International Journal of Molecular and Physical Gastronomy

Courses (Molecular Gastronomy and Education)

How many significant figures should be used for means and standard deviations?

Hervé This vo Kientza[1,2]

1. Université Paris Saclay, INRAe, AgroParisTech, UMR SayFood 0782, 91120 Palaiseau, France
2. Group of Molecular Gastronomy, INRAE-AgroParisTech International Centre for Molecular and Physical Gastronomy, 22 place de l'agronomie, 91120, Palaiseau, France.

* *Correspondence: herve.this@inrae.fr*

Abstract
This course is given to students attending the Educational Module "Physical Chemistry for Food Structuring" of the Master's Program "Engineering – Products - Processes", in AgroParisTech (France). The example of culinary recipes is used to study the question of the precision of the mean and of the standard deviation of a variable for a population and for a subset of a population (used for estimating the characteristics of the population). After examining illustrations of mathematical results concerning these two quantities, the internationally accepted rules for propagating uncertainties are recalled and applied to calculate the number of significant figures in the expression of experimental means and standard deviations.

Keywords
mean, standard deviation, significant figures, population, estimation

Introduction
The following text is part of a university course on methodology in food science, food technology and food engineering. It has been observed that students in these disciplines, even at PhD level and irrespective of the country in which they have studied, know the concepts of mean and standard deviation, but cannot always distinguish between these concepts applied to a population or to a sample extracted from a population in view of estimating the characteristics of the population (there are over 100,000 internet pages on this subject). Moreover they rarely know how to express these values appropriately. In particular, even when they know how to deal with significant numbers, they hesitate about how many digits to display, as shown by numerous articles published on this issue (Harris, 2014; Clymo, 2019; Cousineau, 2020; Quora, 2024; Scribbr, 2024).
Some students do not know that the rules to be

applied are given in the *Guide for the expression of uncertainty in measurement*, from the Bureau International des Poids et Mesures (BIPM, 2008), and very few are able to determine the number of significant figures that can be displayed in measurement results. Sometimes they are taught to apply rules for reaching this goal, and they can indeed apply them, but they do not understand why the rules hold. This course explains slowly the reasons for the choices to be made.
It should be noted that, unlike scientific articles (for which every sentence must be justified by a reference), this course, like all others in the same teaching module, only quotes texts that students are recommended to read (except for culinary recipes whose origin is given).

More precisely, the two questions addressed in this course are as follows:
1. when calculating a mean, how many digits should be given?
2. when calculating a standard deviation, how many digits should be entered?

These questions are widely discussed online and answers of varying quality are proposed (Clymo, 2019), despite the fact that most of the answers concerning uncertainties and measurements are explicitly or implicitly described in the internationally accepted *Guide for the expression of uncertainty in measurement* (GUM), from the Bureau International des Poids et Mesures (BIPM, 2008). The GUM was first published in 1993 by ISO in collaboration with the BIPM, the International Electrotechnical Commission (IEC), the International Federation of Clinical Chemistry (IFCC), the International Union of Pure and Applied Chemistry (IUPAC), the International Union of Pure and Applied Physics (IUPAP) and the International Organization of Legal Metrology (OIML). It (1) describes the practice in the estimation of uncertainty for a broad range of measurements, (2) sets out the concepts required, (3) establishes the general principles, and (4) provides a procedure applicable to cases where an adequate model of the measurement process is available.

The information given in the GUM should be sufficient, but decades of teaching have shown that many students find the GUM difficult to read. This course does not add anything to the GUM, but it does explain it in a culinary context.
First the concept of the mean is examined for culinary recipes. In this particular case, the question of the choice of the number of significant figures is raised (Section 1, below).
Then, after the definition of standard deviation, it is shown how to calculate the mean and the standard deviation (1) for a population and (2) for a sample extracted from the population and used to get an idea of the population.
On this basis, the variability of sample means and standard deviations is considered numerically, in order to make it clear that these values are only orders of magnitude, from which it must be concluded that it would be absurd to display too many figures in their expression (Section 2).
Finally, using the GUM rules (an extract of which is given in Figure 10), it is shown how to determine the significant figures for means and standard deviations. The general solution is given for groups of 3 measurements, as this is common practice for one of common process in laboratory work, namely using a balance to make mass measurements ("gravimetry") (Harris, 2014). Special cases are considered, to show the application of the general formula (Section 3).

1. Introducing the mean and the standard deviation through recipes of pie crust

In this section, we first review the concepts of mean and standard deviation. It is true that they are taught at a early stage at university (Moore and McCabe, 1989), but it will be seen that questions remain about them. In order to introduce them, the example of pie crust recipes is used.
There are many types of pastry, but only the simplest, made with flour, butter, salt and water, are considered here (This, 2010). Cookery

Table 1. The quantities of ingredients given by some classical recipes for pie crust.

Recipe	1	2	3	4	5	6	7	8	9	10
Flour (g)	500	500	500	500	500	500	500	500	6 600	1 500
Butter (g)	250	300	250	180	250	250	350	180	4 000	1 000
Salt (g)	10	0	12	12	12	30	30	30	125	30
Water (or eggs) (g)	200	370 (2)	300	300	160 (3)	ng	20 (2 yolk)	ng	Ng ("some")	700
Sugar (g)		20	13	0	125	0	0	0	0	0

1. Escoffier et al., *1912; 2. Darenne and Duval, 1909 (*pâte à foncer*); 3. Darenne and Duval, 1909 (*pâte à foncer ordinaire*); 4. Darenne and Duval, 1909 (*pâte à foncer commune*); 5. Darenne and Duval, 1909 (pie crust); 6. Favre, 1905; 7. Favre, 1905 ; 8. Favre, 1905 ; 9. Lacam, 1878; 10. Gouffé, 1873. "Ng" stands for "not given".*

books give different recipes for the crusts. For example, in Escoffier *et al.*, 1912), the proportions given for the ingredients are as follows:
- 500 g flour
- 250 g butter
- 10 g salt
- 200 g water.

In another classic French pastry book, Darenne and Duval (1909) indicate different proportions, but also different ingredients, for a "fine pie crust" (*pâte à foncer fine*):
- 500 g flour
- 300 g butter
- 2 eggs
- 20 g sugar
- 250 g water.

