The American Statistician

Taylor & Francis
Taylor & Francis Group

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://www.tandfonline.com/loi/utas20

Explaining the Importance of Variability to Engineering Students

Pere Grima, Lourdes Rodero & Xavier Tort-Martorell

To cite this article: Pere Grima, Lourdes Rodero & Xavier Tort-Martorell (2015): Explaining the Importance of Variability to Engineering Students, The American Statistician, DOI: 10.1080/00031305.2015.1064478

To link to this article: http://dx.doi.org/10.1080/00031305.2015.1064478

Explaining the Importance of Variability to Engineering Students

Pere Grima, Lourdes Rodero and Xavier Tort-Martorell

Department of Statistics and Operational Research, Universitat Politècnica de Catalunya - BarcelonaTech, Barcelona, 08028, Spain

Author's footnote:

Pere Grima is a Professor at Universitat Politècnica de Catalunya – BarcelonaTech, Barcelona, Spain (e-mail: pere.grima@upc.edu); Lourdes Rodero is an Associate Professor at Universitat Politècnica de Catalunya – BarcelonaTech, Barcelona, Spain (e-mail: lourdes.rodero@upc.edu and Xavier Tort-Martorell is a Professor at Universitat Politècnica de Catalunya – BarcelonaTech, Barcelona, Spain (e-mail: xavier.tort@upc.edu)

ABSTRACT

One of the main challenges of teaching statistics to engineering students is to convey the importance of being conscious of the presence of variability and of taking it into account when making technical and managerial decisions. Often, technical subjects are explained in an ideal and deterministic environment. This article shows the possibilities of simple electrical circuits – the Wheatstone Bridge among them— to explain to students how to characterize variability, how it is transmitted and how it affects decisions. Additionally they can be used to introduce the importance of robustness by showing that taking into account the variability of components allows the design of cheaper products with greater benefits than if one were to simply apply formulas that consider variables as exact values. The results are quite unexpected, and they arouse the interest and motivation of students. Supplementary materials for this article are available online.

KEY WORDS: Teaching, variability, engineering students, robust product, Wheatstone bridge

 ACCEPTED MANUSCRIPT

1. INTRODUCTION

When we explain the importance of variability, the attitude of our listeners is very different depending on whether we are addressing a group of professionals or a group of students with no professional experience.

On the one hand, it is clear to professionals that variability is a problem (many times it is "the" problem), and they are very conscious of the need to characterize it, measure it and minimize its effects. However, we have the feeling that to engineering students –who are used to, almost exclusively, deterministic approaches– our lectures on the subject appear to be merely rhetorical. Our experience is that the usual examples, such as packaging processes (there are no two products that weigh "exactly" the same) or the dimensions of mechanical components, are not useful to capture the students' interest. After all, in the end, who cares if a package of rice weighs a few grams more or less than another? Furthermore, on top of capturing their attention we want them to understand and assimilate the concept of variability –not an easy task, as explained by Garfield and Ben-Zvi (2005).

Among the many proposals that have been published for arousing the interest of students, we find particularly interesting that of Søren Bisgaard in a well-known article (Bisgaard, 1991), where he shares his experiences of explaining statistics to engineering students as well as professionals. A core element of his student course involved designing, building and flying a paper helicopter, which has now become famous (Box, 1992), and in which the problems derived from variability take a central role. Another article that demonstrates an original focus, although it is not oriented specifically to engineering students, is that of Lee (2007), where he explains how he uses chocolate chip cookies in his classes to grab the students' attention in the analysis of their characteristics (such as the number of chips in each cookie). This allows him to demonstrate different techniques and key ideas, among which variability is included.

In this article we present how, in an introductory statistics course —the first time engineering students are exposed to statistics— we use two electrical circuits to make students conscious of the existence of variability and the technical problems it introduces. The course is taught for 15 weeks at 4 hours per week during the spring semester of their second year. We present these examples one after the other as an introduction to random variables and probability distributions. We focus on the concepts; the justification of the results comes later once the needed theoretical material has been covered.

