour. A second ball is drawn. If we look at its colour, does it change the probability for the first ball? If you are hesitating, imagine that there were only two balls in the urn. Looking at the second one immediately gives you the colour of the first one.

The result of the first experiment is not affected by the second experiment. However, the information available on the result of the first experiment is affected by the second experiment: if the second ball is white, the probability that the first ball is also white is 2/9 and the probability that it is black, 7/9.

In this case probability really mean information about the colour of the first ball. It differs from the frequentist approach that would have consisted in repeating the first experiment.

3.1.2 Algebra

Let's first play cards: the probability of drawing an ace in a card game is $p_A = 4/52$. The probability of drawing a red card is $p_R = 1/2$. The probability of drawing a red ace is $2/52 = p_A \times p_R$.

More generally, one writes $p(A)$ the probability that A is true and $p(B)$ that B is true. One writes $A \cap B$ that A and B are both true. One has,

$$p(A \cap B) = p(A) \cdot p(B). \tag{3.1}$$

If one tosses a coin twice, the probability that both trials are head is $p(1).p(2) = 1/4$. Figure 3.1 shows a schematic representation of the condition AND that corresponds to the intersection in green.

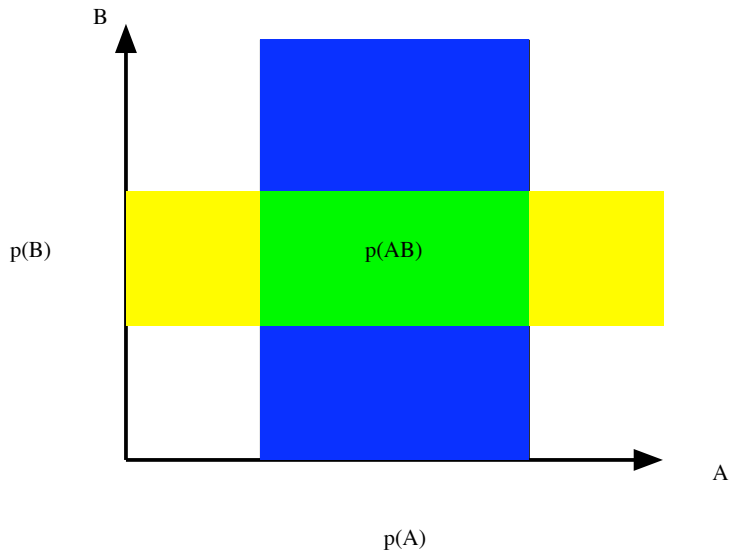


Figure 3.1: Representation of the probability that A AND B are true.

Similarly, one writes $A \cup B$ that A or B are true. The addition rule is

$$p(A \cup B) = p(A) + p(B) - p(A \cap B), \tag{3.2}$$

[1] *Probability theory is nothing but common sense reduced to calculation.*

because one should not count twice the intersection of the ensembles A and B (in green in Fig. 3.1). Therefore, the probability to toss at least one head with two trials is $3/4 = 1/2 + 1/2 - 1/4$.

These two rules were illustrated here with examples having a frequentist interpretation. But they are also correct with a plausibilist interpretation, like "what is the probability that there is life on Mars?". R.T. Cox demonstrated in a famous article [41] that these rules can be deduce from a simple axiomatic. See the enlightening article of W. Tyler Estler [28] and references therein.

But, **be careful**, relation (3.1) is only true when A and B are independent. Let's suppose now that A means "the drawn card is a heart" and B "the drawn card is red". Relation (3.1) is not correct anymore because hearts are always red. A and B are not independent. We need to introduce conditional probabilities to treat such cases. See next section.

A and B are said exclusive if $p(A \cap B) = 0$. It is the case if A means "red card" and B "black card". In such a case, equation (3.2) becomes $p(A \cup B) = p(A) + p(B)$.

Denoting $\bar{A}$, non-A, we have $p(A \cap \bar{A}) = 0$ and also $p(A) + p(\bar{A}) = 1$.

3.1.3 Conditional probability

Previous rules were only valid if A and B were two randomly independent events. To deal with problems where A and B can be correlated, we introduce conditional probabilities: $p(A|B)$ denotes that A is true when B is true. Here, B is a hypothesis. Be careful, $p(A|B)$ and $p(B|A)$ differ. The probability that its rains when it is cloudy is not equal to the probability that it is cloudy when it rains! How link these two probabilities?

With conditional probabilities, previous rules become

$$p(A \cap B|H) = p(A|H) \cdot p(B|AH) = p(B|H) \cdot p(A|BH) \tag{3.3}$$

$$p(A \cup B|H) = p(A|H) + p(B|H) - p(A \cap B|H). \tag{3.4}$$

For the sake of generality, these rules were written assuming an additional hypothesis H.

Let's play cards again to understand. Here H could be "the game is not biased". Let's calculate the probability to draw a black clover $p(B \cap C|H)$ by using eq. (3.3):

$$p(B \cap C|H) = p(B|H) \cdot p(C|BH) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \tag{3.5}$$

or

$$p(B \cap C|H) = p(C|H) \cdot p(B|CH) = \frac{1}{2} \cdot 1 = \frac{1}{4}. \tag{3.6}$$

It works!

Let us consider cases where there are several hypothesis B_i that are exclusive and exhaustive, i.e. that such as $p(B_i \cap B_j) = 0$ and $\sum p(B_i) = 1$. With cards, this could correspond to heart, diamond, clover and spade. This could also be B and $\bar{B}$. We have then the expansion theorem

$$p(A|H) = \sum_i p(A \cap B_i|H) = \sum_i p(A|B_iH) \cdot p(B_i|H). \tag{3.7}$$

For some cases, this sum can be written as an integral. With cards again, this means that the probability to draw a given card is the probability to draw this card among hearts, plus the probability to draw it among diamonds, plus the probability to draw it among clovers plus the probability to draw it among spades, assuming that the game is not biased.

Naturally, with cards, such a relation is not useful because there is the same number of cards for each hypothesis. Let's take another example. Suppose that a trap can catch both mice and shrews without any discrimination (hypothesis H). There are 60% of mice. If there are 57% of females among the shrew and 53% among the mice, what is the probability to catch a female? Formula (3.7) gives

$$p(F|H) = p(F|M) \cdot p(M|H) + p(F|S) \cdot p(S|H) = 0,53 \times 0,60 + 0,57 \times 0,40 = 54,6\%. \tag{3.8}$$

Remember Laplace's words...

3.1.4 Bayes Theorem

Bayes theorem can be easily deduced from equation (3.3),

$$\boxed{p(B|AH) = \frac{p(B|H) \cdot p(A|BH)}{p(A|H)}.} \tag{3.9}$$

As it links two inverse probabilities, $p(B|AH)$ and $p(A|BH)$, it is widely used in many domains of sciences. Bayes derived it in a very specific case [33]. Pierre Simon, Marquis de Laplace, a former student of the university of Caen derived it in this way in his famous "traité philosophique des probabilités" [35].

This general result is always called 'Bayes theorem', although Bayes never wrote it; and it is really nothing but the product rule of probability theory which had been recognised by others, such as James Bernoulli and A. de Moivre (1718), long before the work of Bayes. Furthermore, it was not Bayes but Laplace (1774) who first saw the result in generality and showed how to use it in real problems of inference.

Bayes' manuscript was transmitted to the Royal Society two years after his death by Richard Price, on the 10th of November 1763. Laplace's version appeared 11 years later when he was 25 years old:

> ≪ *Si un événement peut être produit par un nombre* n *de causes différentes, les probabilités de l'existence de ces causes prises de l'événement sont entre elles comme les probabilités de l'événement prises de ces causes, et la probabilité de l'existence de chacune d'elles est égale à la probabilité de l'événement prise de cette cause, divisée par la somme de toutes les probabilités de l'événement prises de chacune de ces causes.* ≫
> Laplace, 1774, *Mémoire sur la probabilité des causes par les événements* [34].

If you can understand, it means that your French has really improved!

There are other expressions of Bayes theorem when there are several possibilities. For a given B_j, expanding the denominator with formula (3.7), equation (3.9) becomes,

$$\boxed{p(B_j|AH) = \frac{p(B_j|H) \cdot p(A|B_jH)}{\sum_i p(A|B_iH) \cdot p(B_i|H)}.} \tag{3.10}$$

The denominator can sometimes be written as an integral as we shall see later. This version is also due to Laplace.

Laplace was also the first to apply what he called "my formula" to physics. In particular, he determined the uncertainty of mass of Saturn:

> *"In order to give some interesting applications of it I have profited by the immense work which M. Bouvard has just finished on the movements of Jupiter and Saturn. His calculations give him the mass of Saturn equal to 3512th part of that of the sun. Applying to them my formulae of probability, I find that it is a bet of 11,000 against one that the error of this result is not 1/100th of its value..."*
> Laplace, 1815, *Essai philosophique sur les probabilités.*

According to the latest value determined by the NASA, he is still winning his bet!

3.1.5 Bayesian Inference

Let us first consider a simple example to understand Bayesian inference. Before proposing surgery to a patient, a medical doctor orders a medical test to confirm his/her diagnostic. Here D means "the patient has the disease". The medical test can be either positive or negative. Here P means "the test is positive". But the test is not 100% reliable: it has a probability $p(P|D)$ to be positive if the patient has the disease. It can also have the probability $p(P|\bar{D})$ to be positive even if the patient does not have the disease. This is called a fake positive.

If the test is positive, was is the probability that the patient has the disease, denoted $p(D|P)$? A way to answer to the question is to repeat the medical tests and perform a statistical analysis. But it is usually not possible to collect a large amount of blood and it is very expensive. How to answer to the question with a single test?

As we are interested in the Bayes theorem gives

$$p(D|P) = \frac{p(D) \cdot p(P|D)}{p(P)}. \tag{3.11}$$

Here:

- $p(D)$ is the probability that the patient has the disease estimated by the medical doctor before the test. It is usually called the **prior** [prior]. It includes all the knowledge before the measurement or the test here.
- $p(D|P)$ is called the posterior and it consists in the probability of having the disease after a positive test.
- $p(P|D)$, i.e. the probability that the test is positive if the patient has the disease, is given by the company after having tested it on many sick patients.
- The denominator $p(P)$ includes all the ways to have a positive test: $p(P) = p(P|D)p(D) + p(P|\bar{D})p(\bar{D})$. Here $p(\bar{D}) = 1 - p(D)$ and $p(P|\bar{D})$ is the probability to have a fake positive result, also determined by the maker of the test.

Finally, Bayes theorem leads to

$$p(D|P) = \frac{p(D) \cdot p(P|D)}{p(P|D)p(D) + p(P|\bar{D})(1 - p(D))}. \tag{3.12}$$

Let's do some numerical applications to understand. We shall consider several cases:

- The test is 100% reliable: $p(P|D) = 1$ et $p(P|\bar{D}) = 0$. Then $p(D|P) = 1$, whatever the prior. The patient has the disease for sure if the test is positive, as expected.
- The maker of the test guarantees $p(P|D) = 0.90$ and $p(P|\bar{D}) = 0.05$. If the medical doctor is not sure of the diagnostic and determines the prior to $p(D) = 1/2$, then $p(D|P) = 0.90/0.95 = 0.95$. Test has improved the knowledge.
- One still has $p(P|D) = 0.90$ and $p(P|\bar{D}) = 0.05$. But the medical doctor strongly suspects the disease because of clinical symptoms and determines the prior to $p(B) = 0.80$. Then, $p(D|P) = 0.99$! The medical doctor and the test can be both wrong, but with an extremely low probability.
- Let's see what's happen if the doctor and the medical test give contradictory results. Suppose that the medical doctor determines the prior to $p(D) = 0.10$, but the test is positive. Then, $p(D|P) = 0.67$, which is close to $1/2$. Other medical tests are necessary.
- Eventually, let's consider the case of positive test in the framework of the screening of a rare disease. As a prior, we can choose the occurrence of the disease $p(D) = 0.001$ which leads to $p(D|P) = 0.02 = 2\%$. Having a positive test in such a case does not mean that you have the disease. Other tests should be done before concluding.

As a conclusion, we can see that Bayes theorem is used for two purposes. It shows how to improve the knowledge the result contained in the prior by including the result of an experiment. It also shows how to inverse the probability: in the example we requested $p(D|P)$ knowing $p(P|D)$.

Bayes theorem allows to express what we learn from experiences, starting from what we new before. It is completely different way of thinking than the frequentist approach that corresponds to the Shadok's logic, see the cartoon of figure 3.2. How to take into account that the multiple failures change the likelihood of success?

3.1.6 Another presentation

As the Bayes theorem is not easy to explain for a general audience, Bayesian inference is sometimes presented as a deduction tree, see figure 3.3. It is very easy for such an example with only four possibilities: yes or no for the disease and for the test.

Let's consider the last example of a medical test during a screening. Considering 1 000 persons, if the occurrence of the disease is $p(D) = 0.001$, one have the disease and 999 don't. The sick person has 90% chance

Figure 3.2: Shadoks are stupid animals who want to escape their planet. Prof. Shadoko, the most intelligent Shadok, calculated that there is a chance over a million that their rocket succeeds. So they hurry to have 999 999 failures to have one success.
Shadok saying: *"If we keep trying, we finally succeed. Hence, the more we fail the likelier we will succeed."* (Cartoon found on the web)

to get a positive test. Among the 999 healthy persons 50 might have a positive test. As a consequence, among the positive tests, one is a true positive and 50 are fake positives. Eventually, the probability to have the tested disease if the test is positive is only 2%.

The final result is nothing else than Bayes theorem applied to a medical test.

3.2 Determination of the prior

3.2.1 Problematic

Bayesian approach has long being considered too subjective because of the prior determination. We need rules to have each person doing a Bayesian evaluation of the uncertainty using the same prior. How to express the knowns and the unknowns in terms of probability?

Imagine that one tosses a biased coin without knowing if it head or tail that is favoured. If we use a Bayesian inference to determine the answer, we have no other choice than assigning a probability $\frac{1}{2}$ to each face as a prior.

Two assumptions underly the approach:

1. Probabilities represent the state of our knowledge and ignorance. They don't estimate the frequency of an event, but its likelihood.

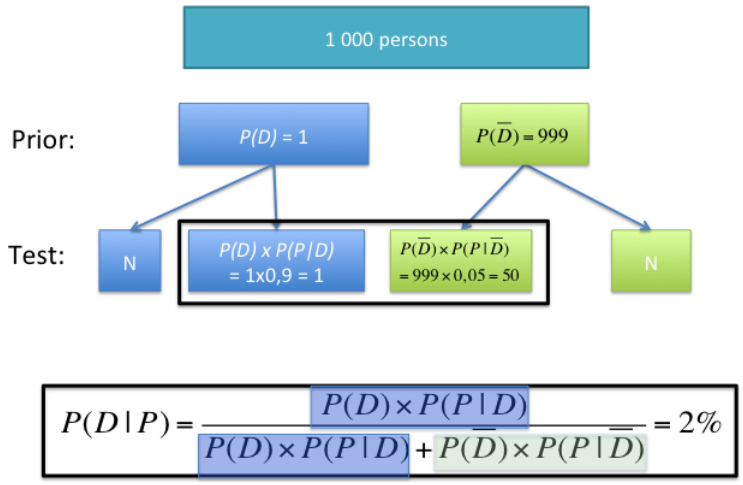


Figure 3.3: Deduction tree to evaluate the probability to have a given disease after a positive test during a screening. See text for explanations.

2. If one has no reason to think that A is more likely than B one writes $p(A) = p(B)$. It is the **principle of insufficient reason** [principe de la raison insuffisante] of Jacob Bernoulli. In 1921 Keynes proposed another formulation: **principle of indifference** [principe d'indifférence].

If one knows that the values taken by a given quantity belong to the interval $[a, b]$, which prior to choose? If one divides this interval into n subintervals of equal length, we assign the same probability to each subinterval. As the result does not depend on the the number of subintervals, this leads to a uniform probability on the interval $[a, b]$. This is exactly the choice done to estimate the reading uncertainty.

Such an approach is difficult to apply to more complicated situations. We rather use the concept of information entropy introduced by Shannon [42]. The rule is to maximise the information entropy taking into account the available robust information following a method developed by Lagrange.

3.2.2 Information theory

This subsection on information theory was inspired by the book of Emmanuel Dion [18] and the lecture notes of Gibert and Lopes [19].

Information theory has been developed for telecommunication engineering. To evaluate the quantity of information that can be transferred into a cable, one has to define the quantity of information and its flux. That is how the *bit* and the *entropy* of information were introduced. This concepts were later applied to statistical physics and inferences.

How information and probabilities are linked? Suppose that you need to identify an individual knowing his/her name. If only one individual bears that name, identification is straightforward. If N persons share the same name, there are N possibilities over a given population to identify him/her. Additional information has to be brought to identify the person. In other words, rare family names carry more information than common family names because they can precisely identify a person. Common names that designate several persons, carry less information as an ambiguity remains. Similarly, the word "the" does not carry much information.

To have a quantitative definition, let's consider a simple example: suppose that you need to identity a criminal among 10 000 persons. If you don't have any information, you have one chance over 10 000 to find it.

If you know that he is a male and the males represents 50% of the population, you have one chance of 5 000 to identify the criminal. This piece of information has reduced by a factor 2 the number of possibilities. If the police knows that the criminal has white hair and that 1 person over 16 has white hair, the number of possibilities has been reduced by a factor 16. With both pieces of information, the number of possibilities has been reduced by a factor 32.

The operator chosen to quantify the information should be additive. This is why the logarithm is chosen. One also prefers positive quantities. Claude Shannon and some others before him, have defined the quantity of information as $I = \log(1/p) = -\log p$. There are other possible definitions using the total number of possibilities N. Then $\log N$ is used to keep a positive quantity. As most of the information transferred by telecommunications is binary (0 or 1), $\log_2$ is preferred. One bit corresponds to $I = \log_2 2 = 1$.

If one goes back to the previous example, the first piece of information (male) is 1 bit. The second piece of information (white hair) is 4 bit because $16 = 2^4$. Both pieces of information together represent $1 + 4 = 5$ bit because $32 = 2^5$.

Shannon defined the **entropy of information** [entropie de l'information] as the mean information:

$$H = \sum_i p_i \log_2(1/p_i) = -\sum_i p_i \log_2(p_i), \tag{3.13}$$

where the p_i are the probabilities of each possibility. The name entropy was chosen for the analogy of the formula with the physical entropy introduced by Boltzmann. The link between information entropy and Boltzmann entropy is not obvious and goes beyond the scope of this lecture. Having the same name can even be considered as misleading for our purpose. H is sometimes called *uncertainty* or *missing information*.

Back to the previous example, the entropy of information about the gender of the criminal is

$$H = \frac{1}{2}\log_2 2 + \frac{1}{2}\log_2 2 = 1. \tag{3.14}$$

The entropy of information related to the white hair is

$$H = \frac{1}{16}\log_2 16 + \frac{15}{16}\log_2(16/15) = 0,34. \tag{3.15}$$

Equiprobable cases have the largest entropy. This is this property that we use to define the prior that should assign the same probability to any possibilities.

In case of measurements, the lower the number of values that can be assigned to a quantity the more accurate the results. Then the measurement seen as a learning process aims to reduce the number of possibilities.

In the 1980's, John E. Shore and Rodney W. Johnson made an axiomatic derivation of the principle of maximum entropy that is shown to be the uniquely correct method for inductive inference [44]. For a recent review on the subject, see Ref. [30]. I will not reproduce the derivation but rather focus on the way to use the MaxEnt principle to determine the prior.

3.2.3 Maximisation of entropy to determine the prior

Information entropy is defined as

$$S = -\sum_i p_i \ln(p_i) \tag{3.16}$$

for discrete cases and

$$S = -\int p(x)\, \ln(p(x))\, dx \tag{3.17}$$

for continuous one. As we are only interested in the maximum value, one can use any logarithm.

New international standards recommend to apply the principle of maximum entropy under constraints given by the available information to determine the adequate PDF used as the prior. Supplement 1 to the GUM [7] lists several probability density functions that maximise the information entropy under various information constrains. It also explains how to numerically simulate these functions. Some examples are listed in the section 2.6.

Example: Coin as a binary case

Let us consider again the example of the biased coin for which there are only two possibilities, head and tail. p_1 is the probability to have a head after tossing the coin and p_2 tail. One has $p_1 + p_2 = 1$ and the information entropy reads

$$S = -p_1 \ln(p_1) - p_2 \ln(p_2) = -p_1 \ln(p_1) - (1 - p_1) \ln(1 - p_1). \tag{3.18}$$

It is plotted in Fig. 3.4 where it is clear the maximum entropy corresponds to equal possibilities, i.e. $p_1 = p_2 = 0.5$.

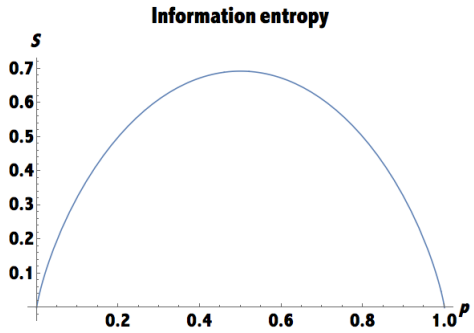


Figure 3.4: Information entropy as a function of the probability p to draw a head (or a tail).

We can also derive S with respect to p_1 to find the extremum. One gets $p_1 = 1/2$ as expected.

With several possibilities

Let us assume now that there N responses and the probability to get the response k is p_k. The information entropy is

$$S = -\sum_{k=1}^{N} p_k \ln(p_k) \qquad \text{with} \qquad \sum_{k=1}^{N} p_k = 1. \tag{3.19}$$

Then,

$$\frac{\partial S}{\partial p_i} = -\ln(p_i) - 1 - \sum_{k \neq i} \frac{\partial p_k}{\partial p_i} \ln(p_k) - \sum_{k \neq i} \frac{\partial p_k}{\partial p_i}. \tag{3.20}$$

As $p_k = 1 - \sum_{j \neq k} p_j$, one has $\frac{\partial p_k}{\partial p_i} = -1$ if $i \neq k$. Inserting it into the above equation,

$$\frac{\partial S}{\partial p_i} = -\ln(p_i) + \sum_{k \neq i} \ln(p_k) + N - 2 = 0. \tag{3.21}$$

We can write a similar equation after deriving the information entropy with respect to p_j and subtract it to the previous result and get

$$-\ln(p_i) + \ln(p_j) + \sum_{k \neq i} \ln(p_k) - \sum_{k \neq i} \ln(p_k) = 2\ln(p_j) - 2\ln(p_i) = 0. \tag{3.22}$$

Finally, the extremum of entropy is achieved when $p_i = p_j \; \forall i, j$.

Maximising with information constrains: rectangular distribution

If the only available information regarding a quantity is a lower limit a and an upper limit b with $a < b$, then, according to the principle of maximum entropy, a rectangular distribution over the interval $[a, b]$ should be assigned to this quantity.

Be $p(x)$ probability density function that maximises the information entropy under the constrains

$$\int_a^b p(x)\,dx = 1. \tag{3.23}$$

By hypothesis, we also know that $p(x) = 0$ if $x < a$ or $x > b$.

Lagrange's method consists in maximising

$$I[p(x)] = S + \lambda[1 - \int_a^b p(x)\,dx] = \int_a^b \left(-p(x)\ln(p(x)) + \lambda[\frac{1}{b-a} - p(x)]\right) dx. \tag{3.24}$$

The additional term multiplied by λ is equal to zero. Here λ is just an arbitrary artificial variable called Lagrange multiplier.

If $p_0(x)$ is the probability density function that maximises $I[p(x)]$, $I[p_0(x) + \delta p(x)] \simeq I[p_0(x)]\,\forall \delta p(x)$ to the first order in $\delta p(x)$. Expanding the integrand to the first order, we get

$$I[p_0(x) + \delta p(x)] \simeq \int_a^b \left(-p_0(x)\ln(p_0(x)) + \lambda[\frac{1}{b-a} - p_0(x)]\right) dx - \int_a^b (\ln(p_0(x)) + 1 + \lambda)\,\delta p(x)\,dx. \tag{3.25}$$

The second term is equal to zero for all $\delta p(x)$ if $\ln(p_0(x)) + 1 + \lambda = 0$ which means that $p_0(x) = e^{-(1+\lambda)}$. One gets a uniform function over the defined interval. Its value or the Lagrange multiplier can be determined with the normalisation constrain given in Eq. (3.23). We eventually get that over the interval $[a, b]$,

$$p_0(x) = \frac{1}{b-a}. \tag{3.26}$$

This corresponds to the rectangular distribution as expected.

Maximising with information constrains: Gaussian distribution

If the mean value and the associated standard deviation are the only information available regarding a quantity, then, according to the principle of maximum entropy, a Gaussian probability distribution should be assigned to it.

Be $p(x)$ a probability density function with a mean value equal to μ and a variance to σ^2. Assuming that this is the only available information, we can determine $p(x)$ maximising information entropy, under the constrains

$$\mu = \int x\,p(x)\,dx, \qquad \sigma^2 + \mu^2 = \int x^2\,p(x)\,dx \qquad \text{and} \qquad 1 = \int p(x)\,dx. \tag{3.27}$$

Using Lagrange multipliers α, β and γ, we look for the function $p(x)$ that maximise the following integral,

$$I[p(x)] = \int p(x)[-\ln(p(x)) + \alpha x + \beta x^2 + \gamma]\,dx. \tag{3.28}$$

Missing terms are constant.

If $p_0(x)$ is the probability density function that maximises $I[p(x)]$, $I[p_0(x) + \delta p(x)] \simeq I[p_0(x)]$ to the first order in $\delta p(x)$. Expanding the integrand to the first order, we get

$$I[p_0 + \delta p] = \int (p_0 + \delta p)[-\ln[p_0 + \delta p] + \alpha x + \beta x^2 + \gamma]\,dx \tag{3.29}$$

$$= \int p_0(x)(-\ln[p_0(x)] + \alpha x + \beta x^2 + \gamma)dx + \int \delta p(x)[-1 - \ln[p_0(x)] + \alpha x + \beta x^2 + \gamma]dx \tag{3.30}$$

Second term of the right hand side of the previous equation should equal to zero whatever $\delta p(x)$. Hence,

$$-1 - \ln[p_0(x)] + \alpha x + \beta x^2 + \gamma = 0. \tag{3.31}$$

As a consequence, $p_0(x)$ reads,

$$p_0(x) = \exp[\gamma - 1 + \alpha x + \beta x^2]. \tag{3.32}$$

Multipliers α, β and γ are determined using the constrains given by equations (3.27). One finally gets,

$$p_(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{3.33}$$

which is a Gaussian distribution! As a conclusion the probability density function that maximises the information entropy when we only know the mean value and the variance is a Gaussian.