However, during a molecular gastronomy seminar (This, 2021), it was observed that pie crusts made with eggs (whole, or yolk only, or white only) could not be distinguished from the same pastry made with water, so this recipe can be simplified. In their book, Darenne and Duval (1909) noted that some pastry chefs use up to 375 g of butter for 500 g of flour, and they also gave another recipe for an "ordinary pie crust" (*pâte à foncer ordinaire*) and for a "common pie crust" (*pâte à foncer commune*). These recipes and others are given in the Table 1.

For a statistical study, the 10 recipes in Table 1 can be considered as a "population", and any element in this population is called - it depends on the particular community describing it - an "element" (using the language of set theory and probability), or a "unit" (more in the statistic circles), or an "experimental unit", or an "object", or an "individual", or a "member" (Moore and MacCabe, 1989; Evans and Rosenthal, 2006). For the pie crust analysis, each element (that is: each particular recipe) is characterised by many variables (such as the quantity of butter, or the quantity of flour), so that a complete representation requires more than the three spatial dimensions, and no visual display can correspond to the table.

However the description can be reduced to a single dimension if salt and sugar are not taken into account and if the amount of water is considered irrelevant, as it depends on the particular quality of the flour (which is not addressed in the recipes). Using these simplifying assumptions, the recipes can be described by the amount of butter in g per 500 g

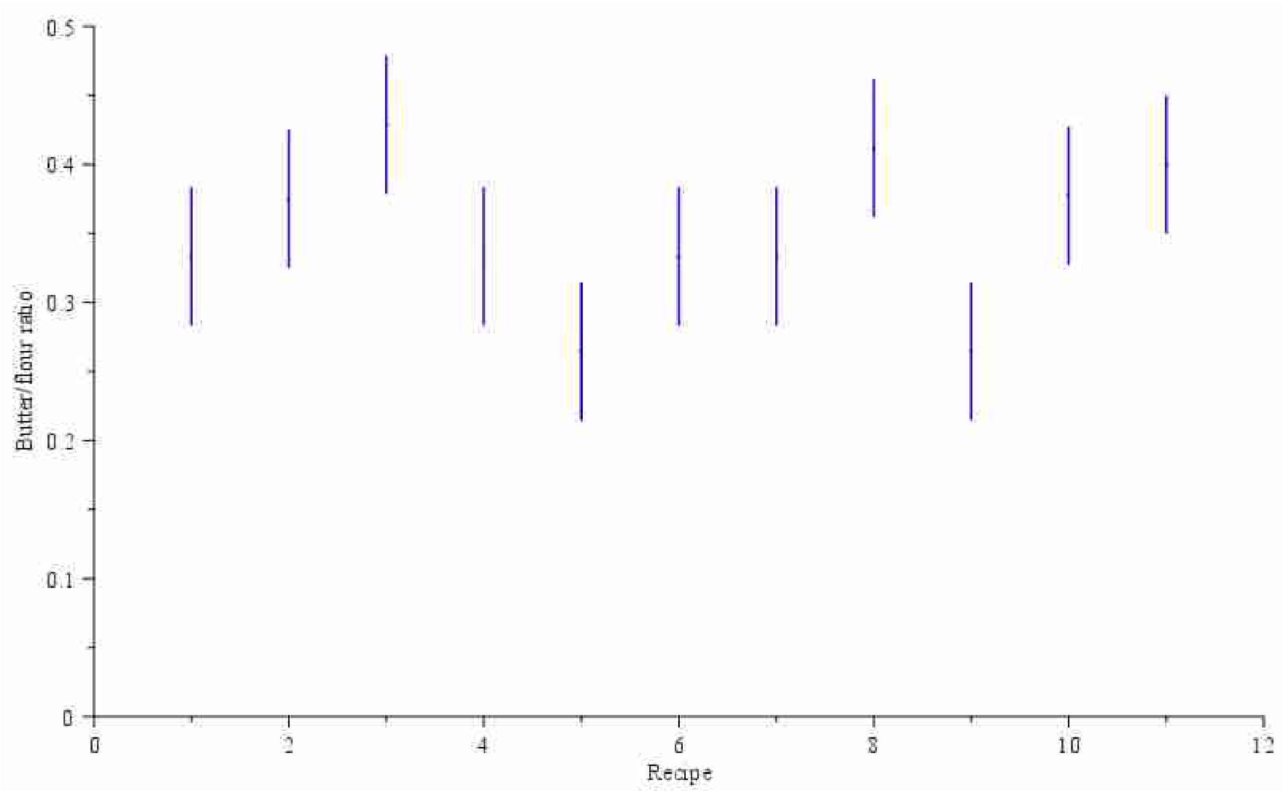

Figure 1. Butter proportions for some classic pie crust recipes. The data points are displayed twice: in blue, the lines take into account the possible variations associated to a number or digits equal to 2. In red, in the middle of these blue lines, the excessively small red dot represent the values with 15 digits (they are barely visible, because too small).

of flour (250, 300, 375, 250, 180, 250, 250, 350, 180, 303, 333), so that the butter/flour ratios are: 250/750, 300/800, 375/875, 250/750, 180/680, 250/750, 250/750, 350/850, 180/680, 303/803, 333/833.

A calculator displaying (for example) 10 digits would find these ratios equal to: 0.3333333333, 0.3750000000, 0.4285714286, 0.3333333333, 0.2647058824, 0.3333333333, 0.3333333333, 0.4117647059, 0.2647058824, 0.3773349938, 0.3997599040. In this list, the rounding of ratios has been done correctly (NIST, 2019; Yale, 2023), but the number of digits (10) is arbitrary, and a more meaningful approach should be based on:

- the uncertainties about the mass of the ingredients,
- or the possible perception of differences in flavour: if no difference is perceived with pie crusts for which the butter/flour ratio differs only by the second digit, for example, it is proper to avoid this digit (as long as it is explained why this is done). Finally it has to be added that for the graphical display of ratios (Kamat *et al*., 2014), the size of data points must be chosen according to the number of digits: in Figure 1, the size of the blue bars is based on the use of the arbitrary (and perhaps bad) choice of only two digits.

While the mean quantitatively locates the most likely value of the variable for the population, the standard deviation measures the extent of the distribution of the values of the variable around the mean. Here, using a 15-digit calculator, we could calculate a standard deviation equal to 0.0546393131 for the 10 recipes. But once again, this number of digits is arbitrary and meaningless, especially considering that quantities in pastry-making are rarely known to an accuracy of less than 1 g. Mentally, many people would prefer to write 0.05 for this standard deviation, but are they right? Should they instead write 0.054? or 0.055? or 0.0546? Here again, a reason is needed to decide, but before considering it, *in silico* experiments will be examined to better appreciate the value of averages and standard deviations for samples.