2. APPETIZER: A VERY SIMPLE CIRCUIT

We begin by assuming that we have received an order to manufacture a batch of very simple electrical circuits: just a battery connected to a resistor. For simplification, we consider that all the batteries have an identical voltage of $V = 100\,V$, and that —more realistically— not all resistors deliver the same resistance (R). As the resistance is different for each circuit, the intensity (I) that flows through each one is also different ($I = V/R$, Ohm's law). The question we then pose to the students is: If our client requests lots of circuits with an average intensity of $10A$, what should the nominal resistance value be?

Most of them do not even bother to give the "obvious" answer: 10Ω. Students begin to show some interest when we say that if the resistors have variability, their nominal value should not be 10Ω. We explain that resistors producing values below 10Ω will harm more than resistors with values above it; and we illustrate the idea with resistances of 7Ω and 13Ω. Obviously, the intensities flowing would be $100V/7\,\Omega = 14.29A$ (a difference of 4.29A with respect to the objective) and $100V/13\,\Omega = 7.69A$ (a difference of only 2.31A). Assuming that the resistance values were distributed symmetrically around this central value, it is clear that a nominal value bigger than 10 Ω is needed to obtain an average intensity of $10A$.

If we suppose that $\mathrm{Var}(R) = 1\Omega^2$, the desired nominal value is 10.1 Ω. Later on –once we have introduced the mathematical expectation and variance of random variables-- we propose the exercise of finding the sought value (the expectation of I) expanding V/R in Taylor series (see section A.1 on the Appendix).

We then proceed to pose, using the same example, the following more realistic and easy problem: Almost surely, the buyer's real interest is minimizing the proportion of circuits with intensity outside specifications rather than having the average intensity on target. It is then easy to show that, assuming the specifications are 10 ± 3A and that, on top of being symmetric and having $\mathrm{Var(R)} = 1\Omega^2$, resistance values follow a normal distribution, we get $I > 13$ when $R < 7.69$ and $I < 7$ when $R > 14.29$. And, given the symmetry, it is obvious that to minimize the number of non-conforming resistors we want them centered at $(7.69 + 14.29)/2 = 11.0\ \Omega$. Two weeks later, once the normal distribution has been covered, the students find these values themselves as part of an exercise.

3. THE WHEATSTONE BRIDGE

Following the case just discussed, we present a slightly more sophisticated example that is more realistic and that has more didactic possibilities: the Wheatstone Bridge. The Wheatstone Bridge is an assembly for measuring electrical resistance; it was employed by Genichi Taguchi (Taguchi and Wu, 1979) to explain his robust design techniques, and also by George Box to demonstrate the weaknesses of Taguchi's methods while at the same time illustrating his proposals (Box and Fung, 1994). Although it seems to have currently fallen out of style and is no longer present in university physics textbooks, it continues to be an excellent example for demonstrating what we lose when we ignore variability.

Figure 1 presents a circuit diagram. The resistor with unknown resistance is placed in R_x. The values of R_1 and R_3 are known, and R_2 is a variable resistor that, once R_x is connected, is regulated such that the current does not pass through ammeter A.

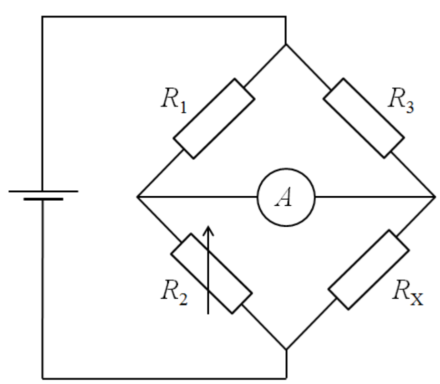

Figure 1: Diagram of a Wheatstone Bridge

In this situation, by applying Kirchhoff's laws, we deduce that $R_2/R_1 = R_x/R_3$ and, thus:

$$R_x = \frac{R_2\, R_3}{R_1} \tag{1}$$

If we ignore variability, any value of R_1 and of R_3 will be equally good. One comfortable solution, as suggested by some textbooks, is that $R_1 = R_3$ and, therefore, $R_x = R_2$; but this is a bad solution, as discussed below.

