A battery-resistor analogy for further insights on measurement uncertainties

To cite this article: Gabriel L A de Sousa and George C Cardoso 2018 *Phys. Educ.* **53** 055001

View the article online for updates and enhancements.

Related content

- The autocorrelation function to characterize time series
  Thomas J Witt

- Approach to determine measurement uncertainty in complex nanosystems with multiparametric dependencies and multivariate output quantities
  B Hampel, B Liu, F Nording et al.

- Assessment of measurement efficiency in laser- and phase-Doppler techniques: an information theory approach
  Miguel R Oliveira Panão

A battery-resistor analogy for further insights on measurement uncertainties

Gabriel L A de Sousa and George C Cardoso

Department of Physics, FFCLRP, Universidade de São Paulo, Av. Bandeirantes 3900, Ribeirão Preto, SP, 14040-901, Brazil

E-mail: gcc@usp.br

CrossMark

Abstract

We use analogies to provide introductory laboratory students intuition into measurement uncertainties. Using a battery-resistor circuit we discuss uncertainty concepts and derive expressions for uncertainty of the mean and sums of uncertainties. Finally, we draw attention to the fact that the interpretation of standard deviation as uncertainty depends on the statistical distribution of the data, while the interpretation of uncertainty of the mean is largely insensitive to such distribution, especially for large samples. If the resistor in the battery-resistor circuit is a resistive loudspeaker, the uncertainty or noise is literally the acoustic power of the sound produced by such noise. In the sound analogy the statistical distribution of the noise is related to its timbre.

1. Introduction

Uncertainty analysis is probably one of the least appreciated aspects of the introductory laboratory among students [1]. Students need to determine uncertainty, understand how to add uncertainties and how repeated measurements or improvements in methodology reduce uncertainty. Despite good books available, e.g. [2], even a plausibility level of understanding the expressions used in experimental uncertainties analysis is normally not accessible without mathematics that is above the scope of an introductory course. Giving physical intuition into uncertainty calculations to students would help increase their level of comfort with the subject [3]. Good progress has been made to make measurement and uncertainties more accessible to students [1, 4–7]; the approaches range from fun activities to nail down the concepts [1] to deeper discussions into 'point' versus 'set' paradigms [7]. However, adequate insight into

the mathematical formulas used is left for future courses such as statistics and statistical physics, usually without a laboratory context. This scenario often prevents the development of a mental picture on the connection between the measured quantity, the 'set of measurements' and respective uncertainty analysis and reporting [8]. Many students are left under the impression that calculation and reporting of experimental uncertainties is just busywork, not steps towards interpretation of results or further refinements in the experimental system and method.

Measurement is, by definition, the measurement of the mean value [9]. In a typical experimental situation we want to *measure* a quantity but only have access to a limited set of readings of the quantity whose true mean (the expected value) we want to estimate. Each given reading in the set most likely differs from the expected value due to random measurement disturbances and to
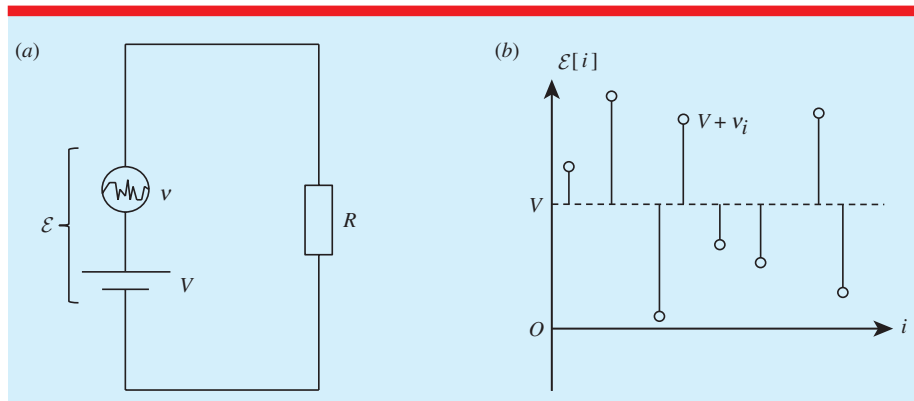
Figure 1. (a) Circuit with battery, noise source and ideal resistive loudspeaker; (b) voltage values $\mathcal{E}$ on the resistor for successive measurements, showing the ith voltage reading. R and V are constants.

intrinsic fluctuations in the measurement process. Therefore we need to determine both how large this noise is, and the uncertainty it causes in the determination of the mean value as calculated from the finite set of readings.

Traditionally the words uncertainty and error have been used interchangeably. However, the new guide for uncertainty measurement (GUM) terminology [9, 10] recommends the use of the word error only for discrepancies in the expected value due to measurement methodology. Otherwise, the GUM recommended word is uncertainty. Uncertainties are defined as type A (determined by repeating measurements) or type B (determined from available knowledge about the measurement system and method). This paper focuses, without loss of generality, on type A uncertainties. For the sake of simplicity, we will assume that the expected values of experimental errors and type B uncertainties are zero.