2. The diversity of means of samples, and the high diversity of estimated standard-deviations (for samples)

What is given - with rigorous demonstrations - in statistics courses will not be repeated here (Moore and McCabe, 1989; Evans and Rosenthal, 2006), but a numerical exploration of the question will be made, in order to convey two main ideas (one for the mean and the other for the standard deviation) that will have to be kept in mind when solving the issue of the significant digits of the mean and of the standard deviation for samples.

2.1. For the mean

Here the difference between a population and a sample is first needed. A population P is a set of elements, which will be noted I_i, $i = 1..n$ (Figure 2) . Each element i can be characterized by the values x_i, y_i, z_i, etc. of the variables x, y, z, etc. For pie crust recipes, P could be a library, and the

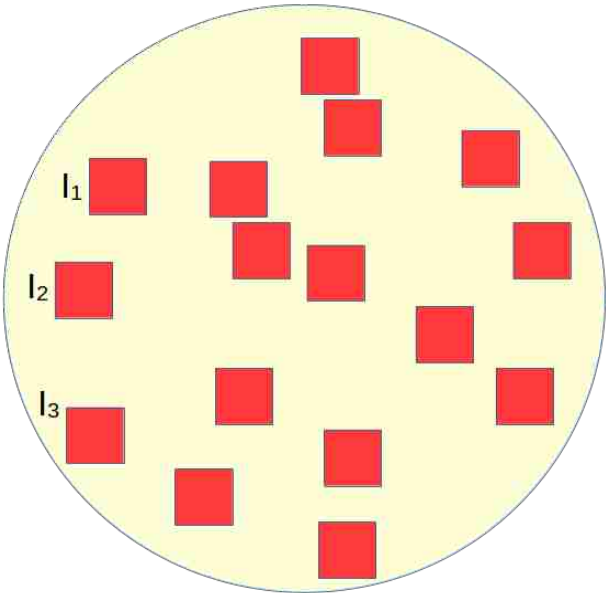

Figure 2. A population is a set, with a finite or infinite number of elements (I_1, I_2, I_3, etc.).

elements would be particular recipes of pie crusts from books in that library. For each recipe, the variables could be the amount of flour, butter, water, salt, sugar, etc.
In order to describe the population P, the data can be organized in a table (Table 2). However such a table can become very cumbersome when the number of elements and the number of variables are large. The mean and the standard deviation are used to describe the population more concisely.
For calculating the arithmetic mean of a variable (for example the butter / flour ratio r) for the elements (recipes) in the population (all culinary

Table 2. The data for a population P.

Elements (recipes)	Variable x *(for example, the quantity of flour)*	Variable y *(for example, the quantity of butter)*	Variable z *(for example, the quantity of sugar)*	Etc.
I_1	x_1	y_1	z_1	
I_2	x_2	y_2	z_2	
…	…	…	…	
I_n	x_n	y_n	z_n	

books in the library), the values taken by the variable r are summed and divided by the number n of elements of the population:

$$\mu=\frac{\sum_{i=1}^{n} r_i}{n} \quad \text{(Equation 1).}$$

The mean, also known as the average, is the point that one would see from a distance on an axis on which the values of the butter / flour ratios r would be displayed (Figure 3).

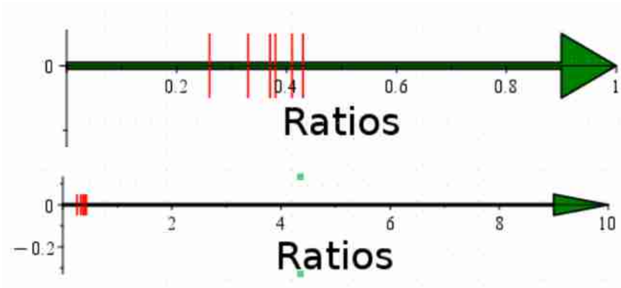

Figure 3. Top*: Results of the different butter / flour ratios from a particular library as displayed on an axis. The mean would appear as one dot when looked from far away (*bottom*).*

In science and technology, the question is often to find a relationship between the mean of the values of a variable for a sample and the mean for the population from which that sample is taken (Moore and McCabe, 1989). Here, for example, the question would be to examine only 3 recipes from the library in order to get an idea of the butter / flour ratio in all recipes of pie crusts from the books in the library.

This question is omnipresent in scientific and technological practice. One of the most common examples is when determining the mass of an object using a balance. Even if the balance has a high precision (for example, giving measurements to the nearest hundredth of a milligram), the real mass of the object weighed cannot be known because of the many causes of error (electronic noise, vibrations, disturbances in the ambient air, etc.) or the limited accuracy of the balance (Harris, 2014).

Students often ask why it is impossible to find the true mass, and the following answer is can be given. Suppose that the accuracy of the balance is 0.001 g and that a value 12.304 g is displayed: this means that the mass can be between 12.3035 g and 12.3045 g. The number of real numbers in the real interval [12.3035-12.3045] is infinite, so that the probability that the true mass would be exactly 12.304 (that is 12.304000… with an infinite numbers of 0) is nil (remember that the probabilities are defined by the ratio of the number of favourable cases by the number of possible cases).

This being established, it can be observed that by repeating the measurements n times with the same balance, a series of values $\{m_i\}$ is obtained, where i is an integer between 1 and n. For this set, the arithmetic mean of the mass values displayed by the balance can be calculated as it was done for the butter/flour ratio. This average value is different from the true (unknown) mass M, and it is also likely to be different from the mean of the masses measured after another group of measurements of the same mass.

The question that we want to illustrate now numerically is to see how the samples means differ from each other and from the mean for the whole population. To study this more quickly than using a real balance, an *in silico* reproduction of mass measurements is proposed: a normal population (of the results that the balance would give) of masses m_i is first created with a mean (for the population) μ equal to 100, and a standard deviation of 1 (reproducing the fluctuations of a scale, admittedly imprecise for this example).