In what follows we will distinguish between the real value of R_x (which we denote simply as R_x) and the measurement result (which we refer to as $\tilde{R}_x$). Let us suppose that $R_1 \sim N(\mu_1;\ \sigma_1)$, $R_3 \sim N(\mu_3;\ \sigma_3)$, and that –as would happen in practice– to determine the value of R_x we would use μ_1 and μ_3 rather than the unknown real values of R_1 and R_3. Assuming that R_2 is affected by a certain experimental error, $\varepsilon \sim N(0;\ \sigma_2)$, we have:

$$\tilde{R}_x = (R_2 + \varepsilon)\frac{\mu_3}{\mu_1} \tag{2}$$

Resolving for R_2 in (1) and substituting in (2) we have:

$$\tilde{R}_x = \left(R_x \frac{R_1}{R_3} + \varepsilon\right)\frac{\mu_3}{\mu_1} \tag{3}$$

It is then clear that the variability of the measurement result ($\tilde{R}_x$) depends at least on the ratio μ_3/μ_1. This dependence can be shown to the students by means of a spreadsheet (available online as supplementary material). In it we calculate the value of $\tilde{R}_x$ for a large set of circuits and observe how the measurement precision changes when the values μ_1, μ_3, σ_1, σ_2 and σ_3 change. The values of R_1 and R_3 are generated from their respective distributions, and the values of R_2 and $\tilde{R}_x$ from expressions (1) and (3), respectively. Figure 2 illustrates the values of $\tilde{R}_x$ obtained in 1000 circuits with components having the parameters shown in the Figure and a value of $R_x = 20\ \Omega$.

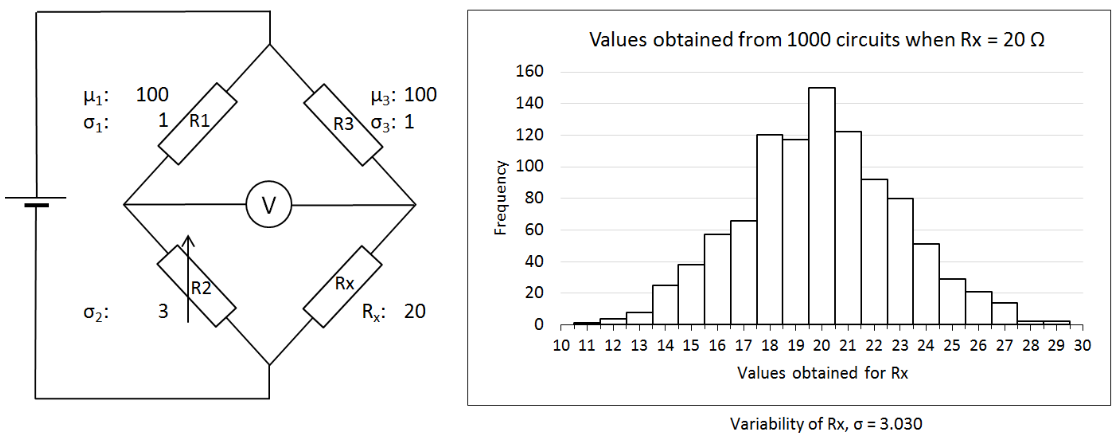

Figure 2: Values obtained for a resistor of 20 Ω in 1000 circuits with the indicated characteristics

We then explain that the resistance value of a common resistor is identified by bands of color printed on its body, with an additional final band that indicates its tolerance (Figure 3). And we show catalogues where it can be seen that the price is related more to the resistor tolerance than to its nominal value.