In this article we propose a thought experiment where we heuristically introduce the notions of mean, standard deviation, uncertainty of the mean and illustrate the need for uncorrelated uncertainties. A simple battery-resistor circuit is used for analogies that help with developing the concepts. Analogies create narratives that stimulate interest and help forming mental models [11–13]. Using the circuit we derive expressions for sums and differences of uncertainties and determine reduction of uncertainty with increased number of measurements. The electric circuit model also clarifies the need for homoscedasticity (homogeneity of variance) for

the derived expressions. Before concluding, we discuss meaning, interpretation and comparisons of quantities with uncertainty. Rudimentary concepts of electric power and summation notation are the only background knowledge needed to follow our arguments.

2. Standard deviation is the effective value of noise

Measurement uncertainty is a combined characteristic of the quantity intended to be measured, the measurement system and, in many cases, the human operator. In this section we introduce the battery-resistor circuit that will be used for our analogies throughout the paper. Noise, the source of uncertainty, is introduced ad hoc in a thought experiment. We show how uncertainty is related to the noise and how the definitions of variance and standard deviation follow naturally from a measurement problem.

In the circuit of figure 1(a), V is an ideal battery and v is an added noise that represents uncertainty in the mean value of the e.m.f. $\mathcal{E}$ observed on the resistor R. Each time an individual reading i of $\mathcal{E}$ is made, v assumes a different random voltage, as represented in figure 1(b). The mean value of the e.m.f. on R is $\bar{\mathcal{E}} = V$, where the bar over the variable indicates mean value. The random noise v fluctuates around zero and its mean value is assumed to be zero.

If resistor R is an ideal resistive loudspeaker, the sound produced by it is proportion to the measurement noise described above. The constant

component V of the battery does not contribute to the loudspeaker noise, since noise is caused only by fluctuations. We will further explore this analogy later.

Let us calculate the mean power dissipated on the resistor of figure 1. The amplitude of the noise v_i changes from measurement to measurement (or with time) as represented in figure 1(b). From the point of view of resistor R, the noise v is the source of uncertainty in the determination of the true mean $\bar{\mathcal{E}}$. Let us assume $R = 1\ \Omega$, for simplicity. For a given measurement or instant i, the ith total electric power dissipated on R is given by:

$$P_i = \frac{\mathcal{E}[i]^2}{R} = \mathcal{E}[i]^2 = (V + v_i)^2, \qquad (1)$$

where $\mathcal{E}[i] = V + v_i$ is the total voltage on R. Using equation (1), the mean electric power over N readings is given by:

$$\langle P \rangle = \frac{1}{N}\sum_{i=1}^{N}(V + v_i)^2. \qquad (2)$$

Expanding the term $(V + v_i)^2$ in the equation above we get:

$$\langle P \rangle = V^2 + \frac{1}{N}\sum_{i=1}^{N}v_i^2 + 2V\sum_{i=1}^{N}\frac{v_i}{N}. \qquad (3)$$

For large N the last term in equation (3) tends to zero, since the mean value of the noise is zero. We can now distinguish two contributions to the mean power $\langle P \rangle$:

$$\langle P \rangle = \langle P_{\text{signal}} \rangle + \langle P_{\text{noise}} \rangle. \qquad (4)$$

The first term is the power $\langle P_{\text{signal}} \rangle$ due to V, the ideal battery–the signal. The second, is $\langle P_{\text{noise}} \rangle$, the contribution of noise v. Since $v_i = \mathcal{E}[i] - \bar{\mathcal{E}}$, we can rewrite the term $\langle P_{\text{noise}} \rangle = (1/N)\sum_{i=1}^{N}v_i^2$ as:

$$\langle P_{\text{noise}} \rangle = \frac{1}{N}\sum_{i=1}^{N}(\mathcal{E}[i] - \bar{\mathcal{E}})^2 \equiv \sigma^2. \qquad (5)$$

For $N \gg 1$, equation (5) is identical to the definition of the statistical variance (σ^2) of the e.m.f. $\mathcal{E}$ applied on R. Therefore, the variance of the voltage on R can be interpreted as the mean power of the noise. By analogy, the variance σ^2

of a set of experimental measurements, not necessarily electrical, is the mean power of the fluctuations around the mean value of the readings. The standard deviation σ, that has units of the quantity being measured, is the root mean square of the noise power, or effective value of the noise amplitude.

In the analogy above it should have become clear that the standard deviation σ is an intrinsic characteristic of the noise. With a larger number of measurements it is possible to better characterize σ. The variance is the analog of the acoustic noise power produced by an ideal resistive speaker in place of R.

3. Sum of uncorrelated uncertainties

In this section we will use concepts developed in section 2 to discuss sums of uncertainties. Figure 2(a) shows a circuit analogous to the one shown in the previous section but now we have independent $\mathcal{E}_1$ and $\mathcal{E}_2$ with means V_1 and V_2 and standard deviations σ_1 and σ_2, respectively, to represent a sum of two quantities.

Figure 2(b) shows independent readings of $\mathcal{E}_1$ and $\mathcal{E}_2$ by themselves, displaying the average values V_1 and V_2 and noises v_1 and v_2 for independent events or readings i, j. Noise sources v_1 and v_2 could, in principle, be either correlated or uncorrelated. A metaphor for correlated noise versus uncorrelated noise is a choir versus a group of students chatting in classroom. The choir has a common conductor and the sounds of its members are correlated. In the classroom, sounds generated by the students are typically uncorrelated. Below we will show that noises that do not have common causes greatly simplify the determination of sums of uncertainties.