As it would be done in a chemistry laboratory, the object is "weighed" 3 times: this means that 3 values are taken at random from the normal population (in *Maple* language, this can be done using the "*rand*" function) (Maple, 2024); the mean (for the sample) of the 3 values is then calculated. If another group of three "digital mass measurements" is extracted from the same normal population, another sample mean is calculated. And so on: using a small

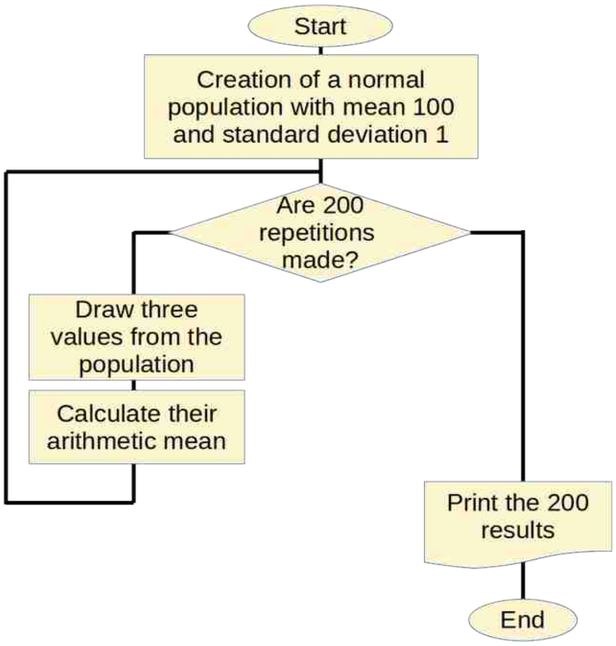

Figure 4. The flowchart for making the figure 5.

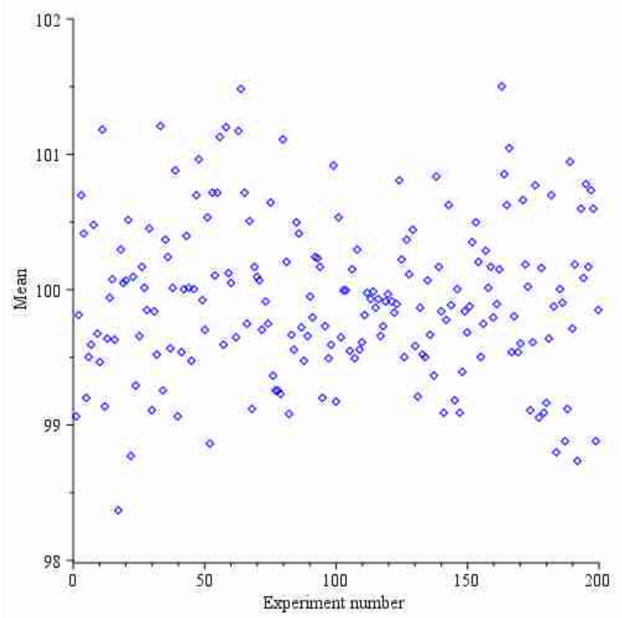

Figure 5. The 200 sample means calculated after drawing randomly 3 samples from a normal population with mean 100 and standard deviation 1.

computer program (Figure 4), this process of "weighing three times" can be reproduced systematically, for example 200 times.
The results of such an experiment are shown in Figure 5: all the means are (of course) around 100, but they are mainly distributed between 99 and 101. Statistics courses can describe this more formally (Reif, 1967; Moore and McCabe, 1989), but the aim here is different: it is only to give an idea of the precision needed to express the mean. Note here that as the interval [99 - 101] is much larger than the precision given by the calculator (10^{-15}), there is no point in using 15 digits (unless when an explicit goal requiring them is set).
These first results can be compared with what would be obtained if the means for 6 "digital mass measurements" were calculated, again for 200 experiments, or if 15 "digital mass measurements" were made for each sample (Figure 6): the distribution of sample means narrows around the population mean (let us repeat: 100, fixed by the construction of the example).

2.2. For the standard deviation

For the standard deviation, the same *in silico* experiment can be carried out as for the mean, which leads to the following idea: the standard deviation for samples is only an estimate of the order of magnitude of the dispersion of the data.
But more details are needed now. First of all, it is important to distinguish the standard deviation for a population and the that for a sample used to estimate the characteristics of the population (Moore and McCabe, 1989). It must be stressed: a group of three mass measurements can be considered as a group in itself (a population), or as a sample that is used for estimating the unknown characteristics of a whole, larger, population from which this sample is extracted. First, when the group is considered as a population in itself, the expression of the

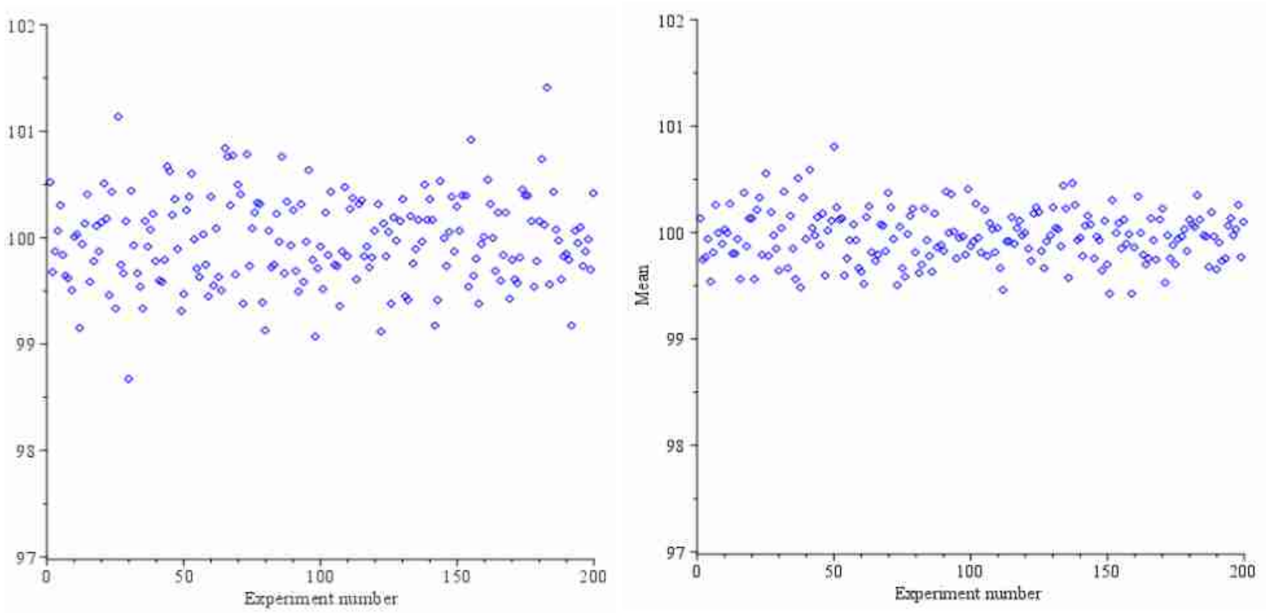

Figure 6. The bigger the samples, the closer their mean is from the mean for the population. Left and right, the means for 200 groups of 6 and 15 measurements respectively are extracted from the same population, as in Figure 5.