This example shows how to treat our ignorance through the entropy and our robust knowledge through Lagrange multipliers to determine the prior that takes both into account.

3.2.4 Jeffrey's prior

Sir Harold Jeffreys (22 April 1891 - 18 March 1989) was a British mathematician, statistician, geophysicist, and astronomer. The book that he and Bertha Swirles wrote *Theory of Probability*, which first appeared in 1939, played an important role in the revival of the Bayesian view of probability.

Examples of priors intended to be non-informative are flat priors although they cannot be normalised when the interval is infinite. But the choice of the relevant parameter is not always well determined. Consider, for example, the radioactive decay of a given isotope: the probability to observe a decay at a time t_i is given by

$$P(t_i|\lambda) = \lambda e^{-\lambda t}, \tag{3.34}$$

Then the probability of observing a given set of times $\{t_1, \ldots, t_n\}$ is

$$P(\{t_i\}|\lambda) = \prod_{i=1}^{n} P(t_i|\lambda) = \lambda^n e^{-\lambda \sum t_i}, \tag{3.35}$$

as each nuclei are independent with each other.

If we want to determine λ from a set of observations $\{t_1, \ldots, t_n\}$, we should determine $P(\lambda|(\{t_i\})$ with the help of the Bayes theorem. But we could also consider $\tau = 1/\lambda$ as a parameter. Assuming a flat non-informative prior for λ leads to

$$P(\tau)d\tau = P(\lambda)d\lambda \Leftrightarrow P(\tau) = P(\lambda)\left|\frac{d\lambda}{d\tau}\right| \propto \frac{1}{\tau^2}. \tag{3.36}$$

Thus, the prior probability distribution for τ is not uniform anymore. In other words, Bayes theorem with λ as a variable with a flat prior does not lead to the same result as with τ as a variable with a flat prior. This cannot be accepted.

However, if we choose as a prior $P(\lambda) = 1/\lambda$ and $P(\tau) = 1/\tau$ which cannot be normalised, one has

$$P(\lambda)d\lambda = P(\tau)d\tau. \tag{3.37}$$

For general cases, Jeffreys' prior is a non-informative prior distribution for a parameter space that is proportional to the square root of the determinant of the Fisher information matrix. It has the key feature to be invariant under reparameterisation of the parameter. It depends upon the statistical model. See books dedicated to Bayesian data analysis for details.

3.2.5 Subjectivity

The present form of Bayes theorem is due to Laplace. Not only did Laplace rediscovered Bayes theorem by himself, in far more clarity, he also put it to good use in solving problems in celestial mechanics, medical statistics and even jurisprudence. Despite Laplace's numerous successes, his developments of probability theory was rejected by many soon after his death, and qualified as "subjective probabilities". The main point of concern tends to be the choice of the prior. However, we have seen how invariance arguments, under the appropriate transformation, can often determine a PDF uniquely. MaxEnt methods also provide unambiguous PDFs [15].

Another concern is related to the fact that some PDFs which represent complete ignorance cannot be normalised. However, the resulting posterior can generally be normalised. As we shall see in the next chapter, the least-squares method relies on the implicit use of a uniform prior, which cannot be normalised.

Note that Ronald Fisher (1890-1962), a famous statistician, was a critical figure in the history of Bayesian analysis. He claimed: "The theory of inverse probability is founded upon an error, and must be wholly rejected". Together with Neyman, they were the two most important persons involved in the demise of inverse probability in the early twentieth century.

3.3 Application to uncertainties

3.3.1 Type A evaluation

In the GUM, the type A evaluation is based on a frequentist approach. These is a Bayesian determination of type A evaluation and articles that call for a revision of the GUM. See e.g. the article of Kacker and Jones [46].

Be n values of a measurement $z_1, \ldots, z_n$ characterised by their mean value $\bar{z}$ and standard deviation $s(\bar{z})$. These data are considered as fixed, not fluctuating. Uncertainty is in the knowledge of the true value of the measured quantity.

Assuming that we have no specific prior information on the values that can be taken by this quantity, thus choosing a non-informative prior, i.e. a uniform value, Bayes theorem leads to difficult calculations because it requires the inversion of Student distribution. Here is the final result.

The probability density function of the true value of the measured quantity is characterised by an expectation

$$E(Z) = \bar{z}, \tag{3.38}$$

and a standard deviation

$$u_B(\bar{z}) = \sqrt{\frac{n-1}{n-3}} s(\bar{z}). \tag{3.39}$$

When n is large enough, the corrective term does not change much. Frequentist and Bayesian approaches meet. For small values of n, it is not the case anymore. Bayesian approach requires that $n > 3$ to be valid.

3.3.2 Radioactivity and Poisson distribution

Definition

When we count a number of events, we should not do any mistake and we generally assume that there is no uncertainty. Except if the detector saturates for example. We will assume here that counting is exact. However, the physical process itself can fluctuate. It is particularly the case with radioactivity. If one counts several times the same quantity, one gets various values.

Poisson distribution gives the probability to count N events when the mean value is equal to μ:

$$P(N|\mu) = \frac{\mu^N}{N!} e^{-\mu}. \tag{3.40}$$

Here N is an integer but not μ. $N!$ means N factorial. Poisson distribution is plotted in figure 3.5.

Siméon Poisson is a French mathematician who claimed that *"Life is good for only two things, discovering mathematics and teaching mathematics."*

The function $P(N|\mu)$ is normalised to one. Mean value of Poisson distribution can be easily calculated and is equal to μ. Variance is also equal to μ and the standard deviation to $\sqrt{\mu}$.

Demonstrations are straightforward using the power series of the exponential function. Normalisation is easy,

$$\sum_{N=0}^{\infty} \frac{\mu^N}{N!} e^{-\mu} = e^{-\mu} \sum_{N=0}^{\infty} \frac{\mu^N}{N!} = e^{-\mu} e^{\mu} = 1. \tag{3.41}$$

Mean value also,

$$\bar{N} = \sum_{N=0}^{\infty} N \cdot P(N|\mu) = e^{-\mu} \sum_{N=1}^{\infty} \frac{N\mu^N}{N!} = e^{-\mu}\mu \sum_{N=1}^{\infty} \frac{\mu^{N-1}}{(N-1)!} = e^{-\mu}\mu e^{\mu} = \mu. \tag{3.42}$$

For the variance, it's a bit longer, but still easy:

$$\begin{aligned}
\overline{N^2} &= \sum_{N=1}^{\infty} N^2 P(N|\mu) = e^{-\mu}\mu \sum_{N=1}^{\infty} N \frac{\mu^{N-1}}{(N-1)!} && (3.43)\\
&= e^{-\mu}\mu \left(\sum_{N=1}^{\infty} (N-1) \frac{\mu^{N-1}}{(N-1)!} + \sum_{N=1}^{\infty} \frac{\mu^{N-1}}{(N-1)!} \right) && (3.44)\\
&= e^{-\mu}\mu \left(\mu e^{\mu} + e^{\mu}\right) = \mu^2 + \mu = (\bar{N})^2 + \mu. && (3.45)
\end{aligned}$$

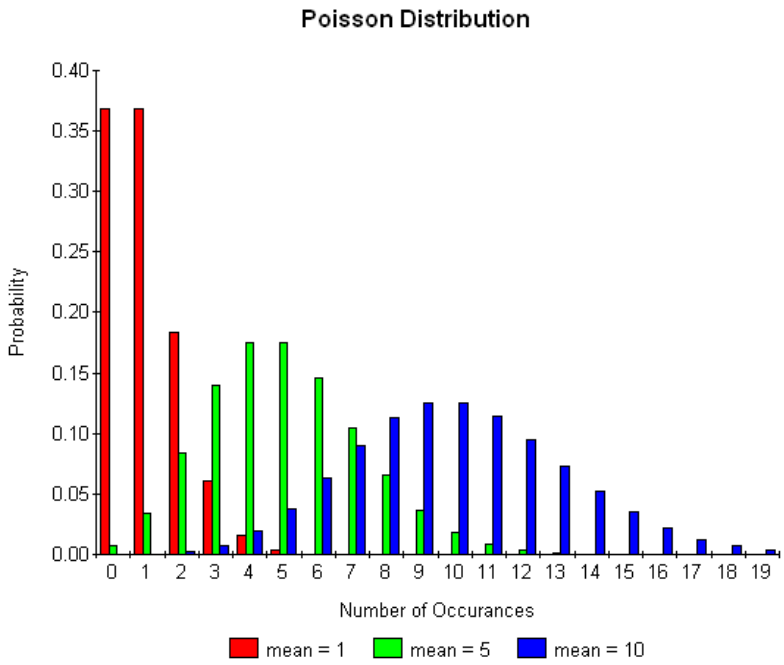


Figure 3.5: Poisson distribution (figure reproduced from `systomath.com`).

When μ is large enough, Poisson distribution tends to be similar to the Gaussian distribution with a mean value μ and a variance μ. As a matter of facts Poisson distribution is close to a Gaussian when $\mu > 9$. We can therefore use what we learned about Gaussian to estimate the coverage factor.

Inversion of Poisson distribution

The number of disintegrations of a radioactive process satisfies the Poisson distribution. This is experimentally well established. Knowing the true value μ, the probability to detect N events is equal to

$$P(N|\mu) = \frac{e^{-\mu}\mu^N}{N!}. \tag{3.46}$$

But during an experiment, we measure N_1 events and we would like to estimate the probability density function of the true value μ, $P(\mu|N_1)$. There are two ways to proceed. We can repeat the measurements a large number of times and calculate the mean value, but it is expensive and time consuming.

Another way is to evaluate $P(\mu|N_1)$ using Bayes theorem:

$$P(\mu|N_1) = \frac{P(N_1|\mu)P_{prior}(\mu)}{P(N_1)}. \tag{3.47}$$

Here $P(N_1|\mu)$ is the Poisson distribution and $P_{prior}(\mu)$, the prior. If we don't have any specific information about the values that can be taken by the measured quantity, the prior is uniform on the interval $[0, +\infty[$: $P_{prior}(\mu) = C$. We know for sure that the activity is positive. Denominator takes into account all the ways to get N_1, from all values of μ,

$$P(N_1) = \int_0^\infty P(N_1|\mu)\, P_{prior}(\mu)\, d\mu = \ldots = C. \tag{3.48}$$

The result is easily obtained integrating N_1 times by parts.

Eventually, the inversion of the Poisson distribution using the Bayes theorem leads to

$$P(\mu|N_1) = \frac{e^{-\mu}\mu^{N_1}}{N_1!}. \tag{3.49}$$

Although this expression looks like a Poisson distribution, it is not one. Here, N_1 is a fixed parameter and μ a random variable. It is the contrary for the Poisson distribution. It is worthwhile to remember that $P(\mu|N_1)$ represent the likelihood of a given value of μ knowing the result of a single measurement N_1. For more details, see ref. [48].

Calculation of the mean value and the variance is easy. Expectation is

$$E(\mu|N_1) = \int_0^{+\infty} \mu P(\mu|N_1)\, d\mu = (N_1+1)\int_0^{+\infty} \frac{e^{-\mu}\mu^{N_1+1}}{(N_1+1)!} = N_1+1. \tag{3.50}$$

Similarly, the variance is also N_1+1. This is a confirmation that $P(\mu|N_1)$ and $P(N_1|\mu)$ differ. They don't have the same mean value and the same variance. But their plot looks similar.

We will know apply these results to the determination of the decision threshold, the detection limit and the limits of the confidence interval. But before, let us consider multiple determinations of the activity.

Multiple evaluation

The previous result relative to the mean value of μ, i.e. $\bar{\mu} = N_1 + 1$ and $u(m\bar{u}) = \sqrt{N_1+1}$, is only valid for a single measurement.

If a second independent measurement is done, finding N_2 counts, we can perform the same calculation using the PDF after the first measurement, eq. (3.49, as a prior for the second measurment in Bayes theorem. After integrating (N_1+N_2) times by parts the normalisation integral, the PDF after two measurements reads

$$P(\mu|N_1,N_2) = 2^{N_1+N_2}\frac{\mu^{N_1+N_2}\times e^{-2\mu}}{(N_1+N_2)!}. \tag{3.51}$$

Then, one can easily evaluate the mean value

$$\bar{\mu} = \frac{N_1+N_2+1}{2}, \tag{3.52}$$

and the variance

$$u^2(\mu) = \frac{N_1+N_2+1}{4} = \frac{\bar{\mu}}{2}. \tag{3.53}$$

With three independent measurements, one would get

$$\bar{\mu} = \frac{N_1+N_2+N_3+1}{3}, \tag{3.54}$$

and

$$u^2(\mu) = \frac{N_1+N_2+N_3+1}{9} = \frac{\bar{\mu}}{3}. \tag{3.55}$$

And so on.

When the number of measurements increase $\bar{\mu}$ and $\bar{N}$ are getting closer and closer as expected. The variance is decreasing like for other experiments.

3.3.3 Near a threshold

Let us consider a case with a poor result of measurement, with a relative uncertainty of 50%: $x = 2\pm1$, whatever the unit. If it is a quantity that we know to be positive, like a mass, a concentration, a length and so on, such a result is problematic. Assuming that the values have a Gaussian distribution, 2.5% of the values are negative, which is impossible. How to take into account the fact that we are certain that the result is positive in case of a poor measurement result?

Let us suppose that it is a mass. We don't measure directly a mass. Depending on the technique used, we measure the deviation of an index in case of a scale or the deviation of a beam in case of mass spectroscopy... Of course, measuring a mass generally gives more accurate results, but in case of neutrinos or celeste bodies uncertainty is still large.

Let us denote x the quantity that is directly measured and m the mass. A model gives the link of the deviation x a function of the mass after calibration. This is generally written with a mathematical function

$x = f(m)$. But after a measurement, we want the inverse, i.e. the mass m as a function of the deviation x. It is again Bayes theorem that will allow us to answer to this problem:

$$P(m|x) = \frac{P_{prior}(m) \times P(x|m)}{\int P_{prior}(m) \times P(x|m)\, dm}. \tag{3.56}$$

As we are sure that the mass is positive, we choose as a prior the Heavyside function $P_{prior}(m) = \theta(m)$ that is equal to 1 if $m > 0$ and to 0 if $m < 0$.

To continue, one needs the probability density function of the experimental results $P(x|m)$. For the sake of simplicity, we will suppose that it is a Gaussian. We can then solve analytically the problem. Then,

$$P(m|x) = \frac{\exp\left[-\frac{(x-f(m))^2}{2u_x^2}\right]}{\int_0^{+\infty} \exp\left[-\frac{(x-f(m))^2}{2u_x^2}\right] dx}. \tag{3.57}$$

Eventually, if we assume that the deviation is well calibrated in kilogramme or any other unit of mass, one directly has $m = x$. Calculation of $P(m|x)$ is then straightforward. It is then easy to calculate $\bar{m}$, the mean value of the masse and its associated uncertainty. Mathematical expressions are heavy, I will not write them here.

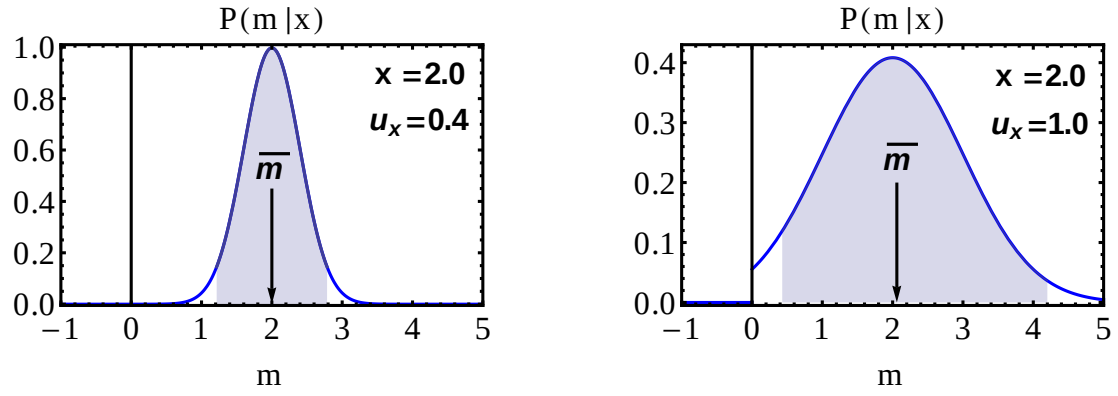


Figure 3.6: Representation of the probability density function of the masse m for a given deviation x for two different values of the associated uncertainty. The confidence interval that includes 95% of the values is also indicated with shade. Figure reproduced from [32].

Figue 3.6 represents the probability density function $P(m|x)$ for two different uncertainties in x. The mean value is indicated by an arrow. The shaded zone corresponds to the 95% confidence interval, i.e. it includes 95% of the values that can be taken by m.

If the measurement result has a small uncertainty, one finds that $\bar{m} = x$ and $P(m|x) \simeq P(x)$, as we always do. But, if the uncertainty is large compared to the value, one has to correct the experimental result by taking into account the fact that the mass cannot be negative. The mean value for the mass, $\bar{m}$, is not equal to x anymore.

If $x = 2 \pm 1$, one gets $\bar{m} = 2.055$. There is a small shift. The mean value is not the most probable one and the confidence interval is not symmetric around $\bar{m}$. Here, the confidence interval that includes 95% of the values is $[0.43; 4.2]$. Without a Bayesian treatment, confidence interval that includes 95% of the values would have been $[0.04; 3.96]$. There is a significative change.

This simple example shows of to take into account the measurement results and other knowledge to evaluate the uncertainty. The Bayesian approach is now used in international standards. This is the case, for example, for the measurement of radioactivity that is presented in the next chapter.

Chapter 4

Testing hypothesis and radiation measurements

"I conceive the mind as a moving thing, and arguments as the motive forces driving it in one direction or the other."

John Craig (1699)

4.1 Introduction

Many measurements are affected by a noise. Thus, the signal that carries information, has to be separated from the noise that is added on to signal due to many natural reasons. This is particularly the case for radiation measurements where background radiations generates a noise.

Figure 4.1 represents a spectrum where a peak is expected in a given region of interest. In this figure, the peak so small that we might wonder whether there is a signal from a contaminent or is it a simple fluctuation of the background? How to decide? From which value can we say that there is a signal? The answer is based on hypothesis testing.

In case of ionising radiations, an international standard (ISO 11929) defines the **decision threshold** [seuil de décision], the **detection limit** [limite de détection] and the **limits of the confidence interval** [extrémités de l'intervalle de confiance]. Note that this international standard is under revision. This new standard [9] is partially based on Bayesian inference and difficult to understand but one can find pedagogical articles that help [29].

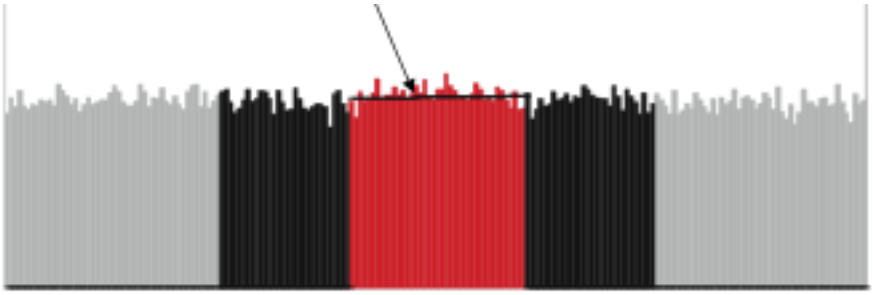


Figure 4.1: Search of a spike in a spectrum. (Figure reproduced from [29])

Moreover, even if we are convinced that there is a contaminant, what is the uncertainty? If one finds for example (2 ± 1) Bq, there is significant probability to have negative values that are a physical non sense. Previous formulas of the GUM have to be corrected. This issue is also addressed by the international standard.

First, we should learn basics of hypothesis testing.

4.2 Hypothesis testing

4.2.1 Basic concepts

A hypothesis is an educated guess about something in the world. It should be testable, either by experiment or observation on a statistical sample. Thus, a branch of the statistics is devoted to hypothesis testing and has wide applications in many branches of sciences.

One generally defines a hypothesis that can be tested, like *the model is right*, called the null hypothesis [hypothèse nulle] and denoted H_0. In the above introduction, it can be that there is no additional contamination. The null hypothesis is generally the accepted fact. We also define an alternative hypothesis [hypothèse alternative].

Then, one tests a statistical sample, with the goal of accepting or rejecting the null hypothesis. In essence, hypothesis testing is a procedure to compute a probability that reflects the strength of the evidence (based on a given sample) for rejecting the null hypothesis. At some point, we have to determine a threshold or cut-off point (called the critical value) to decide when to believe the null hypothesis and when to believe the alternative hypothesis.

Fig. 4.2 shows the distribution of the values that can be taken by the tested quantity within the H_0 hypothesis. If the tested value lies within the 95% confidence interval, the H_0 hypothesis is accepted. However, if it lies in the rejection region located at the two extremes of the distribution in the present case, we reject the hypothesis H_0, as the probability that it is true is lower than $\alpha = 0.05$.

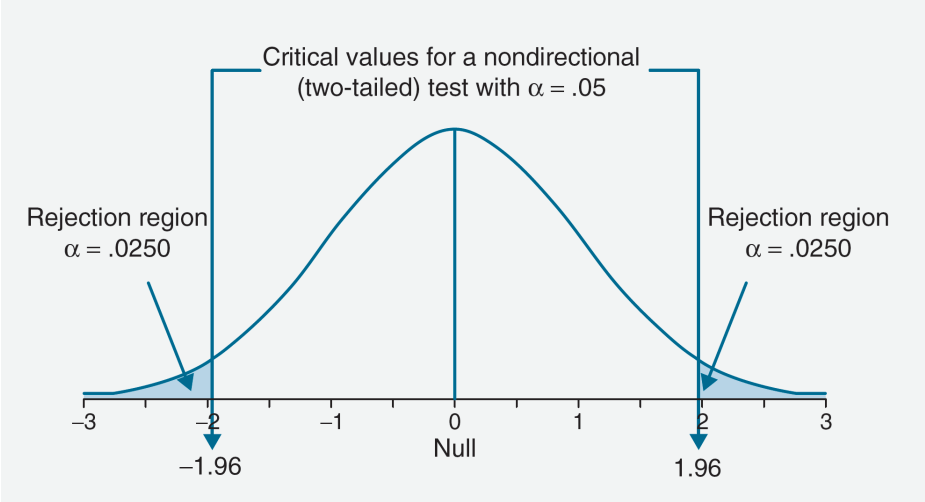


Figure 4.2: Distribution of the values that can be taken by the tested quantity within the H_0 hypothesis. The abscisse quantity is centred on the mean value and divided by the standard quantity.

The rejection region [région de rejet] is the set of values of the test statistic for which the null hypothesis is rejected. It is generally based on the significance level [seuil ou niveau de signification] of the test or the size of the error of rejecting H_0 when H_0 is true. It is usually denoted by α:

$$\alpha = P(\text{Reject } H_0 | H_0 \text{ is true}). \tag{4.1}$$

Although α is frequently chosen to be 0.05, this is not a necessity. Note that there are several statistical tests that depends on the PDF of the tested quantity.

Finally, the process, called hypothesis testing, consists of four steps:

1. State the hypotheses. This involves stating the null and alternative hypotheses. The hypotheses are stated in such a way that they are mutually exclusive. That is, if one is true, the other must be false.

2. Formulate an analysis plan. The analysis plan describes how to use sample data to evaluate the null hypothesis. The evaluation often focuses around a single test statistic.

3. Analyze sample data. Find the value of the test statistic (mean score, proportion, t statistic, z-score, etc.) described in the analysis plan.

4. Interpret results. Apply the decision rule described in the analysis plan. If the value of the test statistic is unlikely, based on the null hypothesis, reject the null hypothesis.

4.2.2 Test of a sample mean

Method

Let $\{x_1, \ldots, x_n\}$ be a random sample drawn from a Gaussian distributed population with a mean μ and a standard deviation σ. Testing the mean implies that we shall use the sample mean $\bar{x}$ to provide some statements on the true mean μ.

Suppose that you suspect that a particular class's performance on a proficiency test is not representative of those people who have taken the test. The national mean score for a student on the test is 558 ± 139. Thus, the null hypothesis is $H_0 : \mu = 558$. Suppose we select a sample of $n = 100$ participants and we record a sample mean equal to 585. Then, there are several alternative hypothesis:

- If you just want to know if the class performance is similar or not to the national score, the alternative hypothesis is $H_a : \mu \neq 558$.
- Suppose we select a sample of $n = 100$ students enrolled in an elite private school and we hypothesise that students at this elite school will score higher than the general population. Then the alternative hypothesis is $H_a : \mu > 558$.

The former case corresponds to a so-called two-tailed test [test bilatéral] and the rejection region is represented in Fig. 4.2 and the latter case is called a upper-tailed test [test unilatéral] represented in Fig. 4.3. Note that the relation between the significance level of the test and the rejection region or the critical value are not the same.

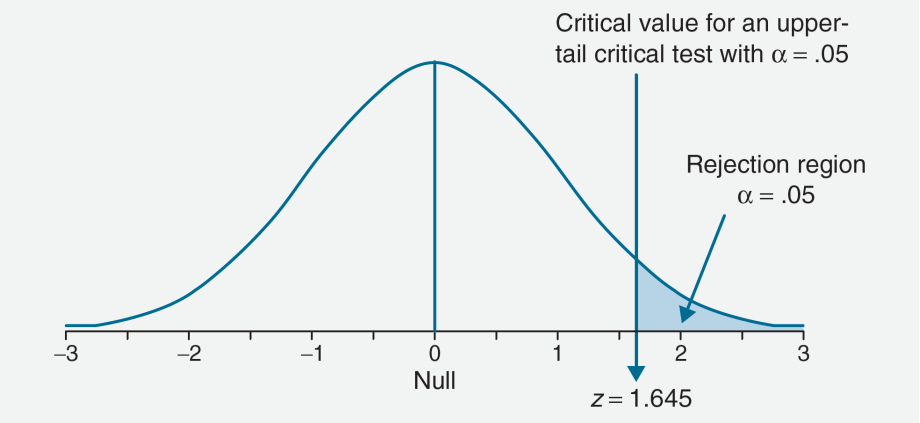


Figure 4.3: Distribution of the values that can be taken by the tested quantity within the H_0 hypothesis. The abscisse quantity is centred on the mean value and divided by the standard quantity.