Following the steps of section 2, let us calculate the total power dissipated on R by $\mathcal{E}_1$ and $\mathcal{E}_2$. For any two readings i and j of the 'noisy' batteries we have that the total power P_{ij} dissipated on $R = 1\ \Omega$ is:

$$P_{ij} = \mathcal{E}[i,j]^2 = (V_1 + V_2 + v_{1i} + v_{2j})^2, \quad (6)$$

where $\mathcal{E}[i,j]$ is the total e.m.f. on R. Indexes i and j are used because v_1 and v_2 are potentially independent entities. The mean total power $\langle P \rangle = \sum_{i,j} P_{ij}/N^2$ dissipated on the resistor is:
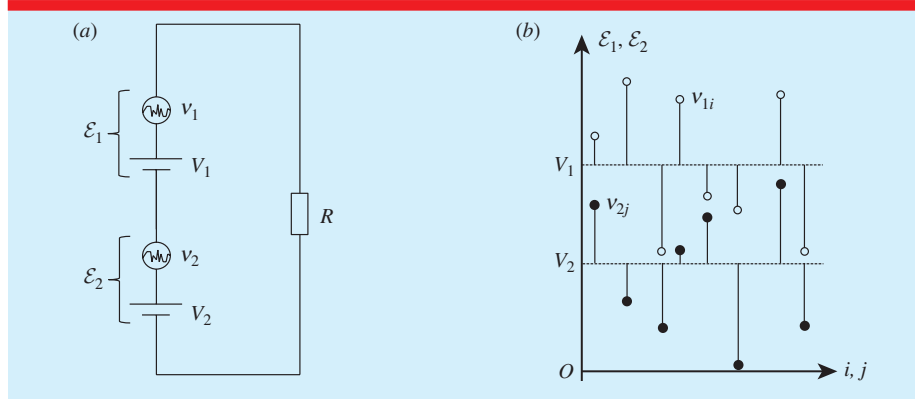
Figure 2. (a) Two voltage sources with noise. (b) Illustration of a series of values of $\mathcal{E}_1$ and $\mathcal{E}_2$ at measurement events i, j, where v_{1i} and v_{2j} (not directly measurable) represent fluctuations around the means V_1 and V_2.

$$\langle P \rangle = (V_1 + V_2)^2 + \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} (v_{1i} + v_{2j})^2$$

$$+ \frac{2(V_1 + V_2)}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} (v_{1i} + v_{2j}). \quad (7)$$

The last term above tends to zero for large N, since $\bar{v}_1 = 0$ and $\bar{v}_2 = 0$, similarly to what happened in equation (3).

Rewriting equation (7) to put in evidence the variances as defined in equation (5), we have:

$$\langle P \rangle = (V_1 + V_2)^2 + \frac{1}{N} \sum_{i=1}^{N} v_{1i}^2$$

$$+ \frac{1}{N} \sum_{j=1}^{N} v_{2j}^2 + 2 \left(\frac{1}{N} \sum_{i=1}^{N} v_{1i} \right) \left(\frac{1}{N} \sum_{j=1}^{N} v_{2j} \right).$$
$$(8)$$

With uncorrelated uncertainties, the last term with factors in parenthesis tends to zero for large N, simplifying the expression. Using definitions of equation (5), we can rewrite equation (8) as:

$$\langle P \rangle = (V_1 + V_2)^2 + \langle P_{\text{noise}_1} \rangle + \langle P_{\text{noise}_2} \rangle. \quad (9)$$

As in section 2, the mean power dissipated on R is again a sum of two terms: $\langle P_{\text{signal}} \rangle + \langle P_{\text{noise}} \rangle$, where now $V_1 + V_2$ is the signal. Remembering equation (5), the term $\langle P_{\text{noise}} \rangle$ is the variance σ_T^2 of the total e.m.f applied to R; $\langle P_{\text{noise}_1} \rangle$ and $\langle P_{\text{noise}_2} \rangle$ are the variances σ_1^2 and σ_2^2 of the quantities $\mathcal{E}_1$ and $\mathcal{E}_2$ being added. Therefore the sum of

the uncertainties characterized by σ_1 and σ_2 is as follows:

$$\sigma_T^2 = \sigma_1^2 + \sigma_2^2. \quad (10)$$

Observe that while the mean values of measurements (signals) add up linearly, $\bar{\mathcal{E}} = V_1 + V_2$, the noises or standard deviations add up quadratically: $\sigma_T^2 = \sigma_1^2 + \sigma_2^2$. Equation (10) is the formula for the sum of uncorrelated uncertainties. We would like to draw attention to the fact that in sums, but also in differences $(\mathcal{E}_1 - \mathcal{E}_2)$ between two uncorrelated values the uncertainties are given by equation (10). To visualize this, just invert one of the batteries in figure 2, changing V_1 for minus V_1 while keeping V_2 constant. The total power of the noise is still given by $\sigma_1^2 + \sigma_2^2$, since the powers of the two independent noises on R add irrespective of the polarity of the batteries. These properties will be further explored in the next sections.