variance *v* is the expected value of the squared deviation from the mean of the masses relative to the average mass $<m>$ for the group of three values:

$$v=\frac{1}{3}\sum_{i=1}^{3}\left(m_i-<m>\right)^2 \quad \text{(Equation 2).}$$

For this (small) population, the standard deviation is :

$$\sigma=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(m_i-<m>\right)^2} \quad \text{(Equation 3).}$$

Here, the denominator used is *n*, because we consider this group as a population that we want to characterize as a population, independently of the whole population of possible outcomes of measurements.

But, as said above, this small group of 3 measurements can be also used to estimate the standard deviation Σ of the whole population from which the 3 masses are extracted (the mean of this larger population would be *M)*. In this case, the Equation 3 would be "biased", and a "good" estimator of the standard deviation Σ of the large population is rather:

$$sd=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(m_i-\bar{m}\right)^2} \quad \text{(Equation 4),}$$

with *n* = 3 in the particular example chosen. In this equation, $\bar{m}$ is the mean for the sample (different from the unknown mean *M* for the whole population).

Having established this, the previous *in silico* experiment can be repeated, again using a normal population of average value equal to 100 and a standard deviation equal to 1. From this population, 3 samples are chosen at random 200 times and, for each group of 3 "digital mass measurements", the estimated standard deviation is calculated. The results are shown in Figure 7.

The variability of the estimated standard deviations is high. Of course, they are all around the standard deviation of the population (remember that this was chosen to be equal to 1), but it appears clearly that one estimated standard deviation chosen at random can only

give an order of magnitude of the value that we want to estimate: consequently, it makes no sense writing many digits to express a standard deviation used for estimating the standard deviation of a population.
Once this result has been obtained, estimated standard deviations for larger samples can be calculated: 6 measurements in Figure 8 top, and 15 in Figure 8 bottom. The larger the samples, the closer the estimated standard deviations are to the standard deviation of the population.

3. Significant figures for means and standard deviations

The objective initially set can now be achieved: to formally determine how many significant digits should be displayed for sample means and estimated standard deviations. This question arises each time measurements are made, but many people hesitate over the answer. In this paragraph, simple examples will be analysed, using official documents, before giving a clear answer to the question.
When there is some uncertainty in the measurements, such as when using a balance to determine mass, or when using a sample of recipes to find the butter /flour ratio to use in pie crust recipes, the last significant figure should be an estimate of the uncertainty of the measurement (Bell, 1999; BIPM, 2008). The definitions given in Figure 10 are taken from the internationally shared *Guide to the expression of uncertainty in measurement* (GUM) of the Bureau International des Poids et Mesures (BIPM, 2008). In particular, the last significant digit given in a value should indicate the uncertainty of the value.
Before doing the maths, let us observe that if you search the internet, you will find (1) pages that give rules without explanation nor justification (for example, Chem21 labs, 2024; LibreTexts Chemistry, 2024), (2) pages that give erroneous or dubious indications (for example, Sribbr, 2024), (3) pages that give so many different indications that readers are lost (Quora, 2024), (4) some interesting articles that nevertheless forget to refer to the international conventions (Cousineau, 2020).
As mentioned above, a recommendable strategy is to understand what one is doing rather than blindly following rules, which is why the solution to the rounding problem should be done using the GUM. A (very common) special case will be examined first, before calculating other examples, and ending with the description of the pie crust.

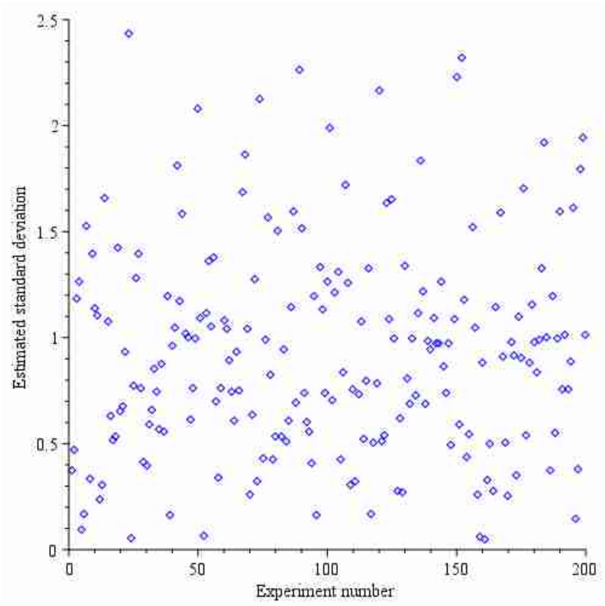

Figure 7. The estimated standard deviations for groups of 3 mass measurements can be very different from one another, and also from the standard deviation (equal to 1) for the population from which these samples were extracted: this means that the estimated standard deviation for a small sample is only an order of magnitude of the standard deviation for the population.

3.1 A general solution in the particular case of three measurements

In this paragraph, the example of mass measurements is used to apply the information from the GUM, in order to determine the number of digits for the mean and the estimated standard deviation.

Let us assume that we use a balance to measure the mass of an object, and repeat this three times, giving the values m_1, m_2, m_3 known with respective uncertainties Δm_1, Δm_2, Δm_3. The mean (*mean*) and the estimated standard deviation (*sd*) for this sample are defined by:

$$mean = \frac{1}{3}\left(m_1 + m_2 + m_3\right) \quad \text{(Equation 5)},$$

and:

$$sd = \sqrt{\frac{1}{2}\left(\left(m_1 - mean\right)^2 + \left(m_2 - mean\right)^2 + \left(m_3 - mean\right)^2\right)} \quad \text{(Equation 6)}$$

The mean and the estimated standard deviation being functions of the three variables m_1, m_2, m_3, the uncertainties can be propagated using the rules given by the GUM. If the same scale and the same measurement conditions are used for the three mass determination, the three uncertainties Δm_1, Δm_2, Δm_3 can be considered to be equal, and written Δm (the calculation with different uncertainties would be only slightly more complex). According to the GUM, the uncertainty for a function $f(x_1, x_2, ..., x_n)$ with real values of n variables $x_1, x_2, ..., x_n$, known respectively with uncertainties Δx_1, Δx_2, ..., Δx_n is:

$$\Delta f\left(x_1, x_2, \ldots, x_n\right) = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2 \Delta x_i^2} \quad \text{(Equation 7)},$$

where the $\left(\frac{\partial f}{\partial x_i}\right)$ are the partial derivatives of the function f relative to the real variable x_i.