Then, two other cases must be distinguished: the population variance is known or unknown. If σ is known, testing the hypothesis $H_0 : \mu = \mu_0$ can be accomplished by using the statistic

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}, \tag{4.2}$$

which is distributed according to the standardised normal distribution. If the σ is unknown, we can replace it by the sample variance estimate s_x and use the statistic

$$t = \frac{\bar{x} - \mu_0}{s_x/\sqrt{n}}, \tag{4.3}$$

which is distributed according to the Student's t distribution with $(n-1)$ degrees of freedom.

Example: Class performance against national examination

Let's go back to the class's performance against the national mean score. The national mean score on the test is 558 ± 139 and we record a sample mean equal to 585 for the class with a sample of $n = 100$ students. The null hypothesis is $H_0 : \mu = 558$ and the alternative hypothesis $H_a : \mu \neq 558$.

The level of significance is fixed at $\alpha = 5\%$. In a nondirectional two-tailed test, we divide the alpha value in half so that an equal proportion of area is placed in the upper and lower tail, as in Fig. 4.2. Thus, the critical z-score is equal to $z = 1.96$. This is the critical value for the upper tail of the standard normal distribution. Since the normal distribution is symmetrical, the critical value in the bottom tail will be the same distance below the mean, or $z = -1.96$.

The obtained z-value is

$$z_{obt} = \frac{585 - 558}{139/\sqrt{100}} = 1.94. \tag{4.4}$$

This is less than the critical value; it does not fall in the rejection region. The decision is to retain the null hypothesis. Actually, we should rather say that we cannot reject the null hypothesis.

Suppose now that we select a sample of $n = 100$ students enrolled in an elite private school with the same sample mean equal to 585. The population mean is 558, and we are testing whether the alternative is greater than this value: the alternative hypothesis becomes $H_a : \mu > 558$. We retain the same level of significance, $\alpha = 5\%$. As can be seen in Fig. 4.3, the the critical z-score is equal to $z = 1.645$. As the obtained value ($z_{obt} = 1.94$) is greater than the critical value, the decision is to reject the null hypothesis. The director of the elite private school will be happy to have chosen this alternative hypothesis!

4.2.3 Comparison of two sample means

Method with equal standard deviations

Be two samples $\{x_1, \ldots, x_{n_x}\}$ and $\{y_1, \ldots, y_{n_y}\}$ drawn from normally distributed populations with different means. By comparing the samples we test the null hypothesis $H_0 : \mu_x = \mu_y$. The suitable statistic is

$$t = \frac{\bar{x} - \bar{y} - (\mu_x - \mu_y)}{s_{xy}}, \tag{4.5}$$

where $\bar{x}$ and $\bar{y}$ are the sample means as usual. If the two populations have equal, but unknown standard deviations,

$$s_{xy} = \sqrt{\left(\frac{1}{n_x} + \frac{1}{n_y}\right) \frac{n_x s^2(x) + n_y s^2(y)}{n_x + n_y - 2}}. \tag{4.6}$$

Then t follows Student's distribution with $\nu = n_x + n_y - 2$ degrees of freedom.

Example: systolic blood pressure difference between men and women

Suppose we wish to assess whether there is a statistically significant difference in mean systolic blood pressures between men and women. Data measured on $n = 3\,539$ participants who attended examination are:

	n	$\bar{x}$	$s(x)$
Men	1 623	128.2	17.5
Women	1 911	126.5	9.7

The null hypothesis is H_0: $\mu_m = \mu_w$ and the alternative hypothesis is H_a: $\mu_m \neq \mu_w$. This is a two-tailed test, with an observed value $t_{obs} = 2.66$. As these are two large samples, t-tests and z-tests give similar results. We reject H_0 because $2.66 \geq 1.960$. We find that there is a statistically significant difference in mean systolic blood pressures between men and women

4.3 Bayesian hypothesis testing

The null-hypothesis significance testing (NHST) is the standard inferential tool in many sciences despite its serious flaws. It consists in estimating the conditional probability that indicates the likelihood of an observed result, given that the null hypothesis is correct: $p(D|H_0)$. However, the probability of the null hypothesis being true, given the obtained data, would be expressed as $p(H_0|D)$, which is not equivalent to $p(D|H_0)$. The only way to translate $p(D|H_0)$, which is readily found by applying NHST methods, into $p(H_0|D)$ is by applying the Bayes theorem:

$$p(H_0|D) = \frac{p(D|H_0)p(H_0)}{p(D)}. \tag{4.7}$$

In the examples treated above, the null hypothesis H_0 is tested against the alternative hypothesis $H_a = \bar{H}_0$. Thus, the denominator becomes

$$p(D) = p(D|H_0)p(H_0) + p(D|H_a)p(H_a), \tag{4.8}$$

where the two priors are related by the sum rule $p(H_0) + p(H_a) = 1$. The difficult term is the likelihood function for H_a, $p(D|H_a)$, because, in general, we cannot do predictions with the only given information that H_0 is false. Unlike the null hypothesis, the alternative hypothesis does not specify one particular a priori value for the effect in question. We need either a well-defined alternative or be able to integrate over all possible alternatives when the alternative hypothesis involves an adjustable parameter.

However, the denominator can be ignored if we are interested in the relative merit of H_0 compared to another hypothesis H_1 that allows predictions. Then,

$$\frac{p(H_0|D)}{p(H_1|D)} = \frac{p(D|H_0)}{p(D|H_1)} \times \frac{p(H_0)}{p(H_1)}. \tag{4.9}$$

Then if it is assumed that the prior ratio is equal 1 (i.e., the two hypotheses are deemed equally likely before the data are collected), the posterior ratio is equal to the what is generally called the Bayes factor [facteur de Bayes],

$$B_{0/1} = \frac{p(D|H_0)}{p(D|H_1)}. \tag{4.10}$$

Inserting Eq. (4.8) into Eq. (4.7) gives

$$p(H_0|D) = \frac{p(D|H_0)p(H_0)}{p(D|H_0)p(H_0) + p(D|H_a)p(H_a)} = \frac{B_{0/a}p(H_0)}{B_{0/a}p(H_0) + p(H_a)}. \tag{4.11}$$

If the two hypothesis are assigned the same prior,

$$p(H_0|D) = \frac{B_{0/a}}{B_{0/a} + 1}. \tag{4.12}$$

Thus, when the Bayes factor $B_{0/a} \gg 1$, $p(H_0|D) \to 1$ and the H_0 hypothesis is accepted, whereas when $B_{0/a} \ll 1$, $p(H_0|D) \ll 1$ and the H_0 hypothesis is rejected. Commonly used thresholds to define significance of evidence are given in table 4.1.

Table 4.1: Commonly used thresholds to define significance of evidence. In French, "Extreme" is said [Décisif] and "Moderate" [Substantiel].

Bayes factor	> 100	30-100	10 - 30	3-10	1-3	1
Evidence category for H_0	Extreme	Very strong	Strong	Moderate	Anecdotal	No evidence
Bayes factor	1	1/3 - 1	1/10 - 1/3	1/30 - 1/10	1/100 - 1/30	$< 1/100$
Evidence category for H_a	No evidence	Anecdotal	Moderate	Strong	Very strong	Extreme

Suppose that we want to test the null hypothesis H_0: $x = x_0$ against H_a: $x \neq x_0$, with $x \in \mathbb{R}$. We need a prior distribution for H_a, denoted $p(x)$. Then,

$$B_{0/a} = \frac{p(D|H_0)}{\int p(D|x)p(x)dx}. \tag{4.13}$$

Thus, applying Bayesian hypothesis testing require the calculation of complicated integrals that go beyond the scope of this lecture. Moreover, in general, such integrals have to be carried numerically. More information can be found in textbooks devoted to Bayesian data analysis [13, 14, 15].

4.4 Radioactivity measurements

As already mentioned, this section is based on the international standard ISO 11929:2010 that is under revision. Note that it is partially Bayesian.

4.4.1 Decision threshold

Definition

Suppose that the spectrum obtained by a gamma spectroscopy experiment looks like the one of figure 4.1. Abscissa corresponds to the channel or energy and ordinate to the number of events. Most of the detected events in the region of interest of this figure come from the background. The background effect is defined as "measurement effect caused by the radiation background (for instance, from natural radiation sources)". The net effect is "net effect contribution of the possible radiation of a measurement object (for instance, of a radiation source or a radiation field) to the measurement effect. Finally, the gross effect is "measurement effect caused by the background effect and the net effect".

Let's assume that in the region of interest (in red) the number of events in the background is $n_{bg} = 46$ and the total number is $n_{gross} = 54$. The net result is then $n_{net} = n_{gross} - n_{bg} = 8$. Are these additional 8 events due to a fluctuation of the background or to an additional radioactivity? If n_{net} is very large, it cannot be a fluctuation of the background and we attribute the additional events to an additional radioactivity. We are sure to have a signal. Where is the limit between fluctuations and signal?

This limit defines the decision threshold. Its definition in the international standard [9] reads

> **Decision threshold***: fixed value of the measurand by which, when exceeded by the result of an actual measurement of a measurand quantifying a physical effect, one decides that the physical effect is present.*

Calculation

We shall assume that there is no additional contamination. It is the H_0 hypothesis. As a consequence, the true value of the net number of events is $\mu_{net} = 0$. But, a measurement will give values of n_{net} that are not necessarily equal to zero. As $n_{net} = n_{gross} - n_{bg}$,

$$u^2(\mu_{net} = 0) = u^2(n_{gross}) + u^2(n_{bg}) = n_{gross} + 1 + n_{bg} + 1 = 2(n_{bg} + 1). \tag{4.14}$$

Here, we took into account the Bayesian result that the variance of the inverse Poisson distribution is equal to $n + 1$. As the result is difficult to explain, the international standard claims that we should rule out $n_{gross} = 0$ that leads to variances equal to zero. This is not realistic because the measuring time is finite and we cannot be sure that $n_{gross} = 0$. The standard suggests that one should change the result of all counting from n to $n + 1$. Actually, there is a robust ground for this claim based on the Bayesian inference as explained in the previous chapter. Finally,

$$u^2(\mu_{net} = 0) = 2(n_{bg} + 1). \tag{4.15}$$

Figure 4.4 represents the probability density function $P(\mu_{net} = 0|n_{net})$ as a function of the measured value n_{net}. The plot is similar to the Poisson distribution and to a Gaussian if n_{bg} is large enough, with a standard deviation equal to $\sqrt{2(n_{bg} + 1)}$.

We can therefore test the hypothesis of no contamination. If the result n_{net} is small compared to $u(\mu_{net} = 0)$ the probability that is due to a fluctuation of the background cannot be neglected. On the contrary, if $n_{net} \gg u(\mu_{net} = 0)$ the probability that n_{net} is due to a fluctuation of the background is very small. To define the frontier between these two extreme cases, we fixe a value of the risk α to be wrong when we say that there is a contamination. This value of α is generally small.

Beyond the decision threshold (DT) we conclude that the physical effect is present and we can claim that there is contamination in the present case. The decision threshold is therefore defined with significance level α such as

$$P(n_{net} > DT | \mu_{net} = 0) = \alpha. \tag{4.16}$$

The probability α is also defined as the probability of the error of the first kind. Its value is generally chosen to 2.5% or 5%.

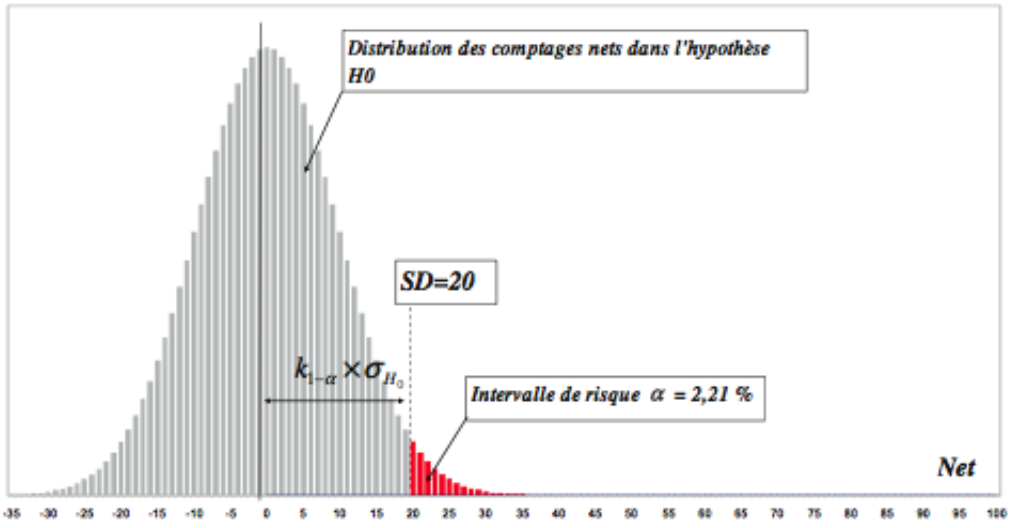


Figure 4.4: Probability to have $\mu_{net} = 0$ as a function of n_{net}. The decision threshold (SD in French) is indicated. (Figure reproduced from [29]).

Finally, with a Poisson statistics, the decision threshold is simply equal to

$$DT = k_{1-\alpha}\sqrt{2(n_{bg} + 1)}, \tag{4.17}$$

where $k_{1-\alpha}$ is the coverage factor corresponding to the probability $1 - \alpha$ for a upper tailed test.

If the Poisson distribution is close to a Gaussian distribution, the relationship between α and $k_{1-\alpha}$ is given by

$$\alpha = \frac{1}{2}\text{erfc}\left[\frac{k_{1-\alpha}}{\sqrt{2}}\right] \qquad \text{with} \qquad \text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^{+\infty} e^{-y^2} dy. \tag{4.18}$$

These formulas are only valid if $n_{bg} \gg 1$. As it is difficult to inverse the error function, one rather uses tables.

Note that if n_{bg} is small, the previous coverage factor is not correct anymore because the Poisson distribution differs from the Gaussian one. For $n_{bg} > 10$, the difference between the DT determined with a Gaussian approximation and the DT determined with a more accurate approach is less than 5%. This is smaller than the relative uncertainty in the DT that is equal to 14% for $n_{bg} = 10$. This is explained in reference [48].

Eventually, one has α chances to get a net value larger than the DT is one measures a sample that is surely not contaminated. In other words, in one measures 100 non-contaminated samples, about α will have a net value larger than DT.

Let's go back to the introductory example with $n_{bg} = 46$. This gives $DT = 20$ with $\alpha = 2.5\%$ ($k_{1-\alpha} = 1.96 \simeq 2$). If $n_{net} > 20$ or if $n_{gross} > 46 + 20 = 66$, the sample is declared contaminated with a probability higher than 97.5%. The larger n_{net} the less chance we have to be wrong by declaring that the sample is contaminated.

Implementation

There are only two possibilities:

First case: If $n_{net} > DT$, we consider that the sample is contaminated. The risk of being wrong is lower than α. The activity of the sample is then given by,

$$A = \frac{n_{net}}{\epsilon t}, \tag{4.19}$$

where t is the counting time and ϵ, the efficiency of the detector.

Second case: If $n_{net} < DT$, the sample might not be contaminated, but we cannot reject the hypothesis that it is contaminated. A very small contamination would be hidden by the fluctuations of the background. The only thing that we can say is that if there is a contamination, it is lower than the so-called detection limit. We will determine it in the next section.

To conclude, here are two notes following the definition of the DT in the standard ISO 11929 [9]:

> *NOTE 1: The decision threshold is defined in a way that the probability of wrongly deciding, that the true value of the measurand is zero, is less than a specified value α.*
> *NOTE 2: If the result is lower than the decision threshold, the result cannot be attributed to the physical effect; nevertheless, it cannot be concluded that the effect is absent.*

4.4.2 Detection limit

Definition

The international standard [9] defines the detection limit [limite de détection] as:

> **Detection limit***: smallest true value of the measurand that ensures a specified probability of being detectable by the measuring method.*

A contamination with a true value $\mu_{net} = DT$ has 50% chances to give a measured value n_{net} lower than the DT. A contamination with a true value slightly higher that the DT also has a probability to have a n_{net} lower than the DT. For such cases, we won't be able to declare that there is a contamination. What we can say is that there might be a contamination that is lower than a value called the detection limit [limite de détection].

Calculation

In the previous section, we evaluated the probability that the true value μ_{net} is equal to zero for various values of n_{net} in order to determine the DT. We will now assume that there is a contamination $\mu_{net} > 0$, corresponding to the hypothesis H_1, and study its distribution as a function of the measured value of n_{net} in order to determine the probability that $\mu_{net} < DT$.

Let us consider figure 4.5. The left Gaussian curve associated to the hypothesis H_0 corresponds to the determination of the DT of the previous section. The right curve, associated to the hypothesis H_1, represents the plot of $Pn_{net})|(\mu_{net}$ as a function of n_{net}. Here, for this value of n_{net}, the risk that $n_{net} < DT$ is β. If μ_{net} were smaller, the risk would be higher and vice versa. Fixing β, the risk to declare the sample non contaminated, i.e of being wrong, we can find the corresponding value of μ_{net} called the detection limit, DL.

The standard states that *"with the decision threshold according to 4.4.1, the detection limit is the smallest true value of the measurand for which the probability of wrongly deciding, that the true value of the measurand is zero, is equal to a specified value β."* β is called the probability of the error of the second kind.

We directly see from figure 4.5 that

$$DL = DT + k_{1-\beta} u(\mu_{net} = DL). \tag{4.20}$$

The problem is that the definition of DL depends DL through $u(\mu_{net} = DL)$. The resolution of such an equation is generally done by iteration. Several cases are shown in the standard.

In the very specific case of the radioactivity satisfying the Poisson statistics, we can exactly solve the problem. As $\mu_{net} = n_{gross} - n_{bg}$, one has

$$u^2(\mu_{net}) = u^2(n_{gross}) + u^2(n_{bg}) = (n_{gross} + 1) + (n_{bg} + 1) = 2(n_{bg} + 1) + \mu_{net}. \tag{4.21}$$

And then, one gets for $\mu_{net} = DL$,

$$DL = DT + k_{1-\beta}\sqrt{2(n_{bg} + 1) + DL}. \tag{4.22}$$

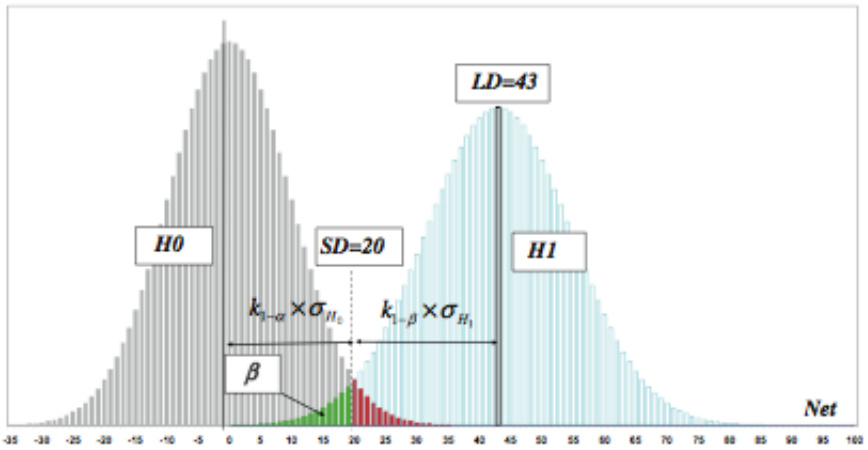


Figure 4.5: Determination of the detection limit. The left Gaussian corresponds to the determination of the decision threshold. (Figure reproduced from [29]). This plot is in French: SD is for "Seuil de Décision" which means decision threshold and LD for "Limite de Détection" which means detection limit.

Fixing $\alpha = \beta$ for the sake of simplicity, one has $k_{1-\alpha} = k_{1-\beta} = k$. Eliminating n_{bg} in eq. (4.22) with the help of the definition of DT, one gets

$$DL = DT + k\sqrt{\frac{DT^2}{k^2} + DL}. \tag{4.23}$$

This can be solved to get

$$DL = 2DT + k^2. \tag{4.24}$$

Such a result is used for radioactivity measurements. As k is of the order of 2, if DT is large, one simply has $DL \simeq 2DT$.

Let us go back to the introductory example. We had $n_{bg} = 46$ and $DT \simeq 20$ with $k = 2$. Therefore, $DL \simeq 44$.

Implementation

It is worth remembering that the detection limit is only used when $n_{net} < DT$. In such a case one writes that the result is lower than the detection limit. The interval $[0, DL]$ corresponds to the confidence interval with a level of confidence equal to $1 - \beta$. The DL is one limit of the confidence interval, not the expected value.

The detection limit is often used to assess the performance of a laboratory.

4.4.3 Calculation of the activity

Results of a measurement are not expressed as a number of counts, but as an activity

$$A = \frac{n_{net}}{\epsilon \cdot t}, \tag{4.25}$$

where t is the counting time and ϵ the efficiency of the detector. Time measurement is generally accurate enough to neglect its uncertainty. This is not the case for the efficiency. How to take into account the uncertainty in the efficiency?

The decision threshold only depends on the background and is not affected by the uncertainties in the efficiency. It is based on the true value $\mu_{net} = 0$ whatever the value of the efficiency of the detector.

As a consequence, the decision threshold in terms of activity reads

$$DT_A = \frac{DT}{t \cdot \epsilon}. \tag{4.26}$$

Indeed, if $n_{net} > DT$, then $\frac{n_{net}}{\epsilon \cdot t} > \frac{DT}{\epsilon \cdot t}$ and then $A > DT_A$.

On the other hand, a large uncertainty in the efficiency affects the detection limit expressed in terms of activity. From eq. (4.25), one immediately gets

$$\frac{u^2(A)}{A^2} = \frac{u^2(n_{net})}{n_{net}^2} + \frac{u^2(\epsilon)}{\epsilon^2} \tag{4.27}$$

and eq. (4.22) is replaced by

$$DL_A = DT_A + k \times u(DL_A). \tag{4.28}$$

Finally, after few lines of calculation, one obtains

$$DL = \frac{k^2 + 2DT}{1 - u_{\%}^2(\epsilon) \times k^2} \qquad \text{and} \qquad DL_A = \frac{DL}{\epsilon \cdot t}, \tag{4.29}$$

where $u_{\%}(\epsilon) = \frac{u(\epsilon)}{\epsilon}$ is the relative uncertainty in the efficiency.

With a large value for this uncertainty, $u_{\%}(\epsilon) = 30\%$, and $\alpha = \beta = 2,5\%$, the detection limit is 1.56 times larger than the detection limit with $u_{\%}(\epsilon) = 0$. In such conditions, $DL \simeq 3\,DT$. However, this uncertainty in the efficiency is too large to be realistic. In general, we can neglect it.

As a conclusion, the decision threshold is not affected by the uncertainty in the efficiency of the detector whereas the detection limit is enlarged by it.

4.4.4 Limits of the confidence interval

If the result is larger than the decision threshold, the physical effect is considered as present. However, if the result is small but with a large uncertainty, the confidence interval can include unrealistic values. This is the case, for example, with $A = (4 \pm 2)$ Bq ($k = 1$). Assuming a Gaussian distribution, 2.5% of the values are negative which is not possible.

As the approach is valid for both n_{net} and A, we shall write y the values of the measurand and $u(y_0)$ its standard uncertainty.

The lower and upper limit of the confidence interval, respectively $y^{\triangleleft}$ and $y^{\triangleright}$, of the measurand, are defined such as the true value of y has a probability $1 - \gamma$ to be included in the interval $[y^{\triangleleft}, y^{\triangleright}]$:

$$\frac{\gamma}{2} = \int_0^{y^{\triangleleft}} f(y)\,dy = \int_{y^{\triangleright}}^{+\infty} f(y)\,dy. \tag{4.30}$$

Here, $f(y)$ stands for the probability density function of y. This is also represented in figure 4.6. Note that the interval is not symmetric.

With a normal distribution, $f(y) = \frac{1}{\sqrt{2\pi}u(y_0)} \exp\left[-\frac{(y-y_0)^2}{2u^2(y)}\right]$, we have

$$\frac{\gamma}{2} = \frac{1}{2}\left(\operatorname{erfc}\left[\frac{-y_0}{\sqrt{2}u(y_0)}\right] - \operatorname{erfc}\left[\frac{y^{\triangleleft} - y_0}{\sqrt{2}u(y_0)}\right]\right) \tag{4.31}$$

$$\frac{\gamma}{2} = \frac{1}{2}\operatorname{erfc}\left[\frac{y^{\triangleright} - y_0}{\sqrt{2}u(y_0)}\right]. \tag{4.32}$$

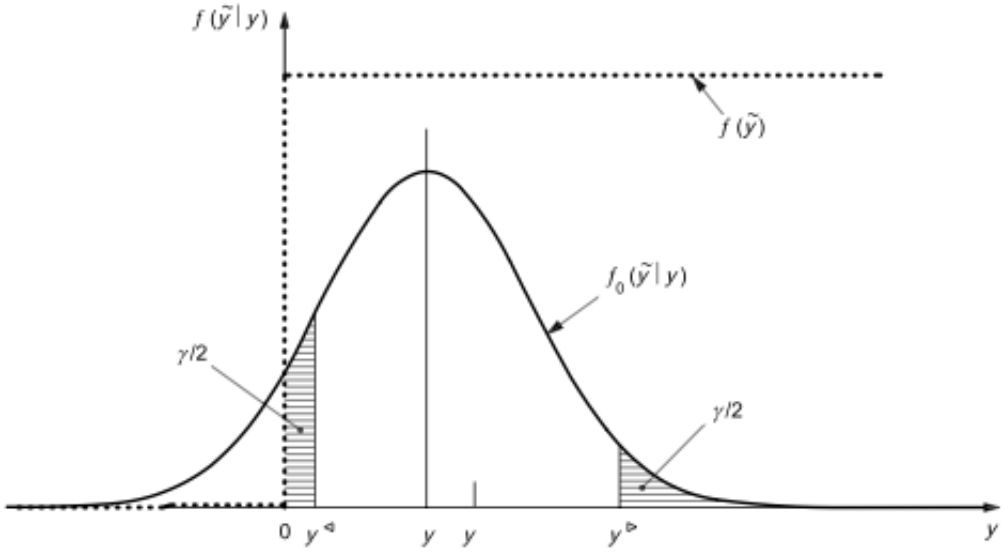


Figure 4.6: Definition of the limits of the confidence interval and the best estimate. (Figure reproduced from the international standard ISO 11929).