Before using equation (10) it is important to understand its limitations. When the noises v_1 and v_2 have a common cause, for example, a vibrating air conditioner in the room, the uncertainties will have some degree of correlation. For correlated uncertainties it is easy to show that equation (10) is not valid. Let us examine a limiting case where $\sigma_1 = \sigma_2$ but also $v_{1i} = v_{2j}$ (perfect correlation). In this case we can simply call $v = 2v_1$ in the definition of noise before equation (5) and the total uncertainty will be given by $\sigma_T^2 = 4\sigma_1^2$. If the uncertainties were uncorrelated the answer would be $\sigma_T^2 = 2\sigma_1^2$, according to

equation (10). For partial correlation $\sigma_T^2 = a\sigma_1^2$, with $2 \leqslant a \leqslant 4$. Correlated uncertainties propagation will not be discussed here and have been treated by Taylor [14].

4. Standard deviation of the mean: the uncertainty of the mean value

In this section we calculate the mean value by averaging measurements of known standard deviations and determine the uncertainty of this new mean. The uncertainty in the determination of the mean is called *standard deviation of the mean*, and often incorrectly called *standard error* (SE) [9]. In this and in the next sections we will discuss the advantages of use of standard deviation of the mean, as opposed to standard deviation, to characterize uncertainty in physical sciences.

Figure 3 is a generalization of figure 2(a) for an arbitrary number n of 'noisy' batteries. Following the argument of section 3 we obtain the following relations for the sum of the means and for the sum of the standard deviations:

$$\bar{\mathcal{E}} = V_1 + V_2 + V_3 + ... + V_n, \qquad (11)$$

$$\sigma_T^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + ... + \sigma_n^2, \qquad (12)$$

where V_k $(k = 1, 2, ..., n)$ are the mean values of voltages of each individual battery, $\bar{\mathcal{E}}$ is the expected mean total voltage on the resistor R, and σ_k are the respective standard deviations.

Consider now the batteries of figure 3 are all of the same type. The mean voltage of one the batteries of the group will be simply $\bar{V} = \bar{\mathcal{E}}/n$. Since the batteries are of the same type, the individual measurements V_k have identical standard deviations σ (same type of noise). We can use equation (12) and get the uncertainty of the mean value $\bar{V}$:

$$\sigma_{\bar{V}} = \frac{\sigma_T}{n} = \frac{\sqrt{n}\,\sigma}{n} = \frac{\sigma}{\sqrt{n}}. \qquad (13)$$

Equation (13) indicates that the uncertainty in the mean value $\bar{V}$, commonly represented by $\sigma_{\bar{V}}$, increases with the standard deviation and decreases with the square root of the number of readings [15]. The *standard deviation* is a characteristic of the noise of the measurement system, method and/or the intrinsic noise of system being measured. If we want to decrease the *standard deviation of the mean* by increasing the number of measurements,

we need, for example, to increase the number of readings by a factor of a hundred to get an improvement in the uncertainty of the mean by a factor of ten. If the measurement noise is caused by the measurement system or method, it might be worth investing in improving the experimental conditions to decrease the standard deviation. As seen in equation (13) a decrease in the standard deviation decreases the uncertainty of the mean much faster than an increase in the number of measurements.

The number of measurements required to bring uncertainty down to a desired level can be determined after the standard deviation σ is estimated early in the measurement process, in a pilot experiment. Derivation of equation (13) assumes that the variances are equal and should not change from measurement to measurement. Therefore to use equation (13) the multiple measurements of the same quantity should come from the same probability distribution (homoscedasticity). In the loudspeaker analogy homocedasticity would be equivalent to noises with equal and constant average powers and timbres.

5. Uncertainty of the mean and statistical distribution of noise

To measure is to compare. We often need to compare different measurements. When measurement uncertainties are involved, we need to be able to quantify agreement or disagreement between values. In this section we discuss the sensitivity of the standard deviation to the statistical distribution of the measurement noise, and the stability of the uncertainty of the mean to the statistical distribution of noise. The statistical distribution of the readings is the timbre of the noise in the loudspeaker analogy. The results of this section are from numerical simulations that we use to illustrate our arguments.

In figure 4(a) we show the histogram of measurement values whose noise has the so-called normal or Gaussian statistical distribution. $\bar{\mathcal{E}}$ is the mean value of the set of readings and the dashed lines represent the mean value plus or minus two standard deviations. The region between the dashed lines represent approximately 95% of the events in a Gaussian distribution. Therefore, on average, in measurements obeying the normal distribution, only around 5% of observations will fall outside the range $\bar{\mathcal{E}} \pm 2\sigma$.

Figure 4(b) shows the distribution of the mean values $\bar{\mathcal{E}}$, where each mean value is determined from a distribution such as the one of figure 4(a). The width of figure 4(b) is proportional to the standard deviation of the mean. On average, in 95% of the experiments with $\mathcal{E}$ noise distribution such as in figure 4(a), the mean value obtained will be in the range $\bar{\mathcal{E}} \pm 2\sigma_{\bar{\mathcal{E}}}$. The mean value of the means distribution $\langle \bar{\mathcal{E}} \rangle$ is the same as the expected value $\bar{\mathcal{E}}$, for very large N; actual values might be slightly different due to statistical fluctuations for a finite number of readings.