In the case of mass measurements, applying the equation 7 to the mass measurements gives the uncertainty $\Delta mean$:

$$\Delta mean = \sqrt{\sum_{i=1}^{3}\left(\frac{\partial mean}{\partial m_i}\right)^2 \Delta m_i^2} \quad \text{(Equation 8)}.$$

In order to calculate it, the partial derivatives have to be calculated first. For example for m_1:

$$\frac{\partial mean}{\partial m_1} = \frac{\partial\left[\frac{1}{3}\left(m_1 + m_2 + m_3\right)\right]}{\partial m_1} = \frac{1}{3} \quad \text{(Equation 9)}.$$

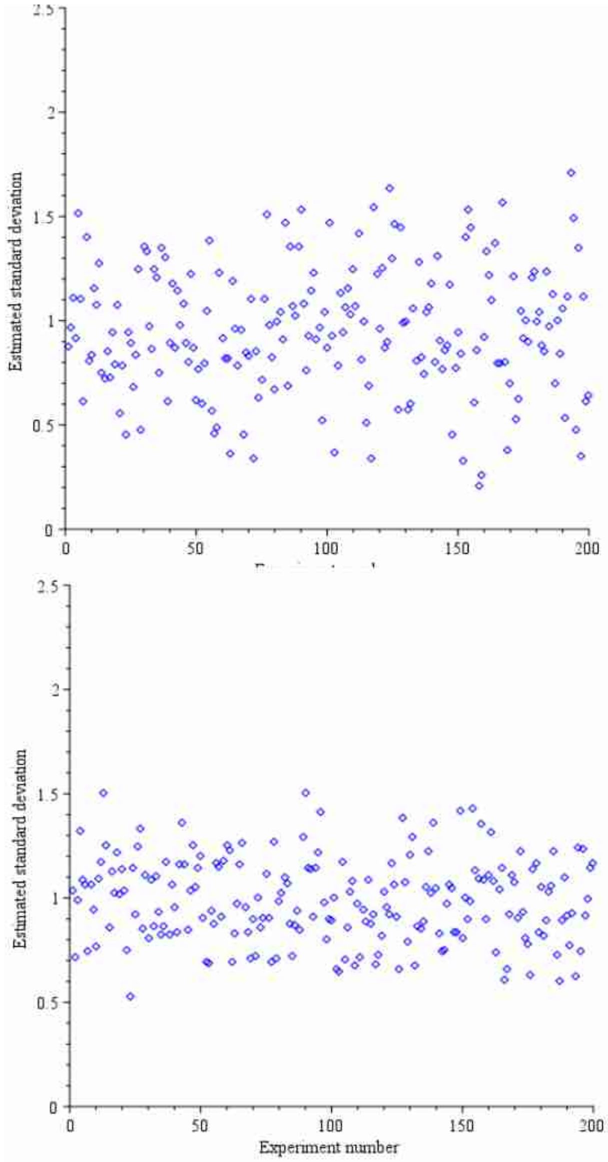

*Figure 8. Display of 200 estimated standard deviations for samples of 6 individuals (*top*), and 200 standard deviations for samples of 15 individuals (*bottom*).*

The result is the same for the two other variables m_2 and m_3. So that the uncertainty of the mean is:

$$\Delta mean = \sqrt{\left(\frac{1}{3}\right)^2 . \Delta m^2 + \left(\frac{1}{3}\right)^2 . \Delta m^2 + \left(\frac{1}{3}\right)^2 . \Delta m^2} = \frac{1}{\sqrt{3}} \Delta m$$

(Equation 10).

As this uncertainty if of the same order of magnitude as the uncertainties on masses, the mean should have the same number of significant figures as the measurement results.

For the estimated standard deviation, the calculation is of the same kind. First the definition of the standard deviation has to be developed in the particular case that is considered:

$$sd=\sqrt{\frac{1}{2}\left(\left(m_1-mean\right)^2+\left(m_2-mean\right)^2+\left(m_3-mean\right)^2\right)}$$ (Equation 11).

This can also be written as:

$$sd=\sqrt{\frac{1}{2}\left(\left(m_1-\frac{1}{3}\left(m_1+m_2+m_3\right)\right)^2+\left(m_2-\frac{1}{3}\left(m_1+m_2+m_3\right)\right)^2+\left(m_3-\frac{1}{3}\left(m_1+m_2+m_3\right)\right)^2\right)}$$ (Equation 12).

As the estimated standard deviation *sd* is a function of the three masses m_1, m_2, m_3, the partial derivatives have to be calculated, finding the equation given in Figure 9.

As for the mean, this expression indicates the number of significant figures that one can display: all figures have to be certain, except the last one. The result will now be explored in particular cases.

3.2. Particular cases and the solution for the pie crust recipes

Example 1 : a small difference between the measurements, and a small uncertainty.

Let us consider the three values: 15.333, 15.334, 15.335. In order to calculate the number of significant figures for the estimated standard deviation, one uses the formula that was found, replacing the letters by the data. An estimated standard deviation is found equal to 10^{-3}, and an uncertainty equal to 7.07 10^{-4}, whose order of magnitude is 10^{-3}. It means that the standard deviation should have only one digit.