These equations are difficult to inverse. The standard gives a table for the coverage factor:

$$y^{\triangleleft} = y_0 - k_p u(y_0) \qquad \text{with} \qquad p = \omega(1-\gamma/2) \tag{4.33}$$

$$y^{\triangleright} = y_0 + k_q u(y_0) \qquad \text{with} \qquad q = 1 - \omega\gamma/2. \tag{4.34}$$

The parameter ω,

$$\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y_0/u(y_0)} \exp\left[-\frac{x^2}{2}\right] dx, \tag{4.35}$$

is also tabulated.

If $y_0 \gg u(y_0)$ ($y_0 > 4u(y_0)$ is enough), $\omega \to 1$ and $p = q = \gamma/2$. One gets a simple result. But for values close to the decision threshold, one should use the above expressions.

As the confidence interval is not symmetric, the mean value is not the most probable value anymore:

$$\bar{y} = y_0 + \frac{u(y_0)}{\omega\sqrt{2\pi}} \cdot \exp\left[-\frac{y_0^2}{2u^2(y_0)}\right], \tag{4.36}$$

with the associated standard uncertainty

$$u(\bar{y}) = \sqrt{u^2(y_0) - (\bar{y} - y_0)\bar{y}}. \tag{4.37}$$

Again, if $y_0 > 4u(y_0)$, $\bar{y} \simeq y_0$ and $u(\bar{y}) \simeq u(y_0)$

Part II

Analysing data

Chapter 5

Least squares fitting

5.1 Historical introduction

To extract information from experimental data, one usually compare them to models or use them to fit parameters from a phenomenological approach. There are statistical methods to fit data. Once the parameters of the models are determined, one has to evaluate the uncertainty in the data and the quality of the fit.

It is very easy to find the equation of a straight ligne that passes through two points. But when one has three, four or more points that are not necessarily well aligned because of the measurement errors or uncertainties, how to do? Facing what Stigler [17] calls a problem of rich man in his book on the history of statistics and the measurement of uncertainty before 1900, physicists studying celestial mechanics or measuring the length of the meridian had far more data than parameters to evaluate.

Actually, this is a very general problem in modelling activity as the model should be simpler than the facts that it pretends to explain. These considerations started in 1686 with Gottfried W. Leibnizs philosophical essay *Discours de métaphysique* (Discourse on Metaphysics), in which he discusses how one can distinguish between facts that can be described by some law and those that are lawless, irregular facts. In addition, as William of Occam noted, given two theories that explain the data, the simpler theory is to be preferred (Occam's razor).Today information theory quantifies information by counting how many bits are needed to encode the information is a possible way to compare the complexity of models and data [31].

There were several attempts to solve the problem of numerous data before Legendre who proposed in 1805 an Appendix on the *Method of the Least Square* in his *Mémoire sur l'orbite des comètes*. But there were no uncertainty in Legendre's work. One has to wait for the theory of probabilities. But what probability density function to choose? Stigler describe several attempts by Laplace. But in 1809, Gauss justifies the method of the least squares that he pretends to have being using since 1795 by using what we now call the Gaussian distribution that was first introduced by De Moivre. Gauss uses Laplace's version of Bayes theorem. He also shows how to solve the problem when it is not linear.

After Laplace found the central limit theorem, he could justify in 1810 the use of the Gaussian distribution and finalize the justification of the least square method with the theory of probabilities. He then applied it to celestial mechanics, geodesy and social sciences. In particular, he could determine the uncertainty of the mass of Saturn and Jupiter as already mentioned.

This chapter is mainly inspired by the book of John R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements* (University Science Books, 1997) [10]. It is an excellent introduction. To go further, I recommend the book of Ifan G. Hughes and Thomas P.A. Hase, *Measurements and their uncertainties, a practival guide to modern error analysis* [12]. And the Bayesian preliminaries to the approach were inspired by the book of Sivia [15].

Note that a supplement to the GUM is under preparation. Please check the homepage of the BIPM.

5.2 Weighted averages

5.2.1 Method

Suppose that the same mesurand is measured by two independent laboratories. Results read

$$x = x_A \pm u_A \quad \text{and} \quad x = x_B \pm u_B \tag{5.1}$$

for the laboratories A and B respectively. How to combine these two results to have a better estimation?

If $|x_a - x_B|$ is larger than $2(u_A + u_B)$, these two results are inconsistent. We shall assume that they are consistent.

Denoting the true value of the measured quantity μ and assuming that the probability distribution functions are Gaussian, the probability density of finding the result of the laboratory A is

$$P(x_A|\mu, u_A) = \frac{1}{\sqrt{2\pi}u_A} \exp\left(-\frac{(x_A-\mu)^2}{2u_A^2}\right). \tag{5.2}$$

Similarly, for the laboratory B. The probability density that the laboratory A gets x_A AND the laboratory B x_B is the product of these two probability densities,

$$P(x_a, x_B|\mu, u_A, u_B) = \frac{1}{2\pi u_A u_B} \exp\left(-\frac{\chi^2}{2}\right) \qquad \text{with} \qquad \chi^2 = \frac{(x_A-\mu)^2}{u_A^2} + \frac{(x_B-\mu)^2}{u_B^2}. \tag{5.3}$$

But to find the most probable value of the rue value μ, we need $P(\mu|x_A, u_A, x_B, u_B)$ that can be obtained with Bayes theorem:

$$P(\mu|x_A, u_A, x_B, u_B) \propto P(\mu) \cdot P(x_a, x_B|\mu, u_A, u_B). \tag{5.4}$$

With a non-informative prior, corresponding to $P(\mu)$ constant, one gets

$$P(\mu|x_A, u_A, x_B, u_B) \propto \exp\left(-\frac{\chi^2}{2}\right). \tag{5.5}$$

The most probable value of μ is also its mean value as the distribution is symmetric. It corresponds to the value that maximises this probability density, i.e. the value that minimises χ^2. Thus, deriving χ^2 with respect to μ leads to

$$x_m = \frac{\frac{x_A}{u_A^2} + \frac{x_B}{u_B^2}}{\frac{1}{u_A^2} + \frac{1}{u_B^2}}. \tag{5.6}$$

It is the weighted average of the results obtained by both laboratories with the weights $w_A = 1/u_A^2$ and $w_A = 1/u_A^2$ respectively,

$$x_m = \frac{w_A\, x_A + w_B\, x_B}{w_A + w_B}. \tag{5.7}$$

This can be easily generalised to N evaluations by independent laboratories:

$$x_m = \frac{\sum_{k=1}^{N} w_k\, x_k}{\sum_{k=1}^{N} w_k} \qquad \text{with} \qquad w_k = \frac{1}{u_k^2}. \tag{5.8}$$

Laboratories with small uncertainties weight more in this average.

Corresponding uncertainty can be easily calculated using the propagation formula that gives,

$$u^2(x_m) = \frac{1}{\sum_{k=1}^{N} w_k}. \tag{5.9}$$

5.2.2 Another approach

We can also use Bayes theorem in an iterative way. Apply it first to the result of laboratory A with a non informative prior, we get a posterior

$$P(\mu|x_A, u_A) \propto \exp\left(-\frac{(x_A-\mu)^2}{2u_A^2}\right). \tag{5.10}$$

This contains the available information after the first measurement. This probability density can then be used as a prior with the result of the laboratory B. Then, Bayes theorem leads to a final posterior similar to the one that we previously got, see equation (5.5).

Then, instead of calculating the most probable value of μ, we could have calculated its mean value and variance, knowing the PDF. If the first prior is non-informative, this is lengthy but easy. For the mean value, we exactly get the same result as the one given equation (5.8) and for the variance, the one given equation (5.9).

5.3 Least squares method

We now want to fit a series of data with a mathematical function. This means adjusting the coefficients or parameters of the function to match the data. We will start with a straight line and then generalise to more complicated cases.

5.3.1 Slope and intercept of a linear function

Let us consider an ensemble of n couples of datum (x_i, y_i) that should be fitted by a straight line. We will consider that the variables x and y are related by a linear relation of the type $y = ax + b$. We want to determine a, b and their uncertainties.

To simplify, we suppose that only the measurements of y suffer uncertainty and the x has no uncertainty, or it is negligible. This is often a reasonable assumption. x can be a number or the time that can be measured with a high accuracy.

If a and b correspond to the true value of the slope and intercept, $ax_i + b$ corresponds then to the true value of y. Assuming that experimental data follow a Gaussian distribution, we have

$$P(y_i|a, b, x_i, u(y_i)) = \frac{1}{\sqrt{2\pi}u(y_i)} \exp\left(-\frac{(y_i - ax_i - b)^2}{2u^2(y_i)}\right). \tag{5.11}$$

And then, il all data supposed to be independent, i.e. if $u(y_i, y_j) = 0$ for all $i \neq j$, we have

$$P(y_1, \ldots, y_n|a, b, x_1, \ldots, x_n, u(y_1), \ldots, u(y_n)) = \frac{1}{\sqrt{(2\pi)^n}\prod_{k=1}^n u(y_k)} \exp\left(-\frac{\chi^2}{2}\right). \tag{5.12}$$

with

$$\chi^2 = \sum_{k=1}^n \frac{(y_k - ax_k - b)^2}{u^2(y_k)} = \sum_{k=1}^n w_k(y_k - ax_k - b)^2 \tag{5.13}$$

using the same weighting factor as before $w_k = 1/u^2(y_k)$.

Eventually, assuming that we have a non-informative uniform prior $P(a, b)$, the Bayes theorem gives

$$\begin{aligned} P(a, b|x_1, \ldots, x_n, y_1, \ldots, y_n, u(y_1), \ldots, u(y_n)) &\propto P(a, b) \cdot P(y_1, \ldots, y_n|a, b, x_1, \ldots, x_n, u(y_1), \ldots, u(y_n)) \\ &\propto \exp\left(-\frac{\chi^2}{2}\right). \end{aligned} \tag{5.14}$$

Estimators of the slope and intercept, denoted $\hat{a}$ and $\hat{b}$ are determined by minimising χ^2. This is why it is called least-squares fitting. Deriving χ^2 with respect to a and b respectively leads to two equations with two

unknown variables,

$$a\sum_{k=1}^{n} w_k\, x_k^2 + b\sum_{k=1}^{n} w_k\, x_k - \sum_{k=1}^{n} w_k\, x_k\, y_k = 0 \tag{5.15}$$

$$a\sum_{k=1}^{n} w_k\, x_k + b\sum_{k=1}^{n} w_k - \sum_{k=1}^{n} w_k\, y_k = 0 \tag{5.16}$$

that are easy to solve. We finally get,

$$\hat{a} = \frac{1}{\Delta}\left(\sum_{k=1}^{n} w_k \cdot \sum_{k=1}^{n} w_k\, x_k\, y_k - \sum_{k=1}^{n} w_k\, x_k \cdot \sum_{k=1}^{n} w_k y_k\right) \tag{5.17}$$

$$\hat{b} = \frac{1}{\Delta}\left(\sum_{k=1}^{n} w_k\, y_k \cdot \sum_{k=1}^{n} w_k\, x_k^2 - \sum_{k=1}^{n} w_k\, x_k\, y_k \cdot \sum_{k=1}^{n} w_k\, x_k\right), \tag{5.18}$$

with

$$\Delta = \sum_{k=1}^{n} w_k\, x_k^2 \cdot \sum_{k=1}^{n} w_k - \left(\sum_{k=1}^{n} w_k\, x_k\right)^2. \tag{5.19}$$

Uncertainties in $\hat{a}$ and $\hat{b}$ and their covariance can be easily determined with the propagation formula:

$$u^2(\hat{a}) = \sum_{k=1}^{n}\left(\frac{\partial \hat{a}}{\partial y_k}\right)^2 u^2(y_k) = \frac{1}{\Delta}\sum_{k=1}^{n} w_k \tag{5.20}$$

$$u^2(\hat{b}) = \sum_{k=1}^{n}\left(\frac{\partial \hat{b}}{\partial y_k}\right)^2 u^2(y_k) = \frac{1}{\Delta}\sum_{k=1}^{n} w_k\, x_k^2 \tag{5.21}$$

$$u(\hat{a},\hat{b}) = \sum_{k=1}^{n}\frac{\partial \hat{a}}{\partial y_k}\frac{\partial \hat{b}}{\partial y_k}u^2(y_k) = -\frac{1}{\Delta}\sum_{k=1}^{n} w_k\, x_k. \tag{5.22}$$

These formulas are easy to compute.

To summarise, the line of best fit depends on the following hypothesis:

1. $u(x_i) = 0\,\forall i$;
2. $u(y_i, y_j) = 0\,\forall i \neq j$;
3. data points follow a Gaussian distribution;
4. the prior is non-informative.

If one of these conditions is not satisfied, we should derive new equations. There are various cases treated in devoted textbooks.

Note that the first least squares method developed by Legendre did not take into account the uncertainty in the experimental data which means that each point had the same weight. This ordinary least squares method is still widely used. Just replace the w_k's by 1 in the above equations to get the formulas.

5.3.2 With an uncertainty in both x_i and y_i

If we want to fit data with a straight line of equation $y = ax + b$ with a set of n couples of data (x_i, y_i) with uncertainties in both the x_i's and the y_i's, most of the books propose the same method as previously, with the following weighting factors,

$$w_k = \frac{1}{u^2(y_k) + a^2 u^2(x_k)}. \tag{5.23}$$

As a is unknown, one proceeds by iterations: one first starts from a reasonable value of a to determine the weights. Using the least squares method we get a new value $\hat{a}$ that is used to determine new weighting factors. We repeat the operations until the value of $\hat{a}$ and $\hat{b}$ converge to fixed values.

If $y = f(x)$, a is replaced by $f'(x_k)$ in the previous method.

5.3.3 Other functions

We can do exactly the same method with polynomial of higher orders. What is important, is the linearity with respect to the coefficients that we want to determine. Similarly, we can easily apply the least squares fitting to functions like

$$y = a\cos x + b\sin x. \tag{5.24}$$

For an exponential function like $y = be^{ax}$ that is very frequent in sciences, we use the logarithm to linearise the problem:

$$\ln(y) = ax + \ln b. \tag{5.25}$$

If the function to be tested is not linear with respect to the coefficients, we cannot solve the problem analytically. A numerical procedure is necessary. We generally start by a Taylor expansion and proceed by iteration. There are several algorithms in the scientific literature that go beyond the scope of this lecture. The article from Donald W. Marquardt [47] is frequently quoted. I.G. Hughes et T.P.A. Hase present other methods in their book [12].

5.4 Goodness of the fit: the χ^2 test

5.4.1 What do the uncertainties include?

Finding the best line that minimise the χ^2 does not mean that the experimental points are well fitted. We can always find such a line with a cloud of points!

Note that uncertainties in the slope and intercept do not take into account the quality of the fit. They just estimate their possible variations within the experimental uncertainties.

Let's consider again the simplest case of three estimates of the same quantity measured by three independent laboratories. The normalised PDF of the value with a non-informative prior is exactly,

$$P(\mu|x_A,u_A,x_B,u_B,x_C,u_C) = \frac{\exp\left(-\frac{\chi^2}{2}\right)}{\int_{-\infty}^{+\infty}\exp\left(-\frac{\chi^2}{2}\right)d\mu} \qquad \text{with} \qquad \chi^2 = \frac{(x_A-\mu)^2}{u_A^2} + \frac{(x_B-\mu)^2}{u_B^2} + \frac{(x_C-\mu)^2}{u_C^2}, \tag{5.26}$$

as we have already seen. It can be easily extended to any number of results.

Let us consider two different situations: the first one corresponds to three compatible results, $x_A = 4.9 \pm 1$, $x_B = 5 \pm 1$ and $x_C = 5.1 \pm 1$ whereas the second case corresponds to three conflicting results, $x_A = 0 \pm 1$, $x_B = 5 \pm 1$ and $x_C = 10 \pm 1$. The probability density function is plotted for both cases in figure 5.1. There is no difference! As a consequence, the mean values and variances are also identical.

This example shows that the procedure does not take into account the quality of the fit. One has to check the goodness of the results or fit by other means. In the second example, the values should be rejected as the coverage intervals of the three points do not overlap. This is a hint that there is a problem: either the measurements are wrong or the model assuming that these three values are estimating the same quantity is wrong. This problem will be addressed in the next chapter dedicated to regression analysis.

5.4.2 Goodness of the fit

If we have to test two mathematical functions, how to choose the one that best fits the data?

To answer to this question, we calculate χ^2_{min} that measures the deviation of the data to the curve. The lower the χ^2_{min}, the better the fit. From which value of χ^2_{min} shall we consider the fit as satisfactory? This depends on the number of terms in the sum or the number of data. If each term of the sum is of the order of 1, χ^2_{min} will be of the order of ν, the number of degrees of freedom. In such a case, we can consider the fit as good. This is not the case if χ^2_{min} is larger than ν.

On the contrary, if χ^2_{min} vanishes or is very small, this suspect. This could be due to fake data or over-fitting. What is over-fitting? Consider that you have 100 points. One can find a polynomial that exactly goes through all data and χ^2_{min} is exactly zero. This is over-fitting because one cannot claim that there is a physical law behind. With another set of data, we would find other coefficients. A good fit should not be affected by intrinsic fluctuations of the data.

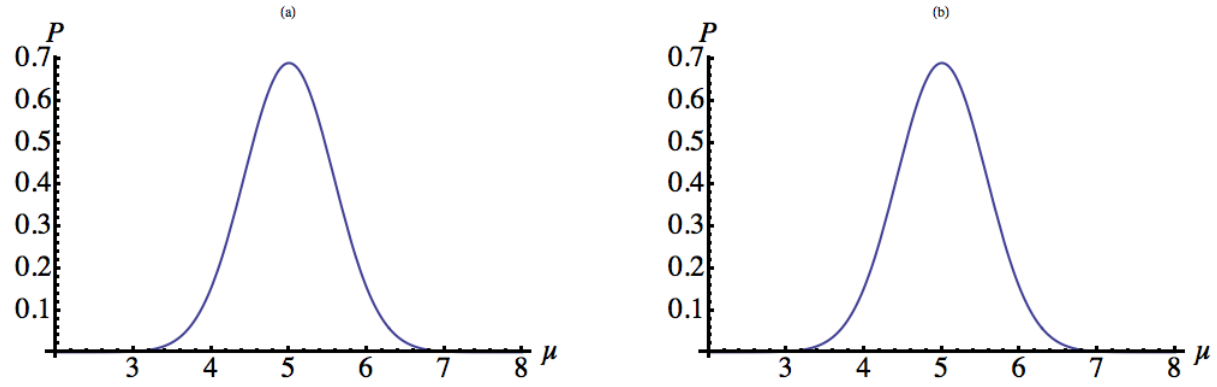


Figure 5.1: Probability density function (pdf) for the mean value after three measurements. In both cases, uncertainty in each value is fixed to one. In the left figure (a), $x_A = 4.9$, $x_B = 5$ and $x_C = 5.1$. For the right figure (b), $x_A = 0$, $x_B = 5$ and $x_C = 10$.

The **reduced chi squared** [chi deux réduit] is defined as $\tilde{\chi}^2_{min} = \chi^2_{min}/\nu$, with ν being the number of degrees of freedom. $\nu = n-2$ in the above example where two parameters a and b were determined. Consequently, $\tilde{\chi}^2_{min}$ should not be large, but it should not be very small neither. $\tilde{\chi}^2_{min}$ is then compared to 1. Statistical studies give the expected range of variation of $\tilde{\chi}^2_{min}$ for a good fit.

5.4.3 The χ^2 distribution

If y is Gaussian random number with a mean value μ and a variance σ^2, $(y-\mu)/\sigma$ is Gaussian random number with a mean value 0 and a variance 1. As a consequence each term of χ^2_{min}, $\frac{y_k - f(x_k)}{\sigma_k}$ is a Gaussian with a mean value 0 and a variance 1, whatever the data and the tested function.

But y^2 is not a Gaussian number. The sum of ν squares of Gaussian random numbers with a mean value 0 and a variance 1 satisfies to the **chi square(d) distribution** [loi de distribution du χ^2]:

$$p(\chi^2_{min}, \nu) = \frac{1}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} [\chi^2_{min}]^{\frac{\nu}{2}-1} \exp\left(-\frac{\chi^2_{min}}{2}\right). \tag{5.27}$$

This fonction is represented in figure 5.2. Its mean value is ν and its variance 2ν. For the reduced chi squared, $\tilde{\chi}^2_{min}$, the mean value is 1 and the variance $2/\nu$.

The probability that χ^2 is larger than the obtained χ^2_{min} is

$$Q = \frac{1}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} \int_{\chi^2_{min}}^{\infty} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}\, dx = \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \int_{\chi^2_{min}/2}^{\infty} t^{\frac{\nu}{2}-1} e^{-t}\, dt = \frac{\Gamma(\nu/2, \chi^2_{min}/2)}{\Gamma(\nu/2)}, \tag{5.28}$$

where $\Gamma(a,x) = \int_x^\infty t^{a-1} e^{-t}\, dt$ is the incomplete Gamma function. Q is tabulated.

As a consequence, if χ^2_{min} is smaller than $\nu - \sqrt{2\nu}$, Q will be close to 1 and the fit is good. On the other hand, if χ^2_{min} is larger than $\nu + \sqrt{2\nu}$, Q will be close to 0 and the fit is bad.

This chi squared distribution is only valid is y_k follow a Gaussian distribution. If it is not the case, the test is less accurate.

There are other chi square tests in the scientific literature to test statistical hypothesis (Is the distribution Gaussian?; Are these two values independent?...). This goes beyond the scope of this lecture.

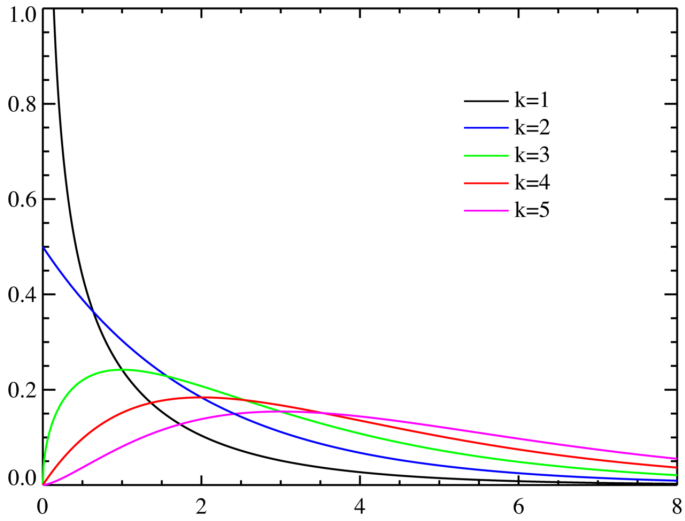


Figure 5.2: Probability density function (pdf) of χ^2 for various numbers of degrees of freedom k. (Figure reproduced from wikipedia.)

5.5 Residual

When we are not sure of the test function that best fits some data points, we can plot the residual to check if there is a hidden structure. The residual is defined as,

$$R_i = \frac{y_i - y(x_i)}{u(y_i)}. \tag{5.29}$$

Figure 5.3 shows two examples with hidden structures that can suggest some physics and lead to a better fit. In both cases the hidden structure is obvious.

These examples were made to show the merit of the test. With real data points, the residual is not that smooth. To find out if there are hidden structures, one can look for correlations between R_i and R_{i-k}. Generally, one uses $k = 1$.

If one plots R_i as a function of R_{i-1}, this should lead to a cloud of points around (0,0) is there is no correlation. Statistically, 91% of the points should be included in the square defined by the limits ± 2. In case of a correlation, points will be closer to the straight line with a slope equal to 1. More points will be out of the square defined by the limits ± 2.

The correlation level can be reduced to a single number,

$$\mathcal{D} = \frac{\sum_{i=2}^{N} [R_i - R_{i-1}]^2}{\sum_{i=1}^{N} R_i^2}. \tag{5.30}$$

This so-called Durbin-Watson number belong to $[0, 4]$. If it is close to 0, residuals are systematically correlated. It it is close to 4, they are systematically anti-correlated. If it is close to 2, they are not correlated.

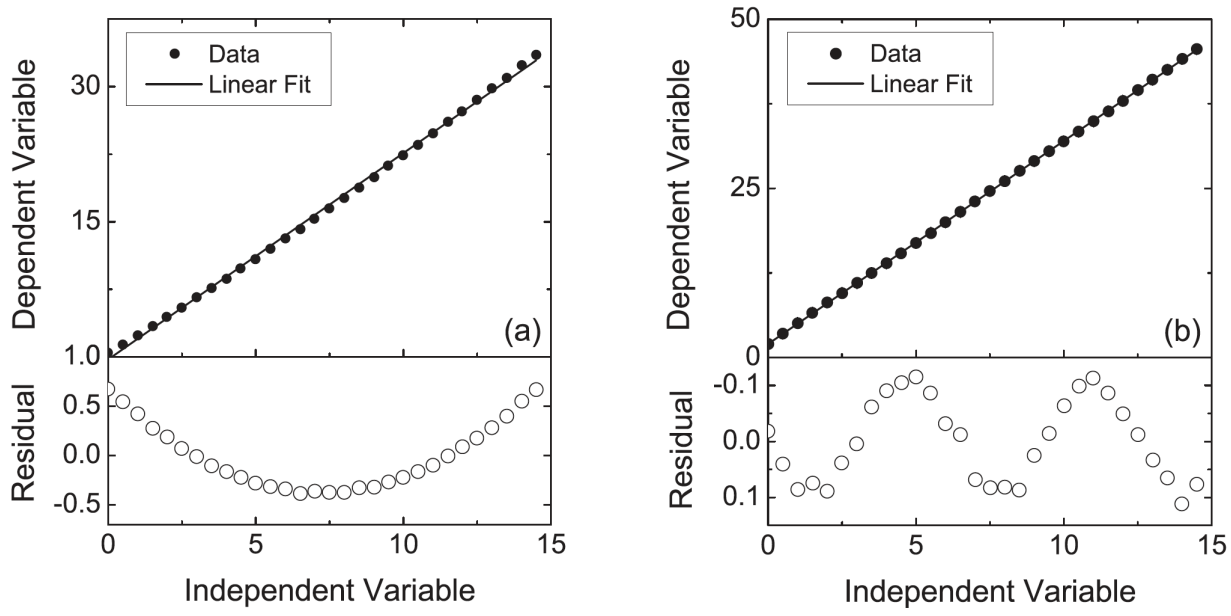


Figure 5.3: Top: Example of a fit with a straight line for two sets of data. Bottom: Residual that shows a structure that is not taken into account by the fit. (Figure reproduced from Ref. [12])

Chapter 6

Ordinary linear regression analysis

We saw in the previous chapter that the uncertainties in the parameters obtained by least squares method do not include the quality of the fit. Regression analysis is a way to address this problem. This chapter is mainly based on two recommended books [20, 21].