Figure 5(a) shows a histogram of readings where the statistics is not Gaussian. The distribution displayed is a uniform distribution of readings of unbiased dice, where the probabilities of integer outcomes 1 through 6 are identical. This truncated type of distribution is common in cases where there are natural physical limitations or when the quantities under consideration are presorted such as in quality control in a factory. The positions of the calculated $\bar{\mathcal{E}} \pm 2\sigma$ for the distribution shown are represented in figure 5(a) by the vertical dashed lines. Observe that, different from the Gaussian distribution, here we are 100% certain that any measurement will lead to a value $\mathcal{E}$ well inside the $\bar{\mathcal{E}} \pm 2\sigma$ interval. If it was a Gaussian distribution, approximately 5% of the observations would fall outside the $\bar{\mathcal{E}} \pm 2\sigma$ interval. This shows that the interpretation of *standard deviations* for measurements depends on the statistical distribution of the noise. For each experimental case we need to have an estimate of the mathematical form of the noise distribution before interpreting the meaning of the standard deviations in statements such as $\bar{x} \pm \sigma$.

Figure 5(b) shows a distribution of mean values when each $\bar{\mathcal{E}}$ comes from an experiment with noise distribution as shown in figure 5(a). Observe that the distribution in figure 5(b) looks Gaussian, despite the fact that figure 5(a) is a uniform distribution. The central limit theorem [15] explains what is happening: if a random variable is measured repeatedly and independently and results are averaged, the distribution of the averages tends to a normal distribution for a large number of measurements. Therefore, independently of the statistical distribution of the noise, the distribution of the means tends to a Gaussian distribution. The interpretation of uncertainty written as
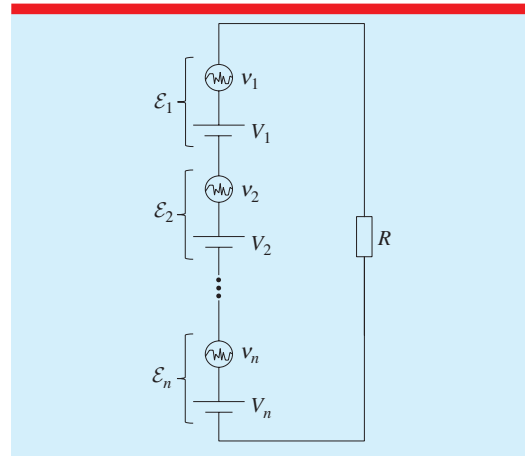

Figure 3. Sum of values $\mathcal{E}_i$, with mean values V_i and simulated sources of uncertainties v_i (with standard deviations σ_i).

standard deviations of the mean follows normal distribution rules for measurements with a large enough number of readings. In practice, with N starting at around 16 the Gaussian formalism may be used for the uncertainty of the mean, since in this case there is only one significant figure used for the uncertainty. The limitation in the number of significant figures in the uncertainty of the mean arises because the estimation of the standard deviation also suffers from uncertainty. There is an uncertainty of approximately $71/\sqrt{(N-1)}$ % in the determination of $\sigma_{\bar{\mathcal{E}}}$. For example, if $N = 50$, there is approximately 10% uncertainty in the determination of $\sigma_{\bar{\mathcal{E}}}$ itself; more than one or two figures to represent $\sigma_{\bar{\mathcal{E}}}$ would not be significant or meaningful. For a discussion on how to display the number of significant figures in the uncertainty of the mean, based on the number of measurements, see [2].