Example 2: a large difference for the data, with a small uncertainty:

$$\Delta sdm := \left(\frac{\left(\frac{2\,m1}{3}-\frac{m2}{3}-\frac{m3}{3}\right)^2 \Delta m^2}{4\left(\frac{\left(\frac{2\,m1}{3}-\frac{m2}{3}-\frac{m3}{3}\right)^2}{2}+\frac{\left(\frac{2\,m2}{3}-\frac{m1}{3}-\frac{m3}{3}\right)^2}{2}+\frac{\left(\frac{2\,m3}{3}-\frac{m1}{3}-\frac{m2}{3}\right)^2}{2}\right)} + \frac{\left(\frac{2\,m2}{3}-\frac{m1}{3}-\frac{m3}{3}\right)^2 \Delta m^2}{4\left(\frac{\left(\frac{2\,m1}{3}-\frac{m2}{3}-\frac{m3}{3}\right)^2}{2}+\frac{\left(\frac{2\,m2}{3}-\frac{m1}{3}-\frac{m3}{3}\right)^2}{2}+\frac{\left(\frac{2\,m3}{3}-\frac{m1}{3}-\frac{m2}{3}\right)^2}{2}\right)} + \frac{\left(\frac{2\,m3}{3}-\frac{m1}{3}-\frac{m2}{3}\right)^2 \Delta m^2}{4\left(\frac{\left(\frac{2\,m1}{3}-\frac{m2}{3}-\frac{m3}{3}\right)^2}{2}+\frac{\left(\frac{2\,m2}{3}-\frac{m1}{3}-\frac{m3}{3}\right)^2}{2}+\frac{\left(\frac{2\,m3}{3}-\frac{m1}{3}-\frac{m2}{3}\right)^2}{2}\right)} \right)^{1/2}$$

Figure 9. The full expression for the uncertainty of the estimated standard deviation.

The objective of a measurement (B.2.5) is to determine the value (B.2.2) of the measurand (B.2.9), that is, the value of the particular quantity (B.2.1, Note 1) to be measured. A measurement therefore begins with an appropriate specification of the measurand, the method of measurement (B.2.7), and the measurement procedure (B.2.8).

In general, the result of a measurement (B.2.11) is only an approximation or estimate (C.2.26) of the value of the measurand and thus is complete only when accompanied by a statement of the uncertainty (B.2.18) of that estimate.

Uncertainty (of measurement): parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.
NOTE 1: The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.

Combined standard uncertainty: standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighted according to how the measurement result varies with changes in these quantities.

A Type A standard uncertainty is obtained from a probability density function (C.2.5) derived from an observed frequency distribution (C.2.18), while a Type B standard uncertainty is obtained from an assumed probability density function based on the degree of belief that an event will occur [often called subjective probability (C.2.1)]. Both approaches employ recognized interpretations of probability.
NOTE: Type B evaluation of an uncertainty component is usually based on a pool of comparatively reliable information (see 4.3.1).

The standard uncertainty of the result of a measurement, when that result is obtained from the values of a number of other quantities, is termed combined standard uncertainty and denoted by μ_c . It is the estimated standard deviation associated with the result and is equal to the positive square root of the combined variance obtained from all variance and covariance (C.3.4) components, however evaluated, using what is termed in this Guide the law of propagation of uncertainty (see Clause 5).

Each input estimate x_i and its associated standard uncertainty $u(x_i)$ are obtained from a distribution of possible values of the input quantity X_i . This probability distribution may be frequency based, that is, based on a series of observations $X_{i,k}$ of X_i , or it may be an a priori distribution. Type A evaluations of standard uncertainty components are founded on frequency distributions while Type B evaluations are founded on a priori distributions. It must be recognized that in both cases the distributions are models that are used to represent the state of our knowledge.

Type A evaluation of standard uncertainty: In most cases, the best available estimate of the expectation or expected value μ_q of a quantity q that varies randomly [a random variable (C.2.2)], and for which n independent observations q_k have been obtained under the same conditions of measurement (see B.2.15), is the arithmetic mean or average $\bar{q}$ (C.2.19) of the n observations:

$$\bar{q}=\frac{1}{n}\sum_{k=1}^{n} q_k$$

Thus, for an input quantity X_i estimated from n independent repeated observations $X_{i,k}$, the arithmetic mean $\bar{X}_i$ obtained from Equation (3) is used as the input estimate x_i in Equation (2) to determine the measurement result y; that is, $x_i=\bar{X}_i$. Those input estimates not evaluated from repeated observations must be obtained by other methods, such as those indicated in the second category of 4.1.3.
The individual observations q_k differ in value because of random variations in the influence quantities, or random effects (see 3.2.2). The experimental variance of the observations, which estimates the variance σ^2 of the probability distribution of q, is given by:

$$s^2(q_k)=\frac{1}{n-1}\sum_{k=1}^{n}(q_j-\bar{q})^2$$

Figure 10. An extract from the GUM (BIPM, 2008).

Let us consider the values 155.323, 151.401, 148.577. Here, we calculate an estimated standard deviation equal to 3.39, but an uncertainty of 7.07 10^{-2}, *i.e.* about 10^{-1}. Here one can give two significant figures for the standard deviation: 3.4.

Example 3 : a small difference between data, and large uncertainty.
Let us have the values 219.2, 220.1, 200.1. One calculates a standard deviation equal to 11, and an uncertainty of 7.07 10^{-2}, which makes 3 significant digits for the estimated standard deviation.

Example 4:
Simple reasoning and applying the official definitions can avoid lengthy searches on the internet, about how practically the mean and the estimated standard deviation can be calculated, instead of using arbitrary rules. As a final application, we propose to observe that in Harris (2014), an often quoted textbook for analytical chemistry, it is written about digits in a sum : "If the numbers being added do not have the same number of significant figures, we are limited by the least-certain one. For example, the molecular mass of KrF_2 is known only to the third decimal place, because we only know the atomic mass of Kr to three decimal places". And the author considers, as an example, the sum 18.998 403 2 +18.998 403 2 + 83.798, with the result 121.794 806 4, of which it is said: "The number 121.794 806 4 should be rounded to 121.795 as the final answer." For sure, the Harris textbook is a good one, but why this rule? From what was given in this article, the Harris rule can be understood, observing that the sum is a function of two variables that can be differentiated:

$$df(x,y)=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy \quad \text{(Equation 13).}$$

Moving to uncertainties, one can write (GUM, 2008):

$$\Delta f=\sqrt{\left(\left(\frac{\partial f}{\partial x}\right)^2\Delta x^2+\left(\frac{\partial f}{\partial y}\right)^2\Delta y^2\right)} \quad \text{(Equation 14).}$$

Here, the sum S can be written:

$$S=2a+b \quad \text{(Equation 15).}$$

So that:

$$\Delta S=\sqrt{\left((2)^2\Delta a^2+(1)^2\Delta b^2\right)} \quad \text{(Equation 16).}$$

The significant figures are given by this uncertainty. For the example from Harris (2014), one would calculate an uncertainty 0.001000002000, which is almost the greatest of the uncertainties for the terms of the sum: no "rule" was to be applied. More generally, one would be wise to recall that the scientific activity is not based on rules, but rather on understanding deeply what one does.