The concept of regression comes from Sir Francis Galton during the late 19th century with the publication of *Regression towards mediocrity in hereditary stature* [36]. Galton, one of the most exceptional statisticians of his time, also introduced many important statistical concepts that are now standard in many statistical analyses, including correlation, standard deviation and percentiles to name a few. Francis Galton, was also an explorer. Early in his career, after he inherited a fortune and quit medical school, he went on two expeditions to Africa, the first to the upper Nile Valley and the second through parts of south-west Africa. Based on what he learned from these adventures he wrote two best-selling books which offered practical advice to future explorers and he introduced a new item of camping gear to the Western world: the sleeping bag.

Galton studied the height of humans that varies a lot with individuals. Are these variations purely random or can they be partly explained by external factors? Is there a relationship between the heights of parents and their children? Regression analysis is one of the most widely used statistical tools because it provides simple methods for investigating functional relationships among variables. It has extensive applications in many subject areas. The relationship is expressed in the form of an equation or a model connecting the *response* or *dependent* variable [variable expliquée ou dépendante] and one or more *explanatory* or *predictor* variables [variables explicatives]. Here, the response variable is the height of the children and the explanatory variable, the mean height of the parents. In general, there might be several explanatory variables. We denote the response variable by Y and the set of explanatory variables by $X_1, X_2, \ldots, X_p$. The true relationship between Y and $X_1, X_2, \ldots, X_p$ can be approximated by the regression model

$$Y = f(X_1, X_2, \ldots, X_p) + \epsilon \tag{6.1}$$

In this course, we shall restrict our study to linear regression analysis with a linear relationship model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \epsilon, \tag{6.2}$$

where $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ are $p + 1$ *coefficients* or *regression parameters* to be determined from the data, and ϵ, the part of the variations that is not explained by the variables $X_1, X_2, \ldots, X_p$. We follow the commonly used notational convention of denoting unknown parameters by Greek letters.

A regression equation containing only one explanatory variable is called a *simple regression equation* [régression linéaire simple]. An equation containing more than one predictor variable is called a *multiple regression equation* [régression linéaire multiple]. We shall start with the simplest case in studying simple linear regression between two variables before extending to more variables.

Why is this called "regression"? Investigating the relationship between the heights of parents and their children, Galton plotted the heights of 930 children who had reached adulthood against the mean height of their parents. See Fig. 6.1. To account for differences due to gender he increased female heights by a factor of 1.08. He observed: *"It appeared from these experiments that the offspring did not tend to resemble their parents in size, but always to be more mediocre than they - to be smaller than the parents, if the parents were large; to be larger than the parents, if the parents were small."* Galton called this phenomenon *"regression towards*

mediocrity"; we now call it *"regression towards the mean."* This statistical phenomenon occurs when repeated measurements are made. It means that, in general, relatively high (or low) observations are likely to be followed by less extreme ones nearer the subjects true mean. Regression to the mean remains an important statistical phenomenon that is often neglected and can result in misleading conclusions. The same thing happens if we start with the children. For the children with height between 70 and 71 inches, the mean height of their parents was 69.0 inches. This is a statistical, not a genetic phenomenon.

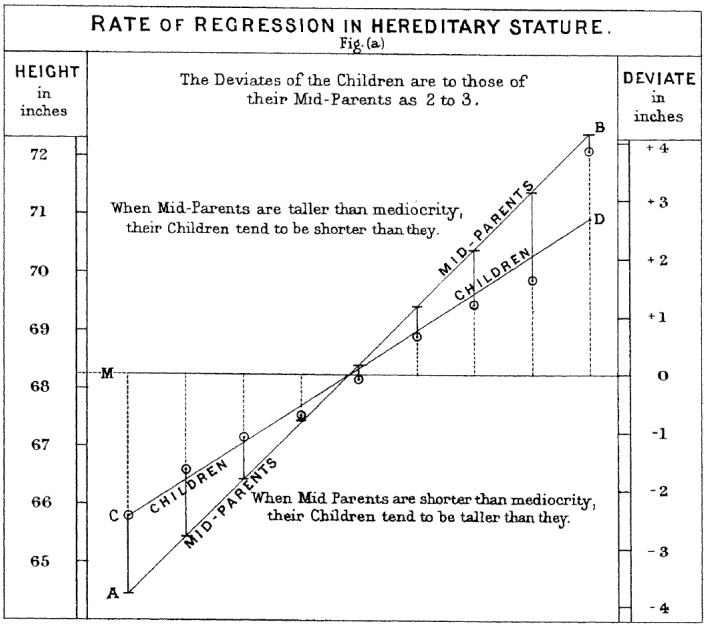


Figure 6.1: Galton's plot reproduced from Ref. [36].

Ironically the term regression, from a Latin root meaning "going back," didn't refer to the statistical procedure Galton used to determine the fit lines for the plotted data points. In fact, Galton didn't even use the least-squares method but just "eyeballed" the data values to draw the fit line. Later, in the 1930's, the term "regression" became associated with the statistical analysis that we now call regression. But it was just by chance that Galton's original results using a fit line happened to show a regression of heights. If his study had showed increasing deviance of children's heights from the average compared to their parents, perhaps we'd be calling it "progression" instead.

6.1 Simple linear regression: relationship between two variables

6.1.1 Deterministic vs stochastic relationships

In some cases, the relationship between two variables X and Y is exact, even if we don't know the coefficients. Considering the case of a simple linear relationship,

$$Y = \beta_0 + \beta_1 X, \tag{6.3}$$

the intercept β_0 and slope β_1 are determined from a set of n paired data points $\{(x_i, y_i), i = 1, \ldots, n\}$ which reflect such a relationship between Y and X, but only approximately. Least-squares method can be used.

But there are also many cases where the relationship between the two variables X and Y is not exact. This is the case for example with the height and weight of individuals. In such cases, we add a new variable ϵ which represents the random error of the model

$$Y = \beta_0 + \beta_1 X + \epsilon. \tag{6.4}$$

This corresponds to regression analysis. It is a probabilistic model that is even used to check whether there is a relationship between X and Y.

Regression analysis is a statistical technique for investigating and modelling the relationship between variables. Applications are numerous; example include engineering, medical, biological and social sciences, physical and chemical sciences, economics, ... The goals of such regression analysis are usually two-fold: to model the relation between Y and X and to predict the expected response value of Y for an arbitrary configuration and estimates its uncertainty.

6.1.2 Coefficient estimation

Based on the available data called $\{(x_i, y_i), i = 1, \ldots, n\}$ we wish to estimate the coefficients β_0 and β_1. This is generally done using the ordinary least square method seen in the previous chapter. For the sake of simplicity, we shall assume that all data points have the same weight as in most textbooks, which means ignoring the experimental uncertainties. As usual in statistics, the estimators [estimateurs] of these coefficients are denoted by $\hat{\beta}_0$ and $\hat{\beta}_1$ (hatted beta zero and hatted beta one).

The estimators of the intercept and slope are obtained by minimising the sum of the squared errors (SSE) [Somme des Carrés des Ecarts (SCE)] with respect to β_0 and β_1. As

$$SSE = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_1)^2, \tag{6.5}$$

the derivation with respect to β_0 gives

$$\left.\frac{\partial SSE}{\partial \beta_0}\right|_{\beta_0=\hat{\beta}_0, \beta_1=\hat{\beta}_1} = -2\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_1) = 0 \qquad \text{or} \qquad \bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}, \tag{6.6}$$

with

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{and} \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

The derivation with respect to β_1 gives

$$\left.\frac{\partial SSE}{\partial \beta_1}\right|_{\beta_0=\hat{\beta}_0, \beta_1=\hat{\beta}_1} = -2\sum_{i=1}^{n} x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_1) = 0. \tag{6.7}$$

Eliminating $\hat{\beta}_0$ with the help of Eq. (6.6), we can easily get an expression for the estimator of the slope [estimateur de la pente]. There are several equivalent formulas in the literature

$$\boxed{\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(X)},} \tag{6.8}$$

It is worth noticing that if there is no correlation between X and Y, the covariance is null, likewise the slope β_1. The estimator of the intercept is then

$$\boxed{\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.} \tag{6.9}$$

The least squares regression line [droite de régression linéaire] is given by

$$\hat{y}(x) = \hat{\beta}_0 + \hat{\beta}_1 x. \tag{6.10}$$

Note that Eq. (6.6) shows that this line goes through the point of coordinates $(\bar{x}, \bar{y})$.

For each observation in the data we can compute

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i, \tag{6.11}$$

that are called fitted values [valeurs estimées]. They allow to estimate the errors ϵ_i

$$\hat{\epsilon}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i, \tag{6.12}$$

which is called ordinary least squares residuals [résidus]. Note that, Eq. (6.6) directly gives

$$\sum_{i=1}^{n} \hat{\epsilon}_i = 0. \tag{6.13}$$

6.1.3 Explained and unexplained deviations

Definitions

If the fit shows that covariance between X and Y is not vanishing, a part of the deviation of Y from the mean can be explained by the explanatory variable X that varies. The residual part remains unexplained. Thus, for any observation, deviation from the mean can be decomposed into the sum of two quantities

$$y_i - \bar{y} = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i). \tag{6.14}$$

The first one, $(\hat{y}_i - \bar{y})$ represents the part explained by the model or the fit, and the second one, $\hat{\epsilon}_i = (y_i - \hat{y}_i)$, the residual that is unexplained by the model or the fit.

Remarkably, such a decomposition is still true with the sum of the squared deviations:

$$\boxed{\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2.} \tag{6.15}$$

This formula is also known as

$$\boxed{SST = SSR + SSE,} \tag{6.16}$$

where SST stands for the Total Sum of the Squared deviations [SCT, Somme des Carrés Totale], SSR denotes the Sum of Squares due to Regression [SCE, Somme des Carrés des variations Expliquées], and SSE represents the Sum of Squared Errors (or residuals) [SCR, Somme des Carrés des Résidus].

To evaluate the proportion of the total variation in Y that is accounted for by the explanatory variables X we define the *coefficient of determination* [coefficient de détermination] as

$$\boxed{R^2 = \frac{\text{Explained variation}}{\text{Total variation}} = \frac{SSR}{SST} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}.} \tag{6.17}$$

R^2, pronounced "R squared", gives us an idea of how the predictor variable X accounts for the response variable Y. It is also a measurement of the goodness of the fit [qualité de la prédiction d'une régression linéaire]. Note that $0 \leq R^2 \leq 1$. If R^2 is near 1, then X accounts for a large part part of the variations of Y.

Example

As an illustrative example, we have randomly generated several sets of 10 paired data points and performed the regression analysis. The y_i's were calculated from an equation $y = 1 + ax$ to which random errors were added. These random errors are normally distributed with a mean value equal to zero and a standard deviation equal to σ. These various cases are plotted in Fig. 6.2.

In the top row of Fig. 6.2, the slope a and the standard deviation in the random contribution to the y_i are identical. The three plots correspond to three different sets of random numbers. The original line is also represented in orange. The coefficients of the regression line and the R^2 vary. They are estimators and their uncertainties need to be evaluated. In the bottom row of Fig. 6.2, the slope a has been either reduced or the standard deviation σ increased. As R^2 is a ratio, decreasing the slope reduced the share of the explained deviations and reduces the determination coefficient. Similarly, increasing the standard deviation increases the share of the unexplained deviations and decreases the determination coefficient. For the last example shown in the panel D, the explanatory variable only explains 40% of the variations of the response variable Y as $R^2 = 0.40$.

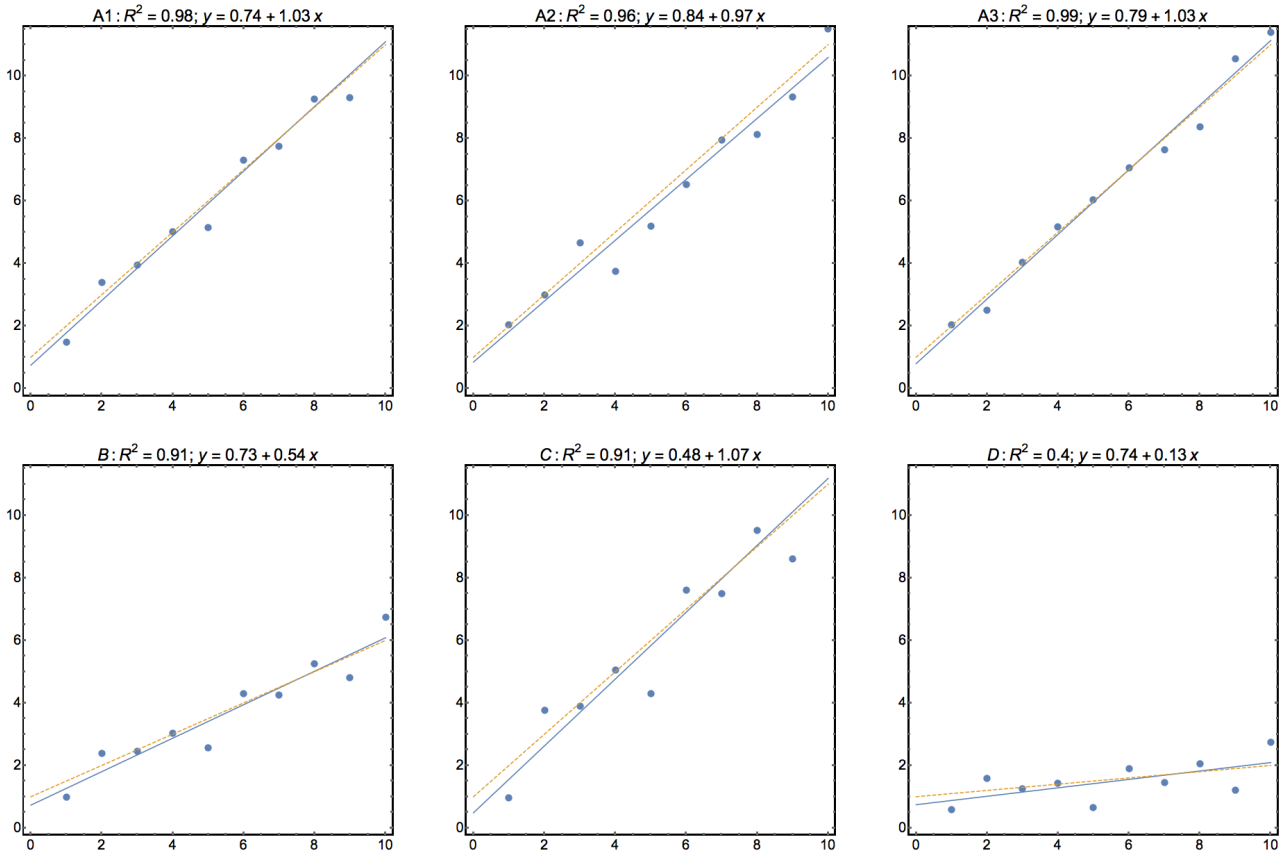


Figure 6.2: Regression lines. For all plots the dashed orange line represents the original equation used to generate the data points and the solid blue line represents the regression line. Corresponding R^2 and the equation of the regression line are indicated.
Top row: Parameters are identical: $a = 1$ and $\sigma = 0.5$. The three plots A1, A2 and A3 correspond to three different draws of data points.
Bottom row: B: $a = 0.5$ and $\sigma = 0.5$ with the same random numbers as A1 - C: $a = 1$ and $\sigma = 1$ with the random numbers of A1 multiplied by 2 - D: $a = 0.1$ and $\sigma = 0.5$ with the same random numbers as A1.
See text for details.

Proof

Here is the proof of Eq. (6.15). As

$$(y_i - \bar{y})^2 = [(\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)]^2 = (\hat{y}_i - \bar{y})^2 + (y_i - \hat{y}_i)^2 + 2(\hat{y}_i - \bar{y})(y_i - \hat{y}_i),$$

we just need to prove that the sum over i of the last term of the r.h.s. vanishes:

$$\begin{aligned}\sum_{i=1}^{n}(\hat{y}_i - \bar{y})(y_i - \hat{y}_i) &= \hat{\beta}_1 \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \hat{y}_i) = \hat{\beta}_1 \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y} + \bar{y} - \hat{y}_i) \\ &= \hat{\beta}_1 \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) - \hat{\beta}_1 \sum_{i=1}^{n}(x_i - \bar{x})(\hat{y}_i - \bar{y})\end{aligned}$$

Using the definition of $\hat{\beta}_1$ given in Eq. (6.8) for the first term of the r.h.s we get

$$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})(y_i - \hat{y}_i) = \hat{\beta}_1^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 - \hat{\beta}_1^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 = 0.$$

No specific assumption was done up to this stage.

6.1.4 Predictions and uncertainties

Regression models predict a value of the Y variable given known values of the X variables. Prediction within the range of values in the dataset used for model-fitting is known informally as interpolation. Prediction outside this range of the data is known as extrapolation. Performing extrapolation relies strongly on the regression assumptions. The further the extrapolation goes outside the data, the more room there is for the model to fail due to differences between the assumptions and the sample data or the true values.

It is generally advised that when performing predictions, one should accompany the estimated value of the dependent variable with a prediction interval that represents the uncertainty. Such intervals is expected to expand rapidly as the values of the independent variable) moved outside the range covered by the observed data.

Adding a point x, what would be the corresponding y? It is natural to take the value given by the estimated coefficients:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \epsilon. \tag{6.18}$$

Uncertainties come from $\hat{\beta}_0, \hat{\beta}_1$ and ϵ.

Method and assumptions

As shown in the plots A1, A2 and A3 of Fig. 6.2, the coefficients of the regression line vary when we select another set of data points obtained in the same conditions. To evaluate their standard deviation we should repeat the trials. The only parameter that changes between the cases A1 to A3 is the error. To repeat the trials, we should do some assumptions on this random error.

In usual regression analysis, it is assumed that for every fixed value x_i, the ϵ_i's are independent random quantities normally distributed with mean zero and a common variance σ^2:

$$\mathrm{E}(\epsilon_i) = 0, \tag{6.19}$$

$$\mathrm{Cov}(\epsilon_i, \epsilon_j) = \delta_{i,j}\sigma^2, \tag{6.20}$$

where $\delta_{i,j}$ is the Kronecker symbol. Here $\mathrm{E}(\epsilon_i)$ denotes the mathematical expectation, i.e. the average over the cases A1, A2... to infinity and $\mathrm{Cov}(\epsilon_i, \epsilon_j)$ the covariance over the same cases. In statistics, a sequence or a vector of random variables is said to be homoscedastic [homoscédastique] if all random variables in the sequence or vector have the same finite variance. This is also known as homoscedasticity [homoscédasticité] or homogeneity of variance. The spelling homoskedasticity is also used. Finally, the assumed PDF for the errors reads

$$P(\epsilon_1, \ldots, \epsilon_n) = \frac{1}{\sqrt{(2\pi)^n}\sigma^n} \exp\left[-\frac{\epsilon_1^2 + \ldots + \epsilon_n^2}{2\sigma^2}\right]. \tag{6.21}$$

This is a strong assumption that is not always verified in practice but allows to get analytical formulas.

Expectation and variance of $\hat{\beta}_1$

To calculate the mean and variance of the estimated slope $\hat{\beta}_1$ over the trials A1 to infinity, we shall use the following expression

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}, \tag{6.22}$$

in which y_i is the only random variable as the x_i's are identical over the trials. Thus, for the mathematical expectation,

$$\mathrm{E}(\hat{\beta}_1) = \frac{\sum_{i=1}^n (x_i - \bar{x})\,\mathrm{E}(y_i)}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

with $\mathrm{E}(y_i) = \beta_0 + \beta_1 x_i + \mathrm{E}(\epsilon_i) = \beta_0 + \beta_1 x_i$. Then,

$$\mathrm{E}(\hat{\beta}_1) = \frac{\sum_{i=1}^n (x_i - \bar{x})(\beta_0 + \beta_1 x_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\beta_0 \sum_{i=1}^n (x_i - \bar{x}) + \beta_1 \sum_{i=1}^n x_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

As $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$, one finally has

$$\mathrm{E}(\beta_1) = \beta_1 \frac{\sum_{i=1}^{n}(x_i - \bar{x} + \bar{x})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \beta_1.$$

The fact that

$$\boxed{\mathrm{E}(\hat{\beta}_1) = \beta_1} \tag{6.23}$$

means that $\hat{\beta}_1$ is an unbiased estimator of β_1.

Similarly, noticing that $\mathrm{Var}(\lambda\epsilon_i) = \lambda^2 \mathrm{Var}(\epsilon_i)$ and that $\mathrm{Var}(\epsilon_i + \epsilon_j) = \mathrm{Var}(\epsilon_i) + \mathrm{Var}(\epsilon_j)$ if $i \neq j$,

$$\mathrm{Var}(\hat{\beta}_1) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \mathrm{Var}(y_i)}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2\sigma^2}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^2}.$$

Finally,

$$\boxed{\mathrm{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}.} \tag{6.24}$$

This variance is adopted for the uncertainty in $\hat{\beta}_1$.

As usual in statistics, we would have expected that $\mathrm{Var}(\hat{\beta}_1) \propto \frac{1}{n}$. This is not explicitly the case, but as the denominator contains n terms and the numerator one, the larger the number of points the better. Note that the larger the deviations of the points x_i to mean $\bar{x}$, the better.

This variance can be used to test the hypothesis $H_0 : \beta_1 = 0$ which means that Y might not depend on X. The problem is that we don't know σ^2. We shall evaluate it later.

Expectation and variance of $\hat{\beta}_0$

We can proceed similarly for the estimated intercept

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.$$

This leads to

$$\boxed{\mathrm{Var}(\hat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right].} \tag{6.25}$$

To show this formula, we shall start by the mathematical expectation,

$$\mathrm{E}(\hat{\beta}_0) = \mathrm{E}(\bar{y}) - \bar{x}\,\mathrm{E}(\hat{\beta}_1) = \mathrm{E}(\bar{y}) - \bar{x}\beta_1.$$

As

$$\mathrm{E}(\bar{y}) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(y_i) = \frac{1}{n}\sum_{i=1}^{n}(\beta_0 + \beta_1 x_i + \mathrm{E}(\epsilon_i)) = \frac{1}{n}\sum_{i=1}^{n}(\beta_0 + \beta_1 x_i) = \beta_0 + \beta_1\bar{x},$$

we finally have

$$\mathrm{E}(\hat{\beta}_0) = \beta_0. \tag{6.26}$$

Thus, $\hat{\beta}_0$ is also an unbiased estimator of β_0.

Regarding the variance,

$$\mathrm{Var}(\hat{\beta}_0) = \mathrm{Var}(\bar{y}) + \bar{x}^2\,\mathrm{Var}(\hat{\beta}_1) - 2\bar{x}\,\mathrm{Cov}(\bar{y}, \hat{\beta}_1),$$

with $\mathrm{Var}(\bar{y}) = \frac{\sigma^2}{n}$, as $\mathrm{Var}(y_i) = \sigma^2\,\forall i$, and

$$\mathrm{Cov}(\bar{y}, \hat{\beta}_1) = \mathrm{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} y_i, \frac{\sum_{j=1}^{n}(x_j - \bar{x})y_j}{\sum_{j=1}^{n}(x_i - \bar{x})^2}\right) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(x_j - \bar{x})\,\mathrm{Cov}(y_i, y_j)}{n\sum_{j=1}^{n}(x_j - \bar{x})^2}.$$

As

$$\mathrm{Cov}(y_i, y_j) = \mathrm{Cov}(\epsilon_i, \epsilon_j) = \sigma^2\delta_{i,j},$$

we finally have that

$$\mathrm{Cov}(\bar{y}, \hat{\beta}_1) = \frac{\sigma^2 \sum_{i=1}^n (x_i - \bar{x})}{n \sum_{j=1}^n (x_j - \bar{x})^2} = 0.$$

Thus,

$$\mathrm{Var}(\hat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right].$$

This variance is adopted for the uncertainty in $\hat{\beta}_0$.

Covariance

As the two estimators of the regression coefficients depend on the same input quantity, they should be correlated. Their covariance reads

$$\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_2) = \mathrm{Cov}\left(\hat{\beta}_1, \bar{y} - \hat{\beta}_1 \bar{x}\right) = \mathrm{Cov}(\hat{\beta}_1, \bar{y}) - \bar{x}\,\mathrm{Var}(\hat{\beta}_1) = 0 - \bar{x}\,\mathrm{Var}(\hat{\beta}_1).$$

Finally,

$$\boxed{\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_2) = \frac{-\bar{x}\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot}. \tag{6.27}$$

Variance of $\hat{y}$

For any value x of X we can predict that the value taken by Y is

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x.$$

The uncertainty in the predicted value by the model is given by

$$\boxed{\mathrm{Var}(\hat{\mu}_y) = \sigma^2 \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right).} \tag{6.28}$$

This expression stands for the mean of the predicted values. To the previous expression, we could add the uncertainty in the error ϵ which gives

$$\boxed{\mathrm{Var}(\hat{y}) = \sigma^2 \left(1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right).} \tag{6.29}$$

This variance is adopted for the uncertainty in all the predicted $\hat{y}$ for a given x.

To show these relations, we shall calculate,

$$\mathrm{Var}(\hat{y}) = \mathrm{Var}(\hat{\beta}_0) + x^2\,\mathrm{Var}(\hat{\beta}_1) + 2x\,\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1).$$

Replacing with the expressions obtained before, we get

$$\mathrm{Var}(\hat{y}) = \frac{\sigma^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})} + x^2 \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} - \frac{2x\bar{x}\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sigma^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \left(\sum_{i=1}^n x_i^2 + nx^2 - 2nx\bar{x} \right).$$

As

$$\sum_{i=1}^n x_i^2 + nx^2 - 2nx\bar{x} = \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) + \left(nx^2 - 2nx\bar{x} + n\bar{x}^2 \right) = \sum_{i=1}^n (x_i - \bar{x})^2 + n(x - \bar{x})^2,$$

we finally get

$$\mathrm{Var}(\hat{y}) = \sigma^2 \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right).$$

This expression stands for the mean of the predicted values. To the previous expression, we should add the uncertainty in the error ϵ which gives

$$\mathrm{Var}(\hat{y}) = \sigma^2 \left(1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right).$$

This variance is adopted for the uncertainty in the predicted $\hat{y}$ plus error.