The advantage of using the uncertainty of the mean instead of the standard deviation of a distribution is to make the uncertainty analysis independent of the probability distribution. The dashed lines in figure 5(b) represent the limits $\bar{\mathcal{E}} \pm 2\sigma_{\bar{\mathcal{E}}}$. Each time we perform the experiment with a finite number of readings the mean will change slightly. Independently of the statistical distribution of the noise, approximately 95% of the experiments performed with the same experimental method and same number of measurements and will give *means* in the range $\bar{\mathcal{E}} \pm 2\sigma_{\bar{\mathcal{E}}}$.
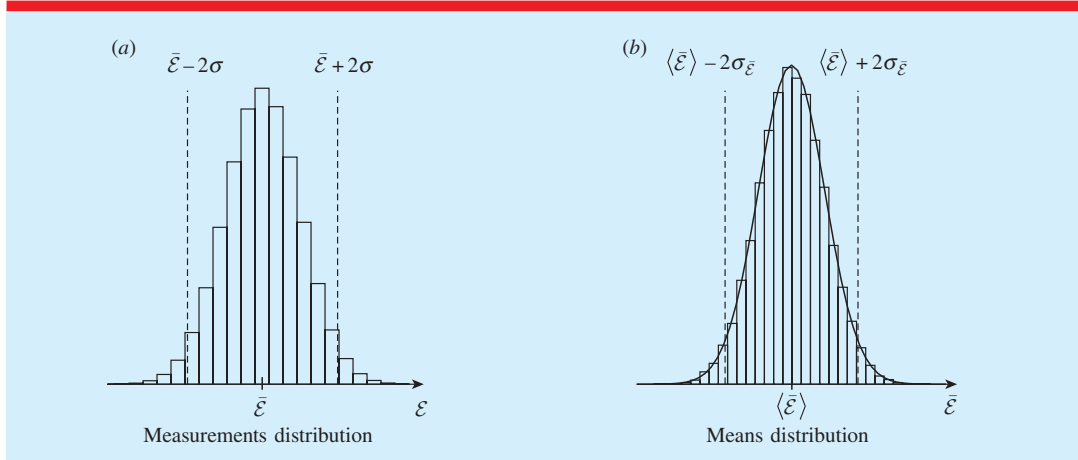
Figure 4. Normal noise. (a) $N = 1000$ observations. As in any Gaussian distribuition, approximately 95% of the $\mathcal{E}$ values measured are in the interval $\bar{\mathcal{E}} \pm 2\sigma$, between the dashed vertical lines. (b) $n = 2000$ repetitions of measurements with $N = 1000$ observations each. σ is the standard deviation of $\mathcal{E}$, $\langle \bar{\mathcal{E}} \rangle \approx \bar{\mathcal{E}}$ and $\sigma_{\bar{\mathcal{E}}} \approx \sigma/\sqrt{2000}$. Once again, approximately 95% of the $\bar{\mathcal{E}}$ means are found between the two vertical lines. The adjusted line on (b) is a Gaussian fit.
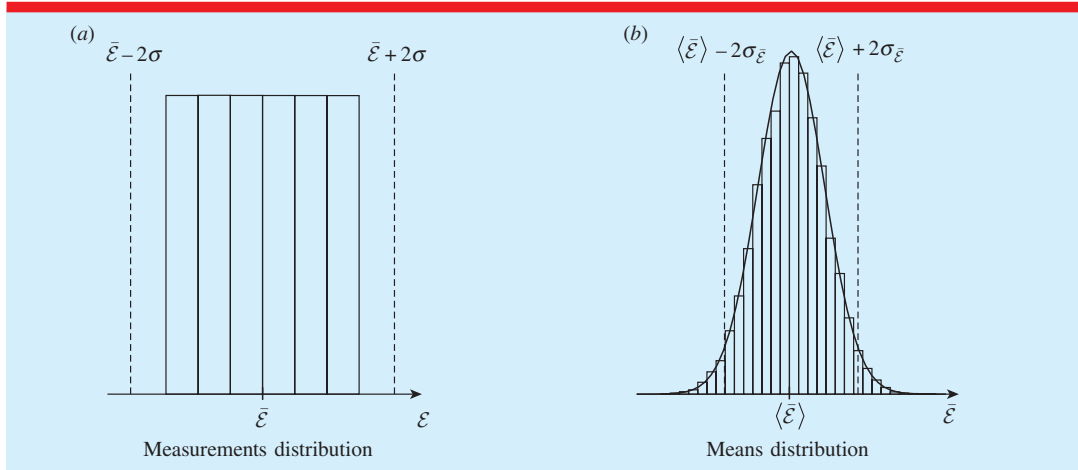

Figure 5. Uniform distribution of six-sided dice. (a) $N = 1000$ observations; the six bins are the probabilities of outcomes 1 through 6, with mean $\bar{\mathcal{E}}$ very close to true exact value 3.5 and $\sigma = 1.708$. The dashed lines at $\bar{\mathcal{E}} \pm 2\sigma$ are expected at positions 0.084 and 6.916, outside of the 1–6 range of the uniform distribution. (b) $n = 2000$ repetitions of the experiment to determine $\bar{\mathcal{E}}$ with $N = 1000$ observations in each repetition. σ is the standard deviation of $\mathcal{E}$, $\langle \bar{\mathcal{E}} \rangle \approx \bar{\mathcal{E}}$ and $\sigma_{\bar{\mathcal{E}}} \approx \sigma/\sqrt{2000}$. The adjusted line on (b) is a Gaussian fit.

Figure 6 shows a different perspective on the data presented in the distributions of figure 5. Figure 6(a) shows a series of experimentally observed values of six-sided dice being thrown. The values $\mathcal{E}$ observed are integers between 1 and 6. The theoretical expected value for the mean value $\bar{\mathcal{E}}$ of the dice results is exactly 3.5. Panel 6(b) shows the results for $\bar{\mathcal{E}}$, σ and $\sigma_{\bar{\mathcal{E}}}$ obtained from the 'experimental' values shown

in figure 6(a), from throw 1 up to the Nth throw. We observe in 6(b) how the mean value fluctuation decreases with the number of measurements (throws) and converges to 3.5 as the number of measurements increases. The standard deviation σ behaves in a similar way: the uncertainty in σ decreases with the number of measurements and σ converges to its expected value of approximately 1.7. Finally the standard deviation of the mean $\sigma_{\bar{\mathcal{E}}}$

also decreases with the number of measurements and tends to zero for a very large number of measurements. Observing the plots of $\mathcal{E}$ and $\sigma_{\bar{\mathcal{E}}}$ we can see a decrease in the uncertainty of the standard deviation of $\bar{\mathcal{E}}$, that is given by $\sigma_{\bar{\mathcal{E}}}$. The dashed line in the $\sigma_{\bar{\mathcal{E}}}$ plot is the theoretically expected decrease in the uncertainty of the mean with the number of experiments, $\sigma_{\bar{\mathcal{E}}} = \sigma/\sqrt{n}$. In a different run of the same dice experiment the curves would be slightly different, especially for the low N region, but the trends would be the same.

In the dice experiment the standard deviation is an intrinsic characteristic of the observed data, not of the experimental method; only an increase in the number of measurements can decrease the uncertainty in an experimental determination of the mean value $\bar{\mathcal{E}}$. On the other hand, in experimental cases where the measurement noise is also due to the method used, a decrease in σ by improvement of the experimental conditions would lead to a faster convergence of the mean value since $\sigma_{\bar{\mathcal{E}}}$ depends linearly on σ but decreases only with the inverse of the square root of the number of measurements.