Example 5:
Finally, the correct expression of the butter to flour ratios for pie crust recipes can be found. Remember that two kinds of expressions were envisioned:
- the uncertainty of the mass determination of the ingredients,
- or the possible perception of differences of flavour.

For the first case, ratios $r=\frac{m_b}{m_f}$ are considered: they are functions of two variables m_b (mass of butter) and m_f (mass of flour), so that the uncertainty on r is:

$$\Delta r(m_b,m_f)=\sqrt{\left(\frac{\partial r}{\partial m_b}\right)^2\Delta m_b^2+\left(\frac{\partial r}{\partial m_f}\right)^2\Delta m_f^2} \quad \text{(Equation 17)}$$

Here, $\frac{\partial r}{\partial m_b}$ is simply equal to $\frac{1}{m_f}$, while $\frac{\partial r}{\partial m_f}$ is equal to $\frac{-m_b}{m_f^2}$. Hence the final expression:

$$\Delta r(m_b,m_f)=\sqrt{\frac{1}{m_f}^2\Delta m_b^2+\left(\frac{-m_b}{m_f^2}\right)^2\Delta m_f^2} \quad \text{(Equation 18)}$$

In the kitchen, using a measuring cup for taking flour, the precision is about 20 g, and for butter a kitchen scale can have a precision of 5 g. Using these values, it appears (please do the calculation) that only the first digit (3) is significant: the average butter/flour ratio is 0.3.
About the differences in perception, a literature survey did not give any answer, so that the experiment remains to be done: can consumer perceive the difference between two pie crusts with 0.300 and 0.301 butter/flour ratio? Between two pie crusts with 0.30 and 0.31 butter/flour ratio? Between two pie crusts with 0.3 and 0.4 butter/flour ratio?

4. Conclusion

In the past, applying the GUM rules was somehow cumbersome, because in practical cases, such as laboratory experiments, the propagation of uncertainties needs many steps, sometimes with long analytical expressions (remember Equation 11 for a simple case). Nowadays, software for formal calculation makes it easier. Anyway, the more important conclusion of this course is: apply rules that you understand.

References

Bell S. 1999. *A Beginner's Guide to Uncertainty of Measurement*, Crown, ISSN 1368-6550.

BIPM. 2008. Évaluation des données de mesure. Guide pour l'expression de l'incertitude de mesure, *JCGM* 100:2008 (F).

Chem21 labs. 2024. Significant figures https://www2.chem21labs.com/labfiles/jhu_significant_figures.pdf, last access 2024-02-20.

Clymo RS. 2019. How many of the digits in a mean of 12.3456789012 are worth reporting?, *BMC Research Notes*, 12, 148, 1-5. DOI: 10.1186/s13104-019-4175-6

Cousineau D. 2020. How many decimals? Rounding descriptive and inferential statistics based on measurement precision, *Journal of Mathematical Psychology*, 97 (2020) 102362. https://doi.org/10.1016/j.jmp.2020.102362.

Darenne E, Duval E. 1909. *Traité de pâtisserie moderne*, Flammarion, Paris.

Escoffier A, Gilbert Ph, Fetu E. 1912. *Guide culinaire*, Flammarion, Paris.

Evans MJ, Rosenthal JS. 2006. *Probability and Statistics: The Science of Uncertainty*, WH Freeman and Co, New York (NY).

Favre J. 1905. *Dictionnaire universel de cuisine pratique*, Guénégaud, Paris.

Gouffé J. 1873. *Livre de pâtisserie*, Hachette, Paris.

Harris DC. 2014. *Quantitative Chemical Analysis*, W. H. Freeman, New York, USA.

Kamat P, Harland GV, Schatz GC. 2014. Graphical excellence, *The Journal of physical chemistry Letters*, 5, 2018-2020.

Lacam P. 1878. *Le Nouveau pâtissier-glacier français et étranger*, Mennessier, Paris.

Libre Texts Chemistry. 2024. Uncertainties in quantities and error propagation, https://chem.libretexts.org/Courses/Hope_College/General_Chemistry_Labs/How-To_Guides/Dealing_with_Quantities_and_their_Errors/Uncertainties_in_Quantities_and_Error_Propagation , last access 2024-02-20.

Maple. 2024. Rand, https://www.maplesoft.com/support/help/maple/view.aspx?path=rand .

Moore DS, McCabe GP. 1989. *Introduction to the practice of statistics*, W.H. Freeman and Company, New York, USA.

NIST. 2019. Good Laboratory Practice for Rounding Expanded Uncertainties and Calibration Values, https://www.nist.gov/document/glp-9-rounding-20190506pdf.

Quora, 2024. How many significant figures should standard deviation have, https://www.quora.com/How-many-significant-figures-should-standard-deviation-have, last access 2024-02-20.

Reif F. 1967. *Berkely Physics Course: Statistical physics*, McGraw Hill, New York, USA.

Sribbr. 2024. How many decimal places do you use in APA style.

https://www.scribbr.com/frequently-asked-questions/how-many-decimal-places-apa-style/ , last access 2024-02-20.

This H. 2010. *Cours de gastronomie moléculaire N°2 : les précisions culinaires*, Quae/Belin, Paris.

This H. 2021. A propos de pâtes, https://hervethis.blogspot.com/2018/03/a-propos-de-pates-tarte.html.

Yale. 2023. A short guide for significant figures, http://www.astro.yale.edu/astro120/SigFig.pdf.

Edited by:
Dr. Róisín Burke, School of Culinary Arts and Food Technology, TU Dublin, Ireland.

Reviewers:
1. Dr. Seán Lacey, Department of Mathematics, Munster Technological University, Bishopstown, Cork, Ireland.

2. Anonymous

Received:
31 February 2024

Accepted :
24 Septembre 2024

Published:
31 October 2024.

Section:
This article was published in the "Courses" section (Part II: Molecular Gastronomy and Education) of the *International Journal of Molecular and Physical Gastronomy*
.

Citation:
This vo Kientza H. 2024. How many significant figures should be used for means and standard deviations?, *International Journal of Molecular and Physical Gastronomy,* 10(1), 8, 1-15. doi: 10.17180/ijmpg-2024-art08