Note that the further the new point x from the mean value of the points $\bar{x}$, the larger the uncertainty in the prediction $\hat{y}$.

Estimation of the variance in the errors

Remarkably, all the previous results are proportional to the variance in the errors σ^2 that is unknown because the errors are unknown. We have to use the estimated residues:

$$\boxed{\hat{\sigma}^2 = s^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 = \frac{SSE}{n-2}.} \tag{6.30}$$

The denominator $n-2$ should be understood as n, the total number of data points, minus 2, the number of coefficients that are determined with the fit. Thus, $n-2$ corresponds to the number of degrees of freedom. We can use this result in all the variances and covariance calculated before.

To find this unbiased estimator, we have to calculate $\mathrm{E}(\sum_{i=1}^{n}\hat{\epsilon}_i^2)$.

First,

$$\sum_{i=1}^{n}\hat{\epsilon}_i^2 = \sum_{i=1}^{n}(y_i-\hat{y}_i)^2 = \sum_{i=1}^{n}y_i^2 + \sum_{i=1}^{n}\hat{y}_i^2 - 2\sum_{i=1}^{n}y_i\hat{y}_i, \tag{6.31}$$

with

$$\hat{y}_i = \bar{y} + \hat{\beta}_1(x_i-\bar{x}).$$

Then

$$\sum_{i=1}^{n}\hat{y}_i^2 = \sum_{i=1}^{n}\bar{y}^2 + \hat{\beta}_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2 + 2\bar{y}\hat{\beta}_1\sum_{i=1}^{n}(x_i-\bar{x}) = n\bar{y}^2 + \hat{\beta}_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2 + 0.$$

From the expression (6.22) for $\hat{\beta}_1$, we can say that

$$\hat{\beta}_1\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{i=1}^{n}y_i(x_i-\bar{x}).$$

Thus,

$$\sum_{i=1}^{n}\hat{y}_i^2 = n\bar{y}^2 + \hat{\beta}_1\sum_{i=1}^{n}y_i(x_i-\bar{x}) = \bar{y}\sum_{i=1}^{n}y_i + \sum_{i=1}^{n}y_i(x_i-\bar{x}) = \sum_{i=1}^{n}y_i[\bar{y}+\hat{\beta}_1(x_i-\bar{x})] = \sum_{i=1}^{n}y_i\hat{y}_i.$$

Back to Eq. (6.31), we have

$$\sum_{i=1}^{n}\hat{\epsilon}_i^2 = \sum_{i=1}^{n}y_i^2 - \sum_{i=1}^{n}\hat{y}_i^2.$$

Now we can calculate the mathematical expectation,

$$\mathrm{E}(\sum_{i=1}^{n}\hat{\epsilon}_i^2) = \sum_{i=1}^{n}\mathrm{E}(y_i^2) - \sum_{i=1}^{n}\mathrm{E}(\hat{y}_i^2) = \sum_{i=1}^{n}[\mathrm{Var}(y_i) + \mathrm{E}^2(y_i)] - \sum_{i=1}^{n}[\mathrm{Var}(\hat{y}_i) + \mathrm{E}^2(\hat{y}_i)],$$

as for any quantity $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X)$. Let us check each term. As,

$$\mathrm{E}(\hat{y}_i) = \bar{y} + \mathrm{E}(\hat{\beta}_1)(x_i-\bar{x}) = \bar{y} + \beta_1(x_i-\bar{x}) = \mathrm{E}(y_i),$$

two terms cancel out. Then, $\mathrm{Var}(y_i) = \sigma^2$ and $\mathrm{Var}(\hat{y}_i)$ is given Eq. (6.28). Thus,

$$\mathrm{E}(\sum_{i=1}^{n} \hat{\epsilon}_i^2) = (n-2)\sigma^2. \tag{6.32}$$

Finally, we found that the unbiased estimator of σ^2 is

$$\hat{\sigma}^2 = s^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{SSE}{n-2}.$$

Inference about the slope

A formal way of measuring the usefulness of the quantity X as an explanatory variable of the quantity Y is to conduct a test of hypothesis about the slope β_1. Note that the null hypothesis $\beta_1 = 0$ means that there is no linear relationship between Y and X. To perform the test, we need to know the PDF of the slope. Let us first summarise the previous results.

We have assumed that the errors ϵ_i are independent random quantities normally distributed with mean zero and a common variance σ^2. With this assumption, we showed that $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased estimates of β_0 and β_1, respectively. Their variances are given in Eqs. (6.25) and (6.24) respectively, and their covariance in Eq. (6.27). The distributions of $\hat{\beta}_0$ and $\hat{\beta}_1$ are Gaussian with means β_0 and β_1 and variances as just mentioned.

However, these variances depend on the unknown parameter σ^2 that is estimated from the data. An unbiased estimator of σ^2 is given in Eq. (6.30). Replacing σ^2 by $\hat{\sigma}^2$ in Eqs (6.25) and (6.24), we get unbiased estimates of the variances of β_0 and β_1. Thus, the standard uncertainty in $\hat{\beta}_0$ and $\hat{\beta}_1$ are

$$s(\hat{\beta}_0) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}, \qquad \text{and} \tag{6.33}$$

$$s(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}, \tag{6.34}$$

respectively. As we know, the standard uncertainty in $\hat{\beta}_1$ is a measure of how precisely the slope has been estimated.

We can show that the quantity

$$t_1 = \frac{\hat{\beta}_1 - \beta_1}{s(\hat{\beta}_1)}$$

is distributed as a Student's t with $n-2$ degrees of freedom. Thus, the confidence interval at the level $1-\alpha$ for the slope is

$$[\hat{\beta}_1 - t_{(\alpha/2,n-2)}\, s(\hat{\beta}_1); \hat{\beta}_1 + t_{(\alpha/2,n-2)}\, s(\hat{\beta}_1)]. \tag{6.35}$$

The hypothesis $\beta_1 = 0$ is rejected if this value is not included in the above interval. This would mean that β_1 is likely to be different from 0, and hence the explanatory variable X is a statistically significant explanatory variable of the response variable Y.

We can do a similar statistical analysis of the intercept β_0, but conclusions will differ, of course.

Let us consider again the cases plotted in Fig. 6.2. For the case A1, the 90% confidence interval for $\hat{\beta}_1$ is $[0.926, 1.144]$. It contains the true value $\beta_1 = 1$ and there is no doubt that X is an explanatory variable of Y. For the case D which had a $R^2 = 0.40$, the 90% confidence interval for $\hat{\beta}_1$ is $[0.026, 0.245]$. It contains the true value $\beta_1 = 0.1$. Note that the 95% confidence interval is $[-0.000\,08, 0.269\,82]$ and contains 0. This case is just at the frontier of statistic significance for the explanatory variable. How to improve the data to be able to answer to the test? One way is to increase the number of data points. Another way is to increase the share of the explained contribution by considering, if possible, a larger range of points to have Δy larger than the errors ϵ_i. With 10 points with x_i's regularly spaced from 10 to 100 instead of 1 to 10 and the same errors as case D, the $R^2 = 0.98$ and the 95% confidence interval for $\hat{\beta}_1$ becomes $[0.090, 0.117]$. It excludes zero without any ambiguity.

Note that the 5% threshold has always been an arbitrary cut-off, proposed by statistician Ronald Fisher in 1925. A chance of 5.1% is not objectively more meaningful than a chance of 4.9%. Fisher acknowledged that it was arbitrary, and he emphasised that it should be used carefully and in conjunction with other information to reach a sensible conclusion.

Predictions and confidence intervals

The fitted regression equation can be used for predictions. There are two kinds of predictions:

1. The prediction of the value of the response variable Y which corresponds to a chosen value x_0 of the explanatory variable X: $\hat{y}(x_0)$.
2. The estimation of the mean response, when $X = x_0$: $\hat{\mu}_Y(x_0)$.

Both predicted values coincide

$$\hat{\mu}_Y(x_0) = \hat{y}(x_0) = \hat{\beta}_0 + \hat{\beta}_1 x_0, \tag{6.36}$$

but the standard deviations differ, see Eqs (6.28) and (6.29). They can be estimated by

$$s^2(\hat{\mu}_Y(x_0)) = \hat{\sigma}^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) \quad \text{and} \tag{6.37}$$

$$s^2(\hat{y}(x_0)) = \hat{\sigma}^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right). \tag{6.38}$$

The standard uncertainty of $\hat{\mu}_Y(x_0)$ is smaller than the standard uncertainty in $\hat{y}(x_0)$. Intuitively, this makes sense. The averaging reduces the variability and uncertainty in the associated estimate.

For both cases Student law with $(n-2)$ degrees of freedom applies. Hence, the confidence interval for the predicted value with confidence level $(1-\alpha)$ is given by

$$[\hat{y}(x_0) - t_{(\alpha/2;n-2)}\, s^2(\hat{y}(x_0)); \hat{y}(x_0) + t_{(\alpha/2;n-2)}\, s^2(\hat{y}(x_0))], \tag{6.39}$$

and the confidence interval for the mean of the predicted values with confidence level $(1-\alpha)$ is given by

$$[\hat{\mu}_Y(x_0) - t_{(\alpha/2;n-2)}\, s^2(\hat{\mu}_Y(x_0)); \hat{\mu}_Y(x_0) + t_{(\alpha/2;n-2)}\, s^2(\hat{\mu}_Y(x_0))]. \tag{6.40}$$

To distinguish between these two cases, the limits (interval) in (6.39) are sometimes referred to as *prediction* or *forecast limits* (interval) [intervalle de prédiction], whereas the limits (interval) given in (6.40) are called *confidence limits* (interval) [intervalle de confiance].

As already mentioned, these two standard uncertainties increase the farther the value of the explanatory variable from the centre of the actual observations. There is another danger in such predictions: the linear relationship that has been estimated may not hold outside the range of observations. Therefore, care should be taken in employing fitted regression lines for predictions far outside the range of observations.

Moreover, it is worth mentioning that the prediction interval does not vanish when the number of observations n increases.

6.1.5 Models with a single parameter

When hypothesis $\beta_0 = 0$ or $\beta_1 = 0$ are not rejected, we might consider a model with a single coefficient. There are naturally two cases.

Regression line through the origin

Sometimes, it may be necessary to fit the model

$$Y = \beta_1 X + \epsilon, \tag{6.41}$$

a line passing through the origin. Such a model is also called the *no-intercept* model [modèle sans constante]. The line may be forced to go through the origin because of subject matter theory or other scientific considerations. We shall not do again the whole demonstration, but just give the main results.

The least-squares estimate of β_1 is

$$\boxed{\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}.} \tag{6.42}$$

The fitted line does not go through the point $(\bar{x}, \bar{y})$. Consequently,

$$\sum_{i=1}^{n} y_i \neq \sum_{i=1}^{n} \hat{y}_i,$$

and

$$SST \neq SSE + SSR.$$

For this reason, some quality measures for models such as R^2 are no longer appropriate for models with no intercepts.

We can show that $\hat{\beta}_1$ is an unbiased estimator of β_1 with a variance

$$\mathrm{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}.$$

Here σ^2 stands for the common variance of the errors, the ϵ_i's. We can show that an unbiased estimator of σ^2 is

$$\boxed{\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-1} = \frac{SSE}{n-1}.} \tag{6.43}$$

Note that the number of degrees of freedom for SSE is $(n-1)$ and not $(n-2)$ anymore. Thus, the standard uncertainty in the estimator of the slope, $\hat{\beta}_1$, is

$$\boxed{s^2(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{\sum_{i=1}^{n} x_i^2}.} \tag{6.44}$$

Confidence intervals should be determined with the Student law for $(n-1)$ degrees of freedom.

Users of regression models through the origin should be careful when using computer software programs because some of them apply the formulas for model with intercept and give incorrect results for the case of regression models through the origin.

Model without explanatory variable

In this section, we treat the trivial example of a regression model without any explanatory variable

$$Y = \beta_0 + \epsilon.$$

Such a case arrises when we wish to test for the mean of a single variable Y based on a random sample of n observations $y_1, \ldots, y_n$. As we already know, the least squares method gives

$$\hat{\beta}_0 = \frac{1}{n} \sum_{i=1}^{n} y_i = \bar{y}. \tag{6.45}$$

With the same hypothesis as before for the errors, we find naturally that the Sum of Squares due to Regression $SSR = 0$. An unbiased estimator in the variance of the errors is

$$\hat{\sigma}^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2, \tag{6.46}$$

as for the type A uncertainties. Thus,

$$s^2(\hat{\beta}_0) = \frac{s^2}{n} = \frac{1}{n(n-1)} \sum_{i=1}^{n} (y_i - \bar{y})^2. \tag{6.47}$$

Once again, confidence intervals are determined with the Student law with $(n-1)$ degrees of freedom.

6.2 Multiple linear regression

Some response quantities might depend on several explanatory quantities. A typical case in nuclear physics is the Liquid Drop Model formula for the binding energy. Several coefficients are fitted on experimental data. The relationship between the response quantity Y and the p predictors or explanatory variables, $X_1, X_2, \ldots, X_p$ is formulated as a linear model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \epsilon, \tag{6.48}$$

where $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ are the model regression coefficients and ϵ is a random error. They are determined from a set of n observations of the response quantity y_i and the p explanatory quantities x_{i1} to x_{ip}.

Multiple linear regression is a generalisation of simple linear regression but formulas are a bit heavy to handle. Therefore, we shall present the standard results of multiple regression analysis in matrix notation.

6.2.1 Least squares method with matrix notation

The n equations

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \epsilon_i,$$

are replaced by a single matrix expression

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1p} \\ 1 & x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{np} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}, \tag{6.49}$$

which is shorten as

$$[y] = [[x]] \cdot [\beta] + [\epsilon]. \tag{6.50}$$

Here $[y]$ is an n vector, $[[x]]$ an $n \times (p+1)$ matrix, $[\beta]$ a $p+1$ vector and $[\epsilon]$ an n-vector.

The least squares estimator $[\hat{\beta}]$ of $[\beta]$ is obtained by minimising

$$SSE = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{k=1}^{p} \beta_k x_{ik} \right)^2 \tag{6.51}$$

$$= [\epsilon]^T \cdot [\epsilon] = ([y] - [[x]] \cdot [\beta])^T \cdot ([y] - [[x]] \cdot [\beta]), \tag{6.52}$$

that is the sum of squared deviations of the observations from their expected values. Minimisation leads to the system of equations

$$[[x]]^T \cdot [[x]] \cdot [\hat{\beta}] = [[x]]^T \cdot [y].$$

Thus, the least squares estimate $[\hat{\beta}]$ can be written explicitly as

$$\boxed{[\hat{\beta}] = ([[x]]^T \cdot [[x]])^{-1} \cdot [[x]]^T \cdot [y],} \tag{6.53}$$

from which it can be seen that $[\hat{\beta}]$ is a linear function of $[y]$. This expression is very compact, but multiplying and inverting matrixes is painful. Computer codes are usually used.

The vector of fitted values $[\hat{y}]$ corresponding to the observed $[y]$ is

$$[\hat{y}] = [[x]] \cdot [\hat{\beta}] = H \cdot [y], \tag{6.54}$$

where

$$H = [[x]] \cdot ([[x]]^T \cdot [[x]])^{-1} \cdot [[x]]^T, \tag{6.55}$$

is an $n \times n$ matrix called the *hat matrix* [matrice chapeau] because it puts a hat to $[y]$. One obviously has that $H^T = H$ and $H^2 = H$.

6.2.2 Explained and unexplained deviations

As we did for the simple regression, for any observation, deviation from the mean can be decomposed into the sum of two quantities

$$y_i - \bar{y} = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i). \tag{6.56}$$

The first one, $(\hat{y}_i - \bar{y})$ represents the part explained by the model or the fit, and the second one, $\hat{\epsilon}_i = (y_i - \hat{y}_i)$, the residual that is unexplained by the model or the fit.

Remarkably, we can show that such a decomposition is still true with the sum of the squared deviations:

$$\boxed{\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2.} \tag{6.57}$$

Thus, for a multiple linear regression, we still have

$$\boxed{SST = SSR + SSE,} \tag{6.58}$$

where SST stands for the Total Sum of the Squared deviations [SCT, Somme des Carrés Totale], SSR denotes the Sum of Squares due to Regression [SCE, Somme des Carrés des variations Expliquées], and SSE represents the Sum of Squared Errors (or residuals) [SCR, Somme des Carrés des Résidus].

To evaluate the proportion of the total variation in Y that is accounted for by the explanatory variables X_k we define again the *coefficient of determination* [coefficient de détermination] as

$$\boxed{R^2 = \frac{\text{Explained variation}}{\text{Total variation}} = \frac{SSR}{SST} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}.} \tag{6.59}$$

R^2 gives us an idea of how the predictor variables X_k account for the response variable Y. It is also a measurement of the goodness of the fit [qualité de la prédiction d'une régression linéaire].

6.2.3 Predictions and uncertainties

Assumptions

As for the simple linear regression, we shall assume that the ϵ_i's are independent random quantities normally distributed with mean zero and a common variance σ^2 that is unknown. With matrix notations, this reads $\mathrm{E}([\epsilon]) = [0]$ and $\mathrm{Var}(\epsilon) = \mathrm{E}([\epsilon] \cdot [\epsilon]^T) = \sigma^2 I_n$ where I_n is the n-dimension identity matrix. Likewise, $\mathrm{Var}(y) = \sigma^2 I_n$.

From these hypothesis we can show that $[\hat{\beta}]$ is an unbiased estimator of $[\beta]$ as $\mathrm{E}([\hat{\beta}]) = [\beta]$. In other words, $\mathrm{E}(\hat{\beta}_k) = \beta_k$ for all indices $k = 0, \ldots, p$. Regarding the covariance matrix of these estimators, it is easy to show that

$$\mathrm{Var}(\hat{\beta}) = \begin{bmatrix} \mathrm{Var}\,\hat{\beta}_0 & \mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) & \ldots & \mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_p) \\ \mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_0) & \mathrm{Var}(\hat{\beta}_1) & \ldots & \mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_p) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(\hat{\beta}_p, \hat{\beta}_0) & \mathrm{Cov}(\hat{\beta}_p, \hat{\beta}_1) & \ldots & \mathrm{Var}(\hat{\beta}_p) \end{bmatrix}, \tag{6.60}$$

is given by

$$\boxed{\mathrm{Var}(\hat{\beta}) = \sigma^2([[x]]^T \cdot [[x]])^{-1}.} \tag{6.61}$$

Off all unbiased estimators of $[\beta]$ that are linear in the observations, the least squares estimator has minimum variance. For this reason, $[\hat{\beta}]$ is said to be the *best linear unbiased estimator* (BLUE) of $[\beta]$.

As for the simple regression, σ^2 is unknown and has to be estimated from the data.

Estimation of the variance in the errors

The vector of residuals is given by

$$[\hat{\epsilon}] = [y] - [\hat{y}] = [y] - H \cdot [y] = (I_n - H) \cdot [y].$$

Thus, the residual sum of squares can be expressed as

$$[\hat{\epsilon}]^T \cdot [\hat{\epsilon}] = [y]^T (I_n - H)^T (I_n - H)[y] = [y]^T (I - H)[y].$$

We can show that an unbiased estimator of σ^2 is

$$\boxed{\hat{\sigma}^2 = s^2 = \frac{[\hat{\epsilon}]^T \cdot [\hat{\epsilon}]}{n-p-1} = \frac{[y]^T (I-H)[y]}{n-p-1} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p-1} = \frac{SSE}{n-p-1}.} \tag{6.62}$$

The denominator corresponds to the number of degrees of freedom, i.e., n, the number of observations minus $(p+1)$, the number of coefficients of the model.

Predictions and uncertainties

The estimated covariance matrix of the coefficients of the model is

$$s^2(\hat{\beta}) = \hat{\sigma}^2 ([[x]]^T \cdot [[x]])^{-1}. \tag{6.63}$$

It allows to test hypothesis on the individual coefficients and the role of a given explanatory quantity.

The predicted value of Y for a given set of values $[x]$ for the $p+1$ explanatory quantities is

$$[y] = [\hat{\beta}]^T \cdot [x] + [\hat{\epsilon}], \tag{6.64}$$

and the associated uncertainty

$$s^2(y) = [x] \cdot s^2(\hat{\beta}) \cdot [x]^T + s^2(\hat{\beta}). \tag{6.65}$$

Once again, this expression is very compact but its implementation requires multiplication and inversion of matrixes.

Regression analysis without constant term

Likewise the simple regression, multiple linear regression without the constant term β_0 is particular. The estimator of the coefficients vector is still

$$[\hat{\beta}] = ([[x]]^T \cdot [[x]])^{-1} \cdot [[x]]^T \cdot [y],$$

and the estimated value of the response quantity for a given vector of explanatory quantities $[x]$ is given by

$$[\hat{y}] = [x]^T \cdot [\hat{\beta}]. \tag{6.66}$$

But the sum of the residues $\sum_{i=1}^n (y_i - \hat{y}_i) \neq 0$ which means that $SST \neq SSE + SSR$ anymore. Thus, the determination coefficient R^2 cannot be used.

The unbiased estimator of σ^2 is given by

$$\boxed{\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p} = \frac{SSE}{n-p},} \tag{6.67}$$

as there are $n-p$ degrees of freedom.

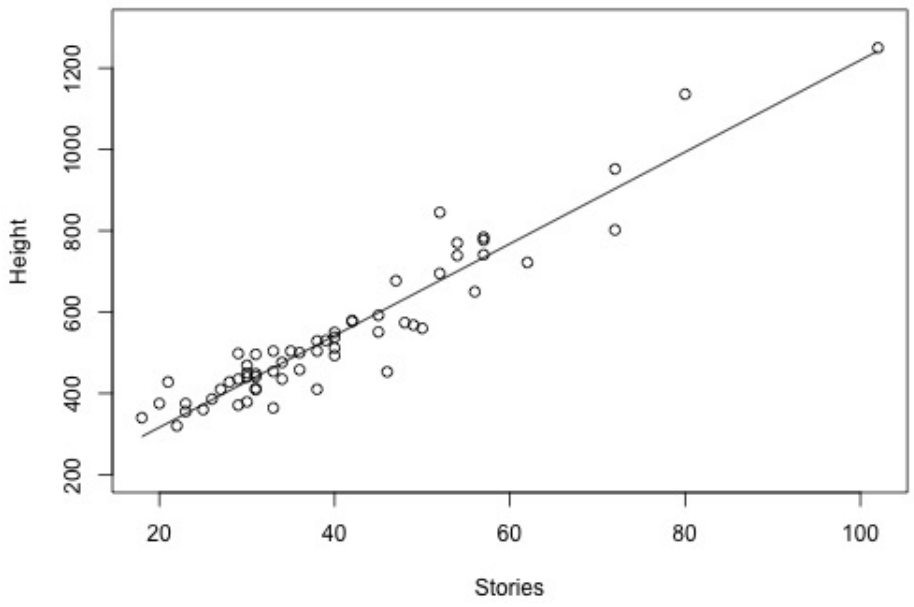


Figure 6.3: Relationship between the number of stories a building and its height.

With experimental uncertainties

If there are uncertainties u_i in the observed y_i's, it is easy take them into account in the least squares estimation and in the predictions. We define the weighting matrix

$$W = \begin{bmatrix} 1/u_1^2 & 0 & \dots & 0 \\ 0 & 1/u_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/u_n^2 \end{bmatrix}, \tag{6.68}$$

and formally implement it into the SSE that is minimised. The new SSE is given by

$$SSE = [\epsilon]^T \cdot W \cdot [\epsilon] = ([y] - [[x]] \cdot [\beta])^T \cdot W \cdot ([y] - [[x]] \cdot [\beta]). \tag{6.69}$$

We shall not go further in the formal developments, but note that the final uncertainty in the predictions will account for the experimental uncertainties and for the uncertainties coming from the errors of the model. This approach corresponds to the full generalisation of the least squares method that we studied in the previous chapter.

6.3 Illustrative examples

6.3.1 R-squared

Relationship between the number of stories a building and its height

Some statisticians compiled data on a set of $n = 60$ buildings reported in the 1994 World Almanac. The fitted line plot is shown in Fig. 6.3.

The software reports $R^2 = 90,4\%$. As expected, there is a strong correlation between the number of stories and the height. More precisely, this tells us that 90.4% of the variation in the height of the building is explained by the number of stories in the building.

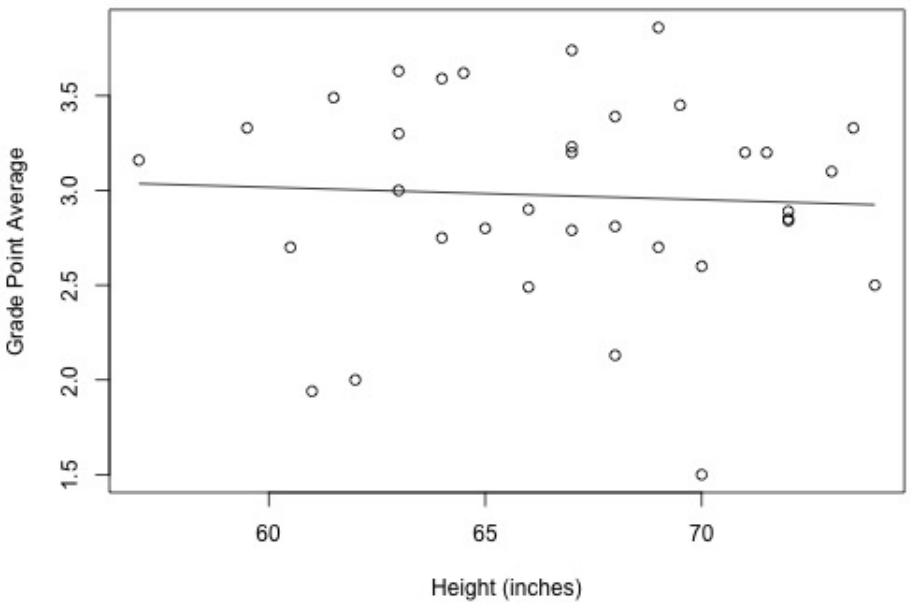


Figure 6.4: Relationship between the height of a student and his or her grade point average.

Relationship between the height of a student and his or her grade point average

Data were collected on a random sample of $n = 35$ students in a statistics course at Penn State University and the resulting fitted line plot is shown in Fig. 6.4.

The software reports that $R^2 = 0.3\%$. Not surprisingly, there is next to no linear relationship between height and grade point average!