6. Comparing quantities that have uncertainties

We have seen in the previous sections that we can interpret *uncertainty of the mean* assuming a Gaussian distribution whenever the number of measurements averaged is large. This is an advantage of use of uncertainty of the mean compared to standard deviation. Therefore, let us focus here on uncertainties described in terms of uncertainties of the means.

Let us say we have two values $X = \bar{X} \pm 2\sigma_{\bar{X}}$ and $Y = \bar{Y} \pm 2\sigma_{\bar{Y}}$. Does X agree or disagree with Y? Due to the probabilistic nature of the measurements, in general we cannot claim with 100% certainty that two values with uncertainty agree or disagree with each other. However, we can estimate the percentage of the cases where similar experiments, if repeated, will lead to a specified degree of agreement or disagreement with each other.

If X agrees with Y, then the difference $X - Y$ must have a very low probability of not being zero. Of course, X and Y agree 100% if $X - Y$ *is* zero, with no uncertainty—but that is never the case when uncertainties are present. In equation (10) and discussion thereafter, we have seen that the difference between two uncertainties is also given by said equation. Therefore, X and Y agree in the confidence interval of 95% if the difference $(\bar{X} - \bar{Y}) \pm 2\sqrt{\sigma_{\bar{X}}^2 + \sigma_{\bar{Y}}^2}$ contains the value zero. In this case only around 5% of similar experiments with the same number of readings will find $(X - Y)$ outside the interval $(\bar{X} - \bar{Y}) \pm 2\sqrt{\sigma_{\bar{X}}^2 + \sigma_{\bar{Y}}^2}$.

Let us take as an example two mass quantities: $X = (100 \pm 4)$ kg, 95% confidence interval (uncertainty approximately $2\sigma_{\bar{X}}$) and $Y = (106 \pm 3)$ kg, 95% C.I. The difference between X and Y is (6 ± 5) kg, 95% C.I. Since this difference does not include zero, X and Y are statistically different at the 95% confidence interval. However with a higher confidence interval we might not be able to say that X and Y are statistically different. In general, we can not say with 100% confidence that two mean values with uncertainties are distinct. Now even if we improve the measurements and get lower uncertainties, we might have, for example $W = (100.0 \pm 0.4)$ kg, 95% C.I. and $Z = (106.0 \pm 0.3)$ kg, 95% C.I., and the difference is $(W - Z) = (6.0 \pm 0.5)$ kg, 95% C.I. Even if we multiply the uncertainty by 3, making a 6σ level uncertainty that increases the confidence interval to 99.999 9998%, we can still have that W and Z can be considered distinct, since the confidence interval does not include zero: (6.0 ± 1.5) kg, 99.999 9998% C.I. At this level of confidence, repetitions of the experiment will have $(W - Z)$ outside the range (6.0 ± 1.5) less than once in every 500 million repetitions. In this case we say that W and Z are distinct for all practical purposes.

Different fields of Science have different requirements on confidence intervals. While applied sciences are less stringent, using 95.45% and 99.73% for confidence intervals (2σ and 3σ), fundamental physics uses at least 5σ standard to accept a physical discovery (99.999 942% C.I.). More systematic and finer comparisons between experimental quantities can be done with help of statistical hypothesis testing.

Before closing, we would like to remind the reader that the statistical uncertainties discussed in this paper do not account for discrepancies (errors) in the measured values due to the methodology used—type B uncertainty due to
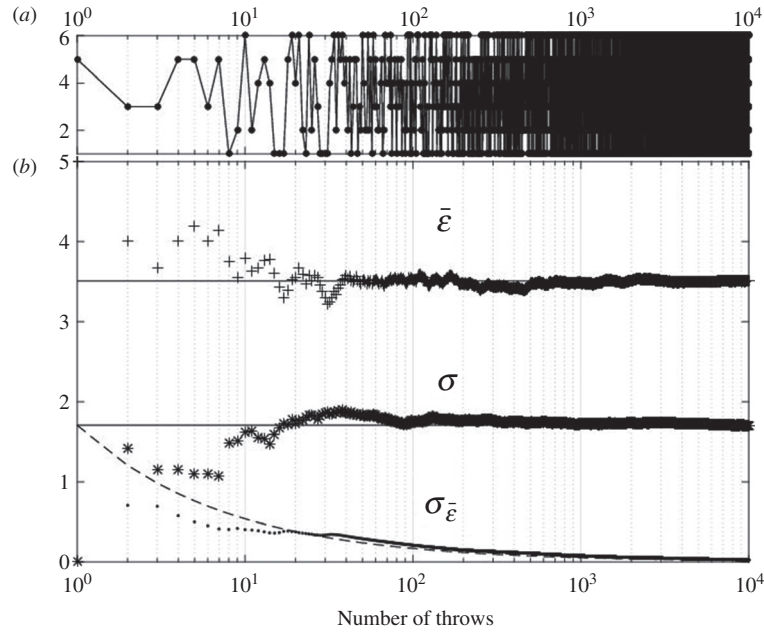
Figure 6. Dice throwing experiment. (a) Values observed in successive throws of dice in an experiment, (b) the crosses are the averages $\bar{\mathcal{E}}$ of the values (a) up to the nth observation, the stars are the standard deviations σ calculated up to the nth observation, and the dots are the standard deviations of the means $\sigma_{\bar{\mathcal{E}}}$, up to the nth mean $\bar{\mathcal{E}}$. The dashed line is the expected dependency $\sigma/\sqrt{n}$ for $\sigma_{\bar{\mathcal{E}}}$. Each of the n throws comes from a uniform distribution with standard deviation σ.