6.3.2 Interpolation versus extrapolation

In practice, when we do prediction for some value of x we havent seen before, we need to be very careful. Predicting y for a value of x that is within the interval of points that we saw in the original data (the data that we fit our model with) is called interpolation. Predicting y for a value of x thats outside the range of values we actually saw for x in the original data is called extrapolation.

For real datasets, even if a linear fit seems appropriate, we need to be extremely careful about extrapolation, which can often lead to false predictions.

Remember that the variance in the predicted value $\hat{y}(x_0)$ for a value x_0 of x reads

$$s^2(\hat{y}(x_0)) = \hat{\sigma}^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right). \tag{6.70}$$

It increases when x_0 departs from $\bar{x}$. Moreover, there is no proof that the relationship remains linear outside the interval used for the regression. This is particularly true when a linear approximation is done over a finite range of values. The danger associated with extrapolation is illustrated in figure 6.5.

This is particularly the case when extrapolating in the past or the future. This is illustrated by a joke and a cartoon. *"If she or he loves you more and more every day, by linear regression she or he hated you before you met"*. The cartoon is in Fig. 6.6.

6.4 Conclusions

This chapter is only an introduction to linear regression analysis. There are several textbooks entirely devoted to the subject if you need to proceed further.

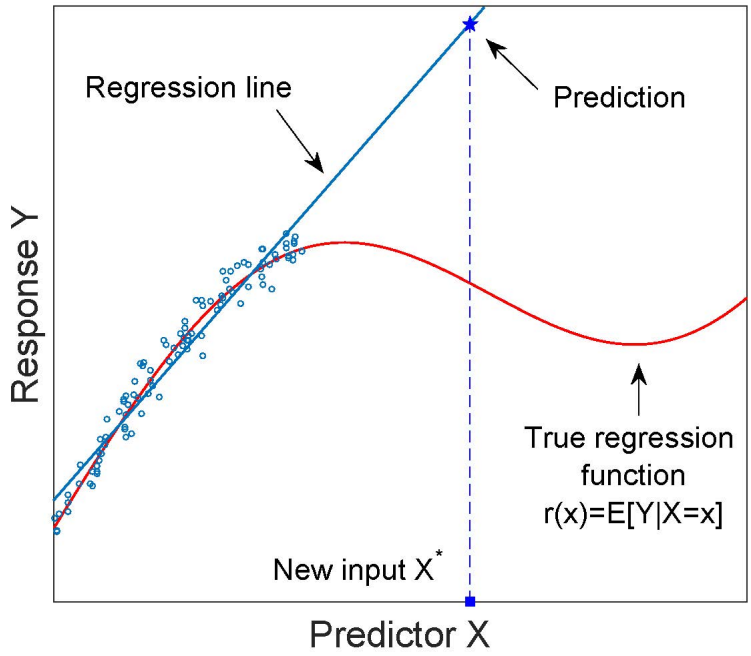


Figure 6.5: Example of wrong extrapolation.

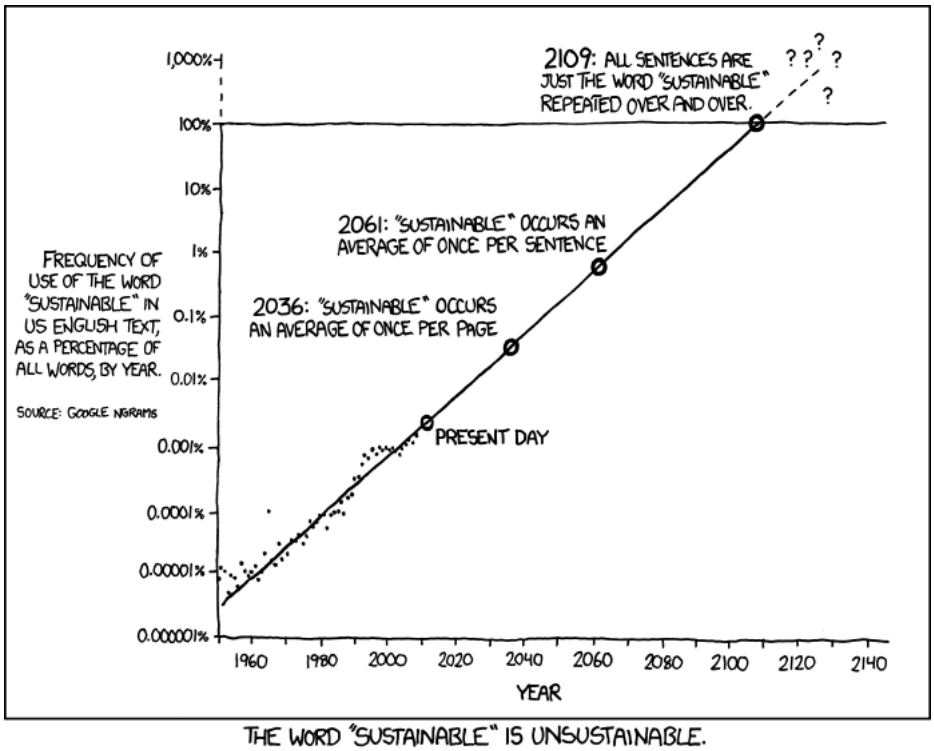


Figure 6.6: Example of wrong extrapolation. Figure reproduced from `www.xkcd.com`.

The linear regression analysis is appealing because it is a simple and powerful method for data analysis as it is based on simple analytical expressions. The least squares method is one of the oldest methods in data analysis that has lead to many progresses in sciences. Linear regression analysis is now used in all fields, including social sciences, as it allows a confrontation of models with data and inferences. However, the expressions are somehow heavy to handle when there are many data points and explanatory quantities. But these equations are implemented in scientific pocket calculators and softwares. Be careful, using these tools as black boxes is dangerous as this could lead to erroneous results. This lecture aims to give you an intuitive understanding of method and the tools.

There are other available methods that are not treated here and mainly devoted to specialists. Some of them are still discussed and developed by mathematicians. But most of them are not developed enough to be considered as a serious alternative to the linear regression. In particular, among others, we can mention the Bayesian linear regression. There are also extensions to more complicated cases such as non-linear regression.

Note that the Joint Committee for Guides in Metrology is preparing a Guide *Evaluation of measurement data - Applications of the least-squares method* (ISO/IEC Guide 98-5).

Part III

Optimising data production

Chapter 7

Design of experiment

"The real voyage of discovery consists not in seeking new landscapes, but in having new eyes."
Marcel Proust

How to be able to draw valid and definitive conclusions from data with the minimum use of ressources? This is the purpose of the design of experiments, also called experimental design, which is a structured and organised way of conducting and analysing tests to evaluate the factors that are affecting a response variable. It serves to improve the efficiency of experiments.

Design of experiments was invented by Ronald A. Fisher in the 1920s and 1930s at Rothamsted Experimental Station, an agricultural research station in England [22]. Although the experimental design method was first used in an agricultural context, the method has been applied successfully in many other contexts successfully since the 1940s.

To prepare this chapter, I have used the French textbooks of Goupy [23] and Sado [24] that emphasise the practical use of these designs rather than the mathematical developments. The former gives some justifications. In English, the most popular textbook seems to be the one of Montgomery [25] that also includes data analysis. The Goos and Jones textbook [26] gives illustrative examples explained in a simple manner.

7.1 Introductory example

This example is attributed to Harold Hotelling [39] in 1944. Here are simplified versions based on the same idea. The weight of objects is measured using a pan balance and set of standard weights.

7.1.1 Weighting of a single object with zeroing

The classical way is to put the object in the left pan for example and adding calibrated weights to the right pan until the balance is in equilibrium. We proceed the same way for zeroing the balance, with an empty left pan. Mathematically, the mass of the object m is just the difference between the results of the two measurements, y_1 and y_2 respectively: $m = y_1 - y_2$. If each measurement is affected by a uniform uncertainty $u(y)$, then $u(m) = \sqrt{2}u(y)$.

Another possibility for the second operation is to put the object in the right pan and the calibrated weights in the left one. The first reading y_1 indicates $y_1 = m \pm m_0$ where m_0 is the unknown zero. The second reading indicates $y_2 = m \mp m_0$. Thus,

$$m = \frac{1}{2}(y_1 + y_2) \qquad \text{and} \qquad u(m) = \frac{1}{\sqrt{2}}u(y). \tag{7.1}$$

The second experiment gives us 2 times as much precision for the estimate.

7.1.2 Weighting of two objects without zeroing

Similarly, to weight two objects, we usually weight them one after the other in order to get directly the answer: $m_1 = y_1$ and $m_2 = y_2$. Thus, $u(m_1) = u(m_2) = u(y)$.

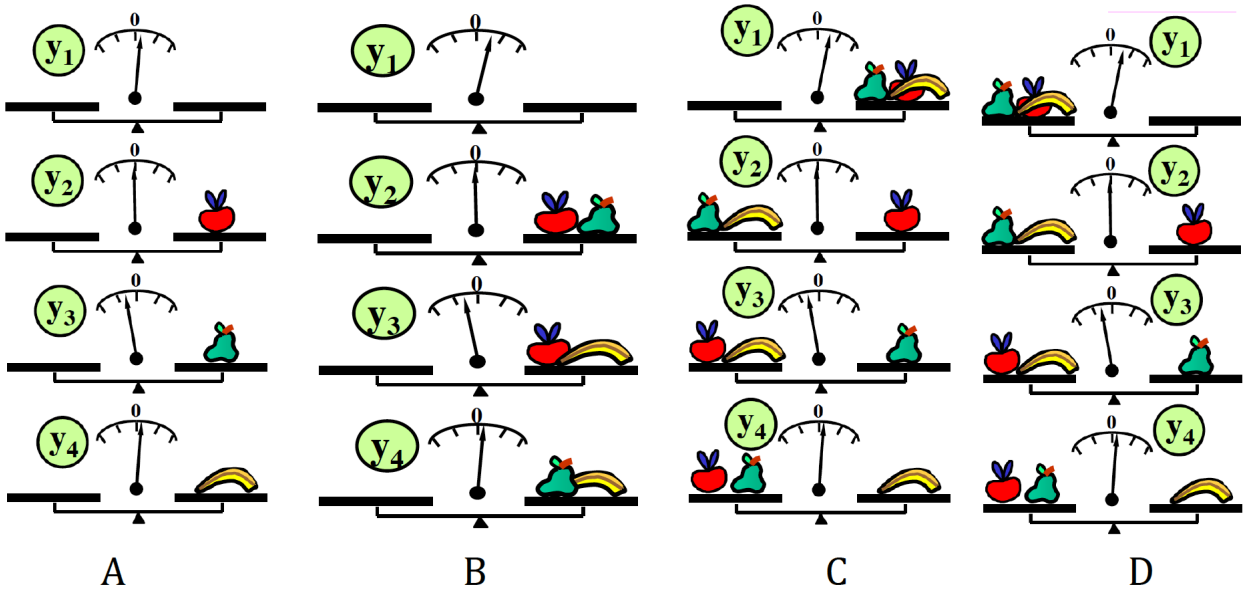


Figure 7.1: Various weighting strategies

Alternatively, we can first put both of them in the left pan a do a measurement: $y_1 = m_1 + m_2$. Then, we can one object in each pan and add calibrated weights to the lighter pan until the balance is in equilibrium: $y_2 = m_1 \pm m_2$. We can easily deduce that

$$m_1 = \frac{1}{2}(y_1 \pm y_2) \qquad \text{and} \qquad m_2 = \frac{1}{2}(m_1 \mp y_2). \tag{7.2}$$

Thus, $u(m_1) = u(m_2) = u(y)/\sqrt{2}$. Once again, we get a better accuracy with the same number of experiments.

7.1.3 Weighting of three objects with zeroing

Let us consider now the case of three objects. Beside the simplest series of measurements consisting in the weighting of each object separately, there are several other possibilities. Which one is optimum? We are going to introduce matrices to differentiate them.

Simple minded measurements

For the simple minded experiment design depicted in Fig. 7.1.A, we can summarise the experiments in a table:

Res.	m_0	m_1	m_2	m_3
y_1	1	0	0	0
y_2	1	1	0	0
y_3	1	0	1	0
y_4	1	0	0	1

Here m_0 stands for the zero and cannot be avoided. Each mass can be easily deduced, $m_i = y_{i+1} - y_1$, thus $u(m_i) = \sqrt{2}u(y)$.

Pair weighting

Let us consider now a weighting by pair of objects as in Fig. 7.1.B. The corresponding table reads

Res.	m_0	m_1	m_2	m_3
y_1	1	0	0	0
y_2	1	1	1	0
y_3	1	1	0	1
y_4	1	0	1	1

From the above table, one can deduce the relation between the results as a function of the masses

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \end{bmatrix} . \tag{7.3}$$

Inverting the matrix, we get the masses as a function the measurement results

$$\begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} . \tag{7.4}$$

Thus, we can easily get that $u(m_i) = u(y)$. It is better than the simple minded design with the same number of experiments, but not optimum as we shall in the next experimental design.

Optimum weighting

Let us now consider a weighting design as in Fig. 7.1.C. For the corresponding table, we assign -1 when the mass is in the other pan:

Res.	m_0	m_1	m_2	m_3
y_1	1	1	1	1
y_2	1	1	-1	-1
y_3	1	-1	1	-1
y_4	1	-1	-1	1

As before, inverting the matrix, we get

$$\begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} . \tag{7.5}$$

Thus, $u(m_i) = \frac{1}{2}u(y)$ with the same number of experiments. The corresponding variance is 8 times smaller than for the simple minded design. This means that from purely statistical point of view we would need 8 times more experiments with the simple minded design to get the same accuracy.

Consider now the weighting design of Fig. 7.1.D. The only difference is for the first experiment for which the other pan is used. You can check, with such a design, we would get $u(m_i) = \frac{1}{\sqrt{2}}u(y)$. This is not as good as the optimum case.

Conclusion

The final uncertainty in the masses depends on the design of experiments and the accuracy of the balance. The matrix that optimise the accuracy is the Hadamard matrix.

7.2 Hadamard matrices

In mathematics, a Hadamard matrix, named after the French mathematician Jacques Hadamard, is a $n \times n$ matrix whose entries are either 1 or -1 and whose rows are mutually orthogonal. It satisfies

$$HH^T = nI, \tag{7.6}$$

where H^T is the transposed matrix and I the identity matrix. Thus $H^{-1} = \frac{1}{n}H^T$ and $u(m) = \frac{1}{\sqrt{n}}u(y)$ when measuring n masses with a design following the Hadamard matrix. This is the optimum design.

The order of a Hadamard matrix must be 1, 2, or a multiple of 4. The first matrices are, for example,

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \ldots \tag{7.7}$$

They are not unique. Note that the 2×2 matrix corresponds to the weighing design that we used with two masses or a single mass with zeroing.

It is apparent that if the rows and columns of an Hadamard matrix are permuted, the matrix remains Hadamard type. It is also true that if any row or column is multiplied by -1, the Hadamard property is retained. Thus, it is always possible to arrange to have the first row and first column of an Hadamard matrix contain only +1 entries. An Hadamard matrix in this form is said to be normalized.

Examples of Hadamard matrices were actually first constructed by James Joseph Sylvester in 1867. Let H be a Hadamard matrix of order n. Then the partitioned matrix

$$\begin{bmatrix} H & H \\ H & -H \end{bmatrix}, \tag{7.8}$$

is a Hadamard matrix of order $2n$. Thus,

$$H_{2k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}, \tag{7.9}$$

for $2 \leq k \in \mathbb{N}$. Sylvester's matrices have a number of special properties. They are symmetric and, when $k \geq 1$, have trace zero. The elements in the first column and the first row are all positive. The elements in all the other rows and columns are evenly divided between positive and negative.

Note that the most important open question in the theory of Hadamard matrices is that of existence, but this out of the scope of this lecture. These Hadamard matrices will be also useful in the following.

7.3 Introduction to factorial designs

"All models are wrong. Some are useful."
G. Box and N. Draper, *Empirical Model Building and Response Surfaces*, Wiley, Hoboken, NJ (1987)

7.3.1 Method

Experimental design deals with comparative experiments that estimate differences in response between treatments. These could be different kinds or amounts of fertiliser in agronomy. Responses are outcomes that we observe after applying a treatment. That is, the response is what we measure to judge what happened in the experiment. We can have more than one response.

Factors combine to form treatments. For example, the baking treatment for a cake involves a given time at a given temperature. The treatment is the combination of time and temperature, but we can vary the time and temperature separately. Thus we speak of a time factor and a temperature factor.

The factors on the experiment are called input variables and denoted x_i. The performance measures resulting from the experiment are called responses and denoted y_i. In such an experiment, we are in control of the factors but we don't know the mechanism. In other words, we are interested in a function $y = f(x_i)$ without knowing it. The goal of the experimental design is to get the maximum information about this function with a minimum number of experiments.

Factorial designs [plans factoriels] are a special case with k factors, each at only two levels. These levels may be quantitative, such as two values of temperature, or qualitative, such as two machines. A complete replicate of such a design requires 2^k observations and is called a 2^k factorial design. Such a design is particularly useful in the early stage of experimental work when many factors are likely to be investigated. It provides the smallest number of runs with which k factors. Therefore, these designs are widely used in factor screening experiments.

Because there are only two levels of each factor, we assume that the response is approximately linear over the range of the factor levels chosen. In many factor screening experiments, when we are just starting to study the process or the system, this is often a reasonable assumption. Thus, we use the Taylor series approximations to the unknown true functional form of the response variable. The trick is to have the simplest model that captures the main features of the data or process. To the second order of the k input factors it reads,

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \sum_{j=1, j\neq i}^{k} \beta_{ij} x_i x_j + \ldots, \tag{7.10}$$

where the parameter β_0 is the overall mean response, β_i, the main effect for the factor i and β_{ij} the interaction between the factors i and j. The values of the β's are obtained by a fitting procedure. The experimental points have to be chosen in way to minimise the uncertainty in the parameters.

Factors are rescaled in order to vary between -1 and +1 or to be assigned value -1 and +1 when they are discrete. This is a code scale. A general recommendation for setting the factor ranges is to set the levels far enough apart so that one would expect to see a difference in the response but not so far apart as to be out of the likely operating range. The use of only two levels seems to imply that the effects must be linear, but the assumption of monotonicity (or nearly so) on the response variable is sufficient. At least three levels of the factors would be required to detect curvature.

By factorial designs, we mean that in a complete trial all possible combinaisons of the levels of the factors are investigated.

7.3.2 Example with one factor

The linear approximation becomes

$$y = \beta_0 + \beta_1 x_1, \tag{7.11}$$

with β_0 and β_1 to be determined. The minimum number of measurements is 2 chosen at the limits of the factor interval in order to have the best determination of the coefficients: $y_1(x_1 = +1) = \beta_0 + \beta_1$ and $y_2(x_1 = -1) = \beta_0 - \beta_1$. We easily get

$$\beta_0 = \frac{1}{2}(y_1 + y_2) \qquad \text{and} \qquad \beta_1 = \frac{1}{2}(y_1 - y_2). \tag{7.12}$$

In matrix notation,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}. \tag{7.13}$$

The first matrix of r.h.s. is a Hadamard matrix, meaning an optimum design. Inversion is easy and gives

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \tag{7.14}$$

This simple example shows the merit to have rescaled the factor to the $[-1, 1]$ interval.

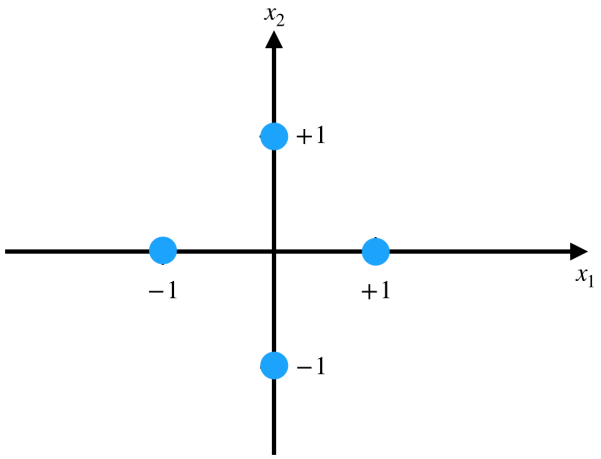


Figure 7.2: Simple design with two factors.

7.3.3 Example with 2 factors

Let us consider a simple example in chemistry, with a reaction rate depending on temperature T varying between 60° and 80°C and pressure P varying between 1 and 2 bar. There are four coefficients to determine:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2, \tag{7.15}$$

with

$$x_1 = \frac{2T - (T_{min} + T_{max})}{T_{max} - T_{min}} \qquad \text{and} \qquad x_2 = \frac{2P - (P_{min} + P_{max})}{P_{max} - P_{min}}.$$

Naively, we could fix the temperature at 70°C and vary the pressure between 1 and 2 bar in order to determine the effects of pressure. Then, similarly, we could fix the pressure at 1.5 bar and vary the temperature. Such a simple design is depicted in Fig. 7.2. But there are problems with the traditional experimental method of changing one factor at a time, i.e., its inefficiency and its inability to determine effects that are caused by several factors acting in combination.

Factorial experimentation is a method in which the effects due to each factor and to combinations of factors are estimated. Factorial designs are geometrically constructed and vary all the factors simultaneously and orthogonally. This is achieved with the following experimental matrix:

Rate (%)	Temperature x_1	Pressure x_2
$y_1 = 60$	-1	-1
$y_2 = 70$	+1	-1
$y_3 = 80$	-1	+1
$y_4 = 95$	+1	+1

Such an optimal design is depicted in Fig. 7.3.

From the results reported in the above table, we notice that at low temperature, the effect of pressure is $(y_3 - y_1)/2 = 10\%$ and at hight temperature, it is $(y_4 - y_2)/2 = 12.5\%$. Similarly, the effect of the temperature is not the same at low and high pressure. There is an interaction that occurs when the effect on the response of a change in the level of one factor from low to high depends on the level of another factor. In other words, when an interaction is present between two factors, the combined effect of those two factors on the response variable cannot be predicted from the separate effects. The effect of two factors acting in combination can either be greater (synergy) or less (interference) than would be expected from each factor separately.

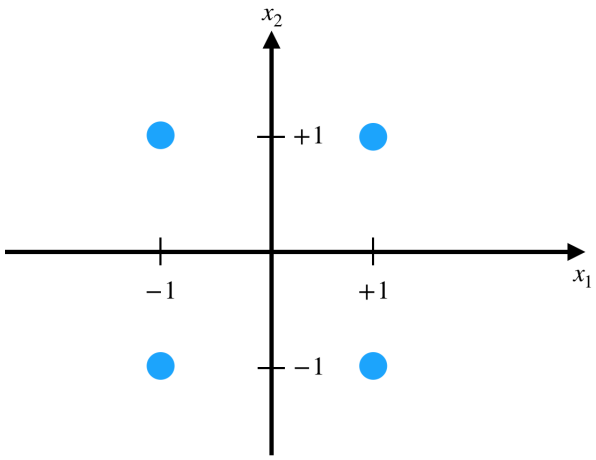


Figure 7.3: Optimal design with two factors.

In this example, the highest rate (y_4) is obtained with the highest temperature and the highest pressure. If that was the question, we can stop there. However, if we need to know the reaction rate for an interpolated value of temperature and pressure, we need to determine the β's. From Eq. (7.15) one has

$$y_1 = \beta_0 - \beta_1 - \beta_2 + \beta_{12} \tag{7.16}$$
$$y_2 = \beta_0 + \beta_1 - \beta_2 - \beta_{12} \tag{7.17}$$
$$y_3 = \beta_0 - \beta_1 + \beta_2 - \beta_{12} \tag{7.18}$$
$$y_4 = \beta_0 + \beta_1 + \beta_2 + \beta_{12}. \tag{7.19}$$

To invert these equations, it is easier to use the matrix formalism. To build up the matrix, we write the experiments table:

Res.	Mean	x_1	x_2	x_1x_2
y_1	+1	-1	-1	+1
y_2	+1	+1	-1	-1
y_3	+1	- 1	+1	-1
y_4	+1	+1	+1	+1

In the mean column, all values are obviously +1. In the first factor column, we alternate ± 1, starting with -1. In the second factor column, we alternate ± 1 every two rows, starting with -1. Finally, in the interaction column, the sign is obtained by multiplying the value in column x_1 with the value in the column x_2. Thus,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix}. \tag{7.20}$$

Once again, the first matrix of the r.h.s. is a Hadamard matrix easy to inverse:

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} +1 & +1 & +1 & +1 \\ -1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 \\ +1 & -1 & -1 & +1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}. \tag{7.21}$$

Actually, the values of the β's can directly be obtained by completing the experiments table:

Res. (%)	Mean	x_1	x_2	x_1x_2
$y_1 = 60$	+1	-1	-1	+1
$y_2 = 70$	+1	+1	-1	-1
$y_3 = 80$	+1	- 1	+1	-1
$y_4 = 95$	+1	+1	+1	+1
	$\beta_0 = 76.25$	$\beta_1 = 6.25$	$\beta_2 = 11.25$	$\beta_{12} = 1.25$

In this table,

- $\beta_0 = \frac{1}{4}(y_1 + y_2 +_3 +y_4)$ is the mean value, obtained by summing all the responses as there are only + in the column and dividing the answer by 4, the number of tests.
- $\beta_1 = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4) = \frac{1}{4}((y_2 + y_4) - (y_1 + y_3))$ is the effect of the first factor, obtained by summing all the responses weighted by the corresponding sign indicated in the column and by dividing the answer by 4. It yields the main effect of x_1.
- β_2 is the effect of the second factor obtained in a similar manner. It yields the main effect of x_2.
- Finally, $\beta_{12} = \frac{1}{4}(+y_1 - y_2 - y_3 + y_4)$ is the interaction between the factors 1 and 2 obtained in a similar manner.

Pressure has a greater influence than temperature ($\beta_2 > \beta_1$). The greater the temperature, the greater the influence of the pressure as $\beta_{12} > 0$.