measurement equipment reproducibility issues. To establish whether the measurement method used was adequate, alternative measurement methods need to be used for cross checking. On a different note, our analogies are directly valid for sums and differences of quantities with uncertainties. We showed how to determine the uncertainty $\sigma_{\bar{f}}$ in an expression of the type $f(x, y) = x \pm y$, where both x and y have uncertainties. The determination of the influence of $\sigma_{\bar{x}}$ and $\sigma_{\bar{y}}$ on $\sigma_{\bar{f}}$ for an arbitrary function $f(x, y)$ can be visualized from an extrapolation of our analogies. For uncorrelated uncertainties, we just need to determine how the noise power in $f(x, y)$ depends on $\sigma_{\bar{x}}$ and on $\sigma_{\bar{y}}$, independently, and add these contributions, since the total power of noise at the resistor of the analogy is the independent sum of the partial powers due to each noise source. For small uncertainties, this calculation is usually done using partial derivatives to determine the independent contributions to the noise in f due to each uncertainty $\sigma_{\bar{x}}$ and $\sigma_{\bar{y}}$, and adding these contributions in quadrature [2]. Alternatively, the contributions can be determined via finite differences as follows:

$$\sigma_{\bar{f}}^2 = |f(\bar{x}, \bar{y}) - f(\bar{x} \pm \sigma_{\bar{x}}, \bar{y})|^2 + |f(\bar{x}, \bar{y}) - f(\bar{x}, \bar{y} \pm \sigma_{\bar{y}})|^2,$$

where, for very small uncertainties, either the positive or negative signs can be used.

7. Conclusion

We have proposed a circuit analogy to help introduce basic concepts in measurements and uncertainties. The analogy drew attention to the need for uncorrelated sources of uncertainty and for homoscedesticity in use of typical expressions used in the introductory laboratory. The loudspeaker analogy clarifies the fact that the standard deviation is a property of the measurement noise, not of the number of points or readings used in the measurement. In the battery-resistive loudspeaker metaphor, the acoustic power of the noise is the variance of quantity being measured; the statistical distribution of the noise is the analogous of the timbre of the noise in the speaker. Finally we discussed how the interpretation of uncertainty of the mean is much less susceptible to the statistical distribution of the experimental noise compared to standard deviation.

ORCID iDs

Gabriel L A de Sousa ⓘ https://orcid.org/0000-0001-5657-1441
George C Cardoso ⓘ https://orcid.org/0000-0001-8459-9812

References

[1] Siegel P 2007 Having fun with error analysis *Phys. Teach.* **45** 232–4
[2] Hughes I and Hase T 2010 *Measurements and Their Uncertainties: a Practical Guide to Modern Error Analysis* (Oxford: Oxford University Press)
[3] Petley B W 1985 *The fundamental Physical Constants and the Frontier of Measurement* (Bristol: Hilger)
[4] Buffler A, Allie S and Lubben F 2008 Teaching measurement and uncertainty the GUM way *Phys. Teach.* **46** 539–43
[5] Lubben F, Buffler A, Allie S and Campbell B 2001 Point and set reasoning in practical science measurement by entering university freshmen *Sci. Educ.* **85** 311–27
[6] Buffler A, Allie S and Lubben F 2001 The development of firrst year physics students' ideas about measurement in terms of point and set paradigms *Int. J. Sci. Educ.* **23** 1137–56
[7] Allie S, Buffler A, Campbell B, Lubben F, Evangelinos D, Psillos D and Valassiades O 2003 Teaching measurement in the introductory physics laboratory *Phys. Teach.* **41** 394–401
[8] Kung R L 2005 Teaching the concepts of measurement: an example of a concept-based laboratory course *Am. J. Phys.* **73** 771–7
[9] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2008 *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement* (Geneva: International Organization for Standardization) (Joint Committee for Guides in Metrology, JCGM 100:2008) p 167
[10] Kacker R, Sommer K D and Kessel R 2007 Evolution of modern approaches to express uncertainty in measurement *Metrologia* **44** 513
[11] Duit R 1991 On the role of analogies and metaphors in learning science *Sci. Educ.* **75** 649–72
[12] Dagher Z R 1995 Analysis of analogies used by science teachers *J. Res. Sci. Teach.* **32** 259–70
[13] Herrmann F and Schmid G B 1985 Analogy between mechanics and electricity *Eur. J. Phys.* **6** 16
[14] Taylor J R 1985 Simple examples of correlations in error propagation *Am. J. Phys.* **53** 663–7
[15] Reichl L 1986 *A Modern Course in Statistical Physics* vol 1998 (London: Edward Arnold)

Gabriel L A de Sousa studied medical physics at the University of São Paulo, earning his B.S. degree on January of 2018. He is currently dedicated to the pursuit of academic training for a career in fundamental theoretical physics.

George C Cardoso is a physicist and professor doctor at the University of São Paulo. Previously he spent several years working as a corporate research scientist in the United States. He completed his D.Sc. at Federal University of Pernambuco, Recife, Brazil. His current research interests include development of optical methods and techniques with applications to medicine and biology.