7.3.4 Example with 3 factors

With three factors, we shall directly apply the scheme presented with two factors. At least 2^3 measurements are necessary to fix the 2^3 coefficients of the Taylor expansion of the response function:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3. \tag{7.22}$$

As an example, let us consider a soft drink bottling problem adapted from [25]: A soft drink bottler is interested in obtaining more uniform fill heights in the bottles. There is variation around the target level and the bottler would like to understand the sources of this variability. The process engineer can control three factors during the filling process: the percent carbonation, the operating pressure in the filler and the line speed in bottle per minute (bpm). The response is the mean deviation from the target. The factors are summarised in the table:

Factor	-	+
Percent carbonation	10%	14%
Operating pressure	175 kPa	210 kPa
Line speed	200 bpm	250 bpm

The experimental matrix and the results are:

Res.	mean	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
$y_1 = -2$	+	-	-	-	+	+	+	-
$y_2 = 4.5$	+	+	-	-	-	-	+	+
$y_3 = -0.5$	+	-	+	-	-	+	-	+
$y_4 = 8$	+	+	+	-	+	-	-	-
$y_5 = -0.5$	+	-	-	+	+	-	-	+
$y_6 = 6.5$	+	+	-	+	-	+	-	-
$y_7 = 1$	+	-	+	+	-	-	+	-
$y_8 = 10.5$	+	+	+	+	+	+	+	+
	$\beta_0 = 3.44$	$\beta_1 = 3.94$	$\beta_2 = 1.31$	$\beta_3 = 0.94$	$\beta_{12} = 0.56$	$\beta_{13} = 0.19$	$\beta_{23} = 0.06$	$\beta_{123} = 0.06$

Like in the two-factor case, the first column corresponding to the mean is filled with +1. In the first factor column, we alternate ±1, starting with −1. In the second factor column, we alternate ±1 every two rows, starting with −1. In the third factor column, we alternate ±1 every four rows, starting with −1. The signs in the interaction columns are obtained by a simple multiplication.

Finally, the β's are the weighted mean of the results by the signs given in each column. For example, $\beta_{12} = \frac{1}{8}(y_1 - y_2 - y_3 + y_4 + y_5 - y_6 - y_7 + y_8)$.

These examples show how clever is the method that allows to get the coefficients with little efforts. Nowadays, there are many softwares available to treat data.

7.3.5 Further developments

Number of experiments

Even for a moderate number of factors, the total number of experiments in a 2^k factorial design is large. For example, a 2^5 design has 32 treatment combinaisons. Thus, the number of replicates that the experimenter can employ may be restricted. When a single replicate of the design is run, there is a risk to fit a model to noise. In order words, if the response y is highly variable, misleading conclusions may result. A good practice in this case is to spread out the factor level aggressively.

Note that when k the number of factors increases, the number of experiments equal to 2^k might become too large to be performed. Fractional factorial designs have been developed. The idea is to test the interactions between parameters to check if some of them can be neglected. See textbooks for details.

Linearity

When possible, an additional measurement can be done at 0 to all factors in order to test the linearity of the model. The response could be compared to the β_0 obtained by averaging all other responses. To test the compatibility of the two results one should know the uncertainty.

7.4 Uncertainty analysis

So far, we have not taken the uncertainty in measurements into account. However, it is important to determine if a factor has a significant effect or not.

We shall distinguish two cases. First, we know the uncertainty in the measurement of the response and we shall use it. Second, we don't know it and we need to perform other measurements for a type-A evaluation of the uncertainty. In the latter case, we shall use what we learned in the regression chapter.

7.4.1 Known uncertainty

Direct calculation

Let us suppose that the uncertainty in the response measurements is known and uniform. As the response is related to the coefficient by a Hadamard matrix, the uncertainty in the coefficient will the smallest. Consider the 2^2 design. Eq. (7.21) gives for each coefficient,

$$\beta_i = \frac{1}{4}\sum_{j=1}^{4} \pm y_j, \qquad \text{thus} \qquad u(\beta_i) = \frac{1}{2}u(y). \tag{7.23}$$

This can be generalised to a 2^k factorial design for which $u(\beta_i) = \frac{1}{\sqrt{2^k}}u(y)$.

If $u(\beta_i) \ll \beta_i$ we can easily conclude that the factor i has an effect on the response. On the contrary, if $u(\beta_i) \gg \beta_i$, the effect is not significant. Then, if $u(\beta_i) \simeq \beta_i$, there is no easy conclusion. One has to do a statistical analysis with hypothesis testing. Or, if it is very important to know if the factor has an effect or not, additional tests should be performed to be able to conclude.

With regression analysis

Considering again a 2^2 design with a single replicate, this is a four-run design with results $[y] = [y_1, y_2, y_3, y_4]^T$. The model we fit to the data is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon. \tag{7.24}$$

The results of the four runs can be written in a matrix form

$$[y] = [[X]]\cdot[\beta]+[\epsilon], \quad \text{with} \quad [y] = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}, [[X]] = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, [\beta] = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix}, \quad \text{and} \quad [\epsilon] = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix}. \tag{7.25}$$

The least square estimate of the model parameters are

$$[\hat{\beta}] = ([[X]]^T \cdot [[X]])^{-1} \cdot [[X]]^T \cdot [y]. \tag{7.26}$$

For the $2^2 = 4$ design

$$[[X]]^T \cdot [[X]] = 4I \qquad \text{and} \qquad ([[X]]^T \cdot [[X]])^{-1} = \frac{1}{4} I, \tag{7.27}$$

where I is the identity matrix. Regarding the uncertainty in the parameters of the model, we found in the chapter on regression analysis that

$$\text{Var}(\hat{\beta}) = \sigma^2([[X]]^T \cdot [[X]])^{-1} = \frac{\sigma^2}{4} I. \tag{7.28}$$

All model parameters have the same variance and they have no covariance.

The variance of the predicted response is a function of the point in the design space where the prediction is made:

$$u^2(\hat{y}(x_1, x_2)) = \frac{\sigma^2}{4}(1 + x_1^2 + x_2^2 + x_1^2 x_2^2). \tag{7.29}$$

The maximum occurs when $x_1 = \pm 1$ and $x_2 = \pm 1$ and is equal to σ^2. The smallest uncertainty is at the centre of the design where $x_1 = x_2 = 0$ and is equal to $\sigma^2/4$.

7.4.2 Unknown uncertainty

When the uncertainty in the response measurement is unknown, we should add some points in order to perform a type A estimate of the uncertainty. Various strategies are generally implemented. Here, we shall study the case of repeated measurements at the centre and the case of n replicates of the whole design. As the latter is very expensive, it is seldom used.

Measurements at the centre

One possibility consists in doing n_0 repeated measurements at the centre (all factors set to zero) when possible and use the type A standard deviation. It is then assumed that this uncertainty is the same for the other points. Let us consider an example.

To study the yield of a chemical process with three variables of interest, temperature, pressure and catalyst concentration, an engineer runs a 2^3 design with four centre points. Each variable can be run at low and high levels. The design and the resulting yields are shown in Table 7.1.

Suppose that the engineer decides to fit a main effects only model that reads

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon. \tag{7.30}$$

htb

Table 7.1: Experimental design for the yield of chemical process

Temperature	Pressure	Concentration	Yield
-1	-1	-1	32
1	- 1	-1	46
-1	1	-1	57
1	1	- 1	65
-1	-1	1	36
1	- 1	1	48
-1	1	1	57
1	1	1	68
0	0	0	50
0	0	0	44
0	0	0	53
0	0	0	56

For this model, the $[[X]]$ matrix and $[y]$ vectors are

$$[[X]] = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad [y] = \begin{bmatrix} 32 \\ 46 \\ 57 \\ 65 \\ 36 \\ 48 \\ 57 \\ 68 \\ 50 \\ 44 \\ 53 \\ 56 \end{bmatrix}. \tag{7.31}$$

This is an orthogonal design, and even with the added centre runs it is still orthogonal. Therefore the $[[X]]^T \cdot [[X]]$ matrix is diagonal:

$$[[X]]^T \cdot [[X]] = \begin{bmatrix} 12 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix} \quad \text{and} \quad [[X]]^T \cdot [y] = \begin{bmatrix} 612 \\ 45 \\ 85 \\ 9 \end{bmatrix} \tag{7.32}$$

Inverting $([[X]]^T \cdot [[X]])$ is then easy and

$$[\hat{\beta}] = ([[X]]^T \cdot [[X]])^{-1} \cdot [[X]]^T \cdot [y] = \begin{bmatrix} 1/12 & 0 & 0 & 0 \\ 0 & 1/8 & 0 & 0 \\ 0 & 0 & 1/8 & 0 \\ 0 & 0 & 0 & 1/8 \end{bmatrix} \cdot \begin{bmatrix} 612 \\ 45 \\ 85 \\ 9 \end{bmatrix} = \begin{bmatrix} 51.000 \\ 5.625 \\ 10.625 \\ 1.125 \end{bmatrix}. \tag{7.33}$$

These parameters could have been obtained by the traditional way. The variance of the model parameters are found from the diagonal elements of $([[X]]^T \cdot [[X]])^{-1}$. That is

$$u^2(\hat{\beta}_0) = \frac{\sigma^2}{12}, \qquad \text{and} \qquad u^2(\hat{\beta}_i) = \frac{\sigma^2}{8},\ i = 1, 2, 3. \tag{7.34}$$

An unbiased estimator of σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{12-4}\sum_{i=1}^{12}(y_i - \hat{y}_1)^2 = \frac{1}{8}([y] - [[X]] \cdot [\hat{\beta}])^T \cdot ([y] - [[X]] \cdot [\hat{\beta}]) \tag{7.35}$$

$$= \frac{1}{8}[y]^T \cdot (I - H) \cdot [y], \tag{7.36}$$

with $H = [[X]] \cdot ([[X]]^T \cdot [[X]])^{-1} \cdot [[X]]^T$, the hat matrix which does not have a simple form. Finally, we get $s = \hat{\sigma} = 3.347$.

It is then possible to test the hypothesis $H_0 : \beta_i = 0$ using the Student distribution.

Repetition of the whole experimental design

Repeating several times the whole experiment, we get n results per point. We can use the matrix formalism from the regression analysis or for a point, the response value is simply replaced by the mean value of the n replicates. Then parameters can be estimated as usual.

The uncertainty in the response at a given point is obtained from

$$s^2(y_i) = \frac{(y_{i1} - \bar{y}_i)^2 + (y_{i2} - \bar{y}_i)^2 + \ldots + (y_{in} - \bar{y}_i)^2}{n-1}. \tag{7.37}$$

If one assumes that the uncertainty in each response point is the same, we get, for all points

$$s^2 = \frac{\nu_1 s^2(y_1) + \ldots + \nu_N s^2(y_N)}{\nu_1 + \ldots + \nu_N} = \frac{1}{N}\sum_{i=1}^{N} s^2(y_i), \tag{7.38}$$

with $N = 2^k$.

According to Goupy [23], it is very rare to repeat the whole design as it costs too much. When possible, the time and money is rather used to test more factors. When one factor turns out to be not influential, we can forget about it and use the repeated results with the above formulas.

7.5 Beyond ordinary designs: A simple comparative experiment

So far, we have only considered simple designs based on several assumptions such as a uniform uncertainty in all response results. This is not always the case. In this section we shall consider other cases with other constraints. This example is adapted from Ref. [26].

Let us consider the comparison of the average thickness of cables produced on two different machines. We want to compare two means, say μ_1 and μ_2 and we have an experimental budget that allows for $n = 12$ observations, n_1 observations from the first machine and $n_2 = n - n_1$ observations from machine 2.

The sample of n_1 observations from the first machine allows us to calculate a sample mean $\bar{x}_1$ with variance σ^2/n_1. Similarly, we can calculate a sample mean $\bar{x}_2$ from the n_2 observations from the second machine with variance σ^2/n_2. Here, we are assuming that the variance in thickness is σ^2 for both machines and that all observations are statistically independent.

We are interested in comparing the two means by calculating the difference between the two sample means, $\bar{x}_1 - \bar{x}_2$. As usual, we want its uncertainty

$$u(\bar{x}_1 - \bar{x}_2) = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \tag{7.39}$$

Table 7.2: Uncertainty in the sample mean difference for different sample size n_1 and n_2 for $\sigma = 1$. The optimum size is in bold characters.

n_1	1	2	3	4	5	**6**	7	8	9	10	11
n_2	11	10	9	8	7	**6**	5	4	3	2	1
$u(\bar{x}_1 - \bar{x}_2)$	1.044	0.775	0.667	0.612	0.586	**0.577**	0.586	0.612	0.667	0.775	1.044

Table 7.3: Uncertainty in the sample mean difference for different sample size n_1 and n_2 for $\sigma_1 = 1$ and $\sigma_2 = 3$. The optimum size is in bold characters.

n_1	1	2	**3**	4	5	6	7	8	9	10	11
n_2	11	10	**9**	8	7	6	5	4	3	2	1
$u(\bar{x}_1 - \bar{x}_2)$	1.348	1.183	**1.155**	1.173	1.219	1.291	1.394	1.541	1.764	2.145	3.015

to be small. We can compute all possible share of the allocated budget between the observations from machines 1 and 2. See table 7.2. As expected the optimum design consist in a balance design for which the number of observations from the two machine is equal.

But if we drop the assumption that both machines have the same uncertainty, what is the optimum design? Let us denote the uncertainties of machine 1 and 2 by σ_1 and σ_2 respectively. The uncertainty in the same mean difference is

$$u(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}. \tag{7.40}$$

Table 7.3 shows the results if $\sigma_2^2 = 9\sigma_1^2$, or $\sigma_2 = 3\sigma_1$. This time, the optimal design requires 3 observations from machine 1 and 9 observations from machine 2. The larger value for n_2 compensates the large uncertainty for the machine 2 and ensures that $u(\bar{x}_2)$ is not excessively large. Thus, balance designs are not always optimal. However, in the present example, the uncertainty for the balance design does not differ much from the optimal design.

Let us now consider the case when the measurements do not have the same cost. Suppose that an observation from population 1 costs twice as much as an observation from population 2. To keep things simple, let the costs be 2 and 1, and let the total budget be 24. Then, we have 11 ways to spend the experimental budget. See table 7.4. This is another example where the balanced design is not optimum.

As a conclusion, the standard design is optimal for one scenario, namely, in the scenario where the number of observations is even, the uncertainties in the observations are identical and the cost of an observation is the same for both populations. The standard design is no longer optimal as soon as one of the usual assumptions is not valid anymore.

Table 7.4: Uncertainty in the sample mean difference for different sample size n_1 and n_2 for $\sigma = 1$ when an observation from population 1 costs twice as much as an observation from population 2. The optimum size is in bold characters.

n_1	1	2	3	4	5	6	**7**	8	9	10	11
n_2	22	20	18	16	14	12	**10**	8	6	4	2
$u(\bar{x}_1 - \bar{x}_2)$	1.022	0.742	0.624	0.559	0.521	0.500	**0.493**	0.500	0.527	0.592	0.769

7.6 In practice

The main steps to implement an experimental design are as follows:

1. State the objective of the study and the hypotheses to be tested.
2. Determine the response variable(s) of interest that can be measured.
3. Determine the controllable factors of interest that might affect the response variables and the levels of each factor to be used in the experiment. It is better to include more factors in the design than to exclude factors, that is, prejudging them to be nonsignificant.
4. Determine the uncontrollable variables that might affect the response variables, blocking the known nuisance variables and randomising the runs to protect against unknown nuisance variables.
5. Design and perform the experiment.
6. Analyse the data from the experiment using the analysis of variance method.
7. Interpret the results and state the conclusions in terms of the subject matter.
8. Consider performing a second, confirmatory experiment if the conclusions are very important or are likely to be controversial.
9. Document and summarise the results and conclusions, in tabular and graphical form, for the report or presentation on the study.

Part IV

Appendix and bibliography

Appendix A

Editorial: Uncertainty Estimates

PHYSICAL REVIEW A**83**, 040001 (2011)

The purpose of this Editorial is to discuss the importance of including uncertainty estimates in paper involving theoretical quantities.

It is not unusual for manuscripts on theoretical fork to be submitted without uncertainty estimates for numerical results. In contrast, papers representing the results of laboratory measurements would usually not be considered acceptable for publication in *Physical Review A* without a detailed discussion of the uncertainties involved in the measurements. For example, a graphical presentation of data is always accompanied by error bars for the data points. The determination of these error bars is often the most difficult part of the measurement. Without them, it is impossible to tell whether or not bumps and irregularities in the data are real physical effects, or artifacts of the measurement. Even papers reporting the observation of entirely new phenomena need to contain enough information to convince the reader that the effect being reported is real. The standards become much more rigorous for papers claiming high accuracy.

The question is to what extend can the same high standards be applied to papers reporting the results of theoretical calculations. It is all too often the case that the numerical results are presented without uncertainty estimates. Authors sometimes say that it is difficult to arrive at error estimates. Should this be considered an adequate reason for omitting them? In order to answer this question, we need to consider the goals and objectives of the theoretical (or computational) work being done. Theoretical papers can be broadly classified as follow:

1. Development of new theoretical techniques or formalism.
2. Development of approximation methods, where the comparison with experiment, or other theory, itself provides an assessment of the error in the method of calculation.
3. Explanation of previously unexplained phenomena, where a semiquantitative agreement with experiment is already significant.
4. Proposals for new experimental arrangements or configurations, such as optical lattices.
5. Quantitative comparisons with experiment for the purpose of (a) verifying that all significant physical effects have been taken into account, and/or (b) interpolating or extrapolating known experimental data.
6. Provision of benchmark results intended as references data or standards of comparison with other less accurate methods.

It is primarily papers in the last two categories that require a careful assessment of the theoretical uncertainties. The uncertainties can arise from two sources: (a) the degree to which the numerical results accurately represent the predictions of an underlying theoretical formalism, for example, convergence with the size of a basis set, or the step size in a numerical integration, and (b) physical effects not included in the calculation from the beginning, such as electron correlation and relativistic corrections. It is of course never possible to state precisely what the error is without doing a larger calculation and obtaining the higher accuracy. However,

the same is true for the uncertainties in experimental data. The aim is to estimate the uncertainty, not to state the exact amount of the error or provide a rigorous bound.

There are many cases where it is indeed not practical to give a meaningful error estimate for a theoretical calculation; for example, in scattering process involving complex systems. The comparison with experiment itself provides a test of our theoretical understanding. However, there is a broad class of papers where estimates can and should be made. Papers presenting the results of theoretical calculations are expected to include estimates for the calculations whenever practicable, and especially under the following circumstances:

1. If the authors claim high accuracy, or improvements on the accuracy of previous work.
2. If the primary motivation for the paper is to make comparisons with present to future high precision experimental measurements.
3. If the primary motivation is to provide interpolations or extrapolations of known experimental measurements.

These guidelines have been used on a case-by-case basis for the past two years. Authors have adapted well to this, resulting in papers of greater interest and significance for our reader.

The Editors

Published 29 April 2011

Bibliography

International standards

[1] Bureau International des poids et mesures, *The International System of Units (SI)* , (8th edition, 2006), `http://www.bipm.org/`. (Free)

[2] *International Vocabulary of Metrology - Basic and General Concepts and Associated Terms*, (2012), `http://www.bipm.org/`. In is in French and English. (Free)

[3] *Essentials of expressing measurement uncertainty,* Physical Measurement Laboratory of the National Institute of Standards and Technology (2000), `http://physics.nist.gov/cuu/Uncertainty/index.html`

[4] Bureau international des poids et mesures, *Le Système international dunités - The International System of Units (SI)*, (2006) `https://www.bipm.org/en/publications/si-brochure/`

[5] *Evaluation of measurement data - Guide to the expression of uncertainty in measurement*, (2008), `http://www.bipm.org/`. (Free)

[6] *Evaluation of measurement data - An introduction to the "Guide to the expression of uncertainty in measurement" and related documents*, JCGM 104:2009, `http://www.bipm.org/`. (Free)

[7] *Evaluation of measurement data - Supplement 1 to the "Guide to the expression of uncertainty in measurement" - Propagation of distributions using a Monte Carlo method*, JCGM 101:2008, `http://www.bipm.org/`. (Free)

[8] *Evaluation of measurement data - Supplement 2 to the "Guide to the expression of uncertainty in measurement" - Extension to any number of output quantities*, JCGM 102:2008, `http://www.bipm.org/`. (Free)

[9] *Determination of the characteristic limits (decision threshold, detection limit and limits of the confidence interval) for measurements of ionizing radiation - Fundamentals and application*, ISO 11929 (2009). A French version is also available (2010). See `www.ortec-online.com/download/Upgrading-GV-ISO-11929.pdf`

Books and lecture notes

[10] John R. Taylor, *An introduction to error analysis, the study of uncertainties in physical environments*, University Science Books, (1982). Il existe une version française publiée chez Dunod (2000).

[11] Les Kirkup and Bob Frenkel, *An Introduction to Uncertainty in Measurement*, Cambridge University Press (2006). One of the rare books on uncertainties based on the GUM.

[12] Ifan G. Hughes et Thomas P.A. Hase, *Measurements and their uncertainties, a practical guide to modern error analysis*, Oxford University Press (2010).

[13] Harold Jeffreys, *Theory of Probability*, Second edition, Oxford University Press (1948)

[14] E.T. Jaynes, *Probability Theory - The Logic of Science*, Cambridge University Press (2003)

[15] D. S. Sivia, *Data analysis. A Bayesian tutorial*, second edition, Oxford University Press (2006).

[16] Ian Hacking, *The Emergence of Probability - A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*, Cambridge University Press (2015)

[17] Stephen M. Stigler, *The History of Statistics, The Measurement of Uncertainty before 1900*, Cambridge, Massachusetts, The Belknap press of Harvard University press 1986. An excellent book on history of sciences with explanation on the calculations.

[18] Emmanuel Dion, *Invitation à la théorie de l'information*, éditions du Seuil, collection Points Sciences (1997). Ouvrage de vulgarisation scientifique.

[19] Dominique Gibert and Fernando Lopes, *Théorie de l'information*, lecture notes downloaded in 2013. Not available online anymore.

[20] Yadolah Dodge et Valentin Rousson, *Analyse de régression appliquée*, Dunod (2004). Livre très pédagogique.

[21] Samprit Chatterjee and Ali S. Hadi, *Regression Analysis by Example*, Wiley (2012).

[22] Ronald A. Fisher, *The Design of Experiments*, Macmillan (1935)

[23] Jacques Goupy, *La méthode des plans d'expériences - optimisation du choix des essais et de l'interprétation des résultats*, Dunod (1996)

[24] Gilles Sado ate Marie-Christine Sado, *De l'expérimentation à l'assurance qualité*, AFNOR (2000)

[25] Douglas C. Montgomery, *Design and Analysis of Experiments*, Wiley (2001)

[26] Peter Goos and Bradley Jones, *Optimal Design of Experiments: A Case Study Appoach*, Wiley-Backwell (2011)

Articles

Pedagogical or review articles

[27] E.T. Jaynes, *Where do we stand on maximum entropy?*, in The Maximum Entropy Formalism, R. D. Levine and M. Tribus (eds.), M. I. T. Press, Cambridge, MA (1979) 15. `http://bayes.wustl.edu/etj/articles/stand.on.entropy.pdf` Fascinating!

[28] W. Tyler Estler, *Measurement as Inference: Fundamental Ideas*, Annals of the CIRP **48** (1999) 611

[29] A. Vivier, R. Fottorino et B. Rousse, *Seuil de décision et limite de détection : estimation, interprétation et optimisation, première partie : les principes de base*, Radioprotection **45** (2010) 321

[30] Steve Pressé, Kingshuk Ghosh, Julian Lee and Ken A. Dill, *Principles of maximum entropy and maximum caliber in statistical physics*, Reviews of Modern Physics, **85** (2013) 1115

[31] Gregory Chaitin, *The Limits of Reason*, Scientific American, March 2006

[32] David Boilley et Yoann Lallouet, *200 ans après Laplace, l'essor des méthodes bayésiennes d'analyse des données*, Bulletin de l'Union des Physiciens Vol. 110 - Février 2016

Historical articles

[33] T. Bayes, *An essay towards solving a Problem in the Doctrine of Chances*, Phil. Trans. Roy. Soc. (1763) 330. Repris dans Biometrika **45** (1958) 293: `http://rstl.royalsocietypublishing.org/content/53/370.full.pdf`. Or here in modern English: `http://www.stat.ucla.edu/history/essay.pdf`. Hard to read.

[34] P.S. de Laplace, *Mémoire sur la probabilité des causes par les événements*, Mémoires de l'Académie royale des sciences de Paris (Savants étrangers), t. VI. p. 621 (1774).

[35] P.S. de Laplace, *Essai philosophique sur les probabilités* (1774). Read also *Théorie analytique des probabilités*, imprimerie royale (1847), réédité par Gabay (1995). Texts available in French on the homepage of the BNF: `http://gallica.bnf.fr`. The first one is enlightening. English versions are surely available.

[36] Francis Galton, *Regression Towards Mediocrity in Hereditary Stature*, The Journal of the Anthropological Institute of Great Britain and Ireland, Vol. 15 (1886), pp. 246-263. `http://galton.org/essays/1880-1889/galton-1886-jaigi-regression-stature.pdf`

[37] Student, *The Probable Error of a Mean*, Biometrika, **6** (1908) 1

[38] B.L. Welch, *The significance of the difference between two means when the population variances are unequal*, Biometrika **29** (1937) 350

[39] Harold Hotelling, *Some Improvements in Weighing and Other Experimental Techniques*, The Annals of Mathematical Statistics **15** (1944) 297

[40] F.E. Satterthwaite, *An approximate distribution of estimates of variance components*, Biometrics Bulletin **2** (1946) 110

[41] R.T. Cox, *Probability, Frequency and Reasonable Expectation*, American Journal of Physics **14** (1946) 1

[42] Claude E. Shannon, *A Mathematical Theory of Communication*, Bell System Tech. J. **27** (1948) 379 et 623. Available online.

[43] E.T. Jaynes, *Information Theory and Statistical Mechanics*, Physical Review **106** (1957) 620

[44] John E. Shore and Rodney W. Johnson, *Axiomatic Derivation of the Principle of Maximum Entropy and the Principle of Minimum Cross-Entropy*, IEER Transactions on Information Theory **26** (1980) 26; *Properties of Cross-Entropy Minimization*, IEER Transactions on Information Theory **27** (1981) 472.

[45] K. Weise et W. Wöger, *A Bayesian theory of measurement uncertainty*, Meas. Sci. Technol. **3** (1992) 1

Other articles

[46] R. Kacker and A. Jones, Metrologia **40** (2003) 235

[47] Donald W. Marquardt, *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*, J. Soc. Industr. Appl. Math. **11** (1963) 431. Available online.

[48] A. Vivier and J. Aupiais, *Optimization of the decision threshold for single radioactive counting,* Radiochim. Acta **95** (2007) 477

[49] Walter Bich, *How to revise the GUM?*, Accred. Qual. Assur. **13** (2008) 271