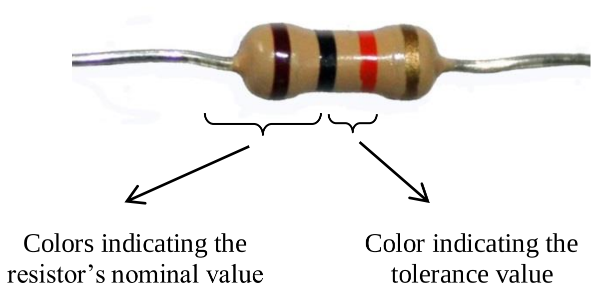

Figure 3: Resistor with bands indicating its nominal value and tolerance

At this point we ask the question: Is it possible to manufacture Wheatstone Bridges with less expensive components and capable of measuring with greater precision? The spreadsheet immediately provides the answer: Yes. For example, in Figure 4 the nominal resistance values R_1 and R_3 have been changed and –despite having multiplied by 10 the values of σ_1 and σ_3, and increased the value of σ_2, from 3 to 10 Ω— the device's precision has improved considerably: the standard deviation of the measurement is 3 times smaller.

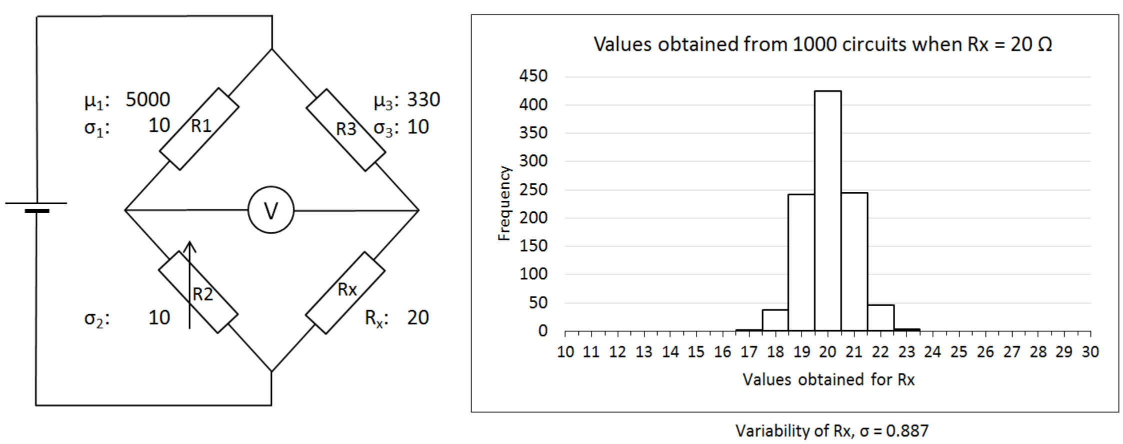

Figure 4: Values obtained for a resistor of 20 Ω for 1000 circuits with the indicated characteristics

At this point we do not justify why this is so. We just show the results and provide the spreadsheet to the students so that they can play with it and experience firsthand the effect of changing the parameters of R_1, R_2 and R_3 distributions. Some weeks later, when students have the needed background, we offer them an exercise in which they must theoretically justify the results.

4. CLASSROOM EXPERIENCE

As mentioned above, we use this material in a first course on statistics taught to engineering students. The course represents 150 hours of student work: 60 hours in the classroom (4 hours per week during 15 weeks) plus 90 hours of personal student work (exercises, mini projects and presentations).

When students take this course they have already taken two courses in calculus and one in electromagnetism, which effectively covers the concepts of circuit theory needed to understand the Wheatstone bridge example.

The first week is introductory; we cover a few motivating examples of the use and importance of statistics in engineering and some basic descriptive statistics.

Variability is introduced the second week. To increase students' interest, we present the simple circuit (one power supply and one resistor) example. It is a great way to show that when variability is involved, things are not what they seem. Then we introduce the more complex Wheatstone bridge, discussing briefly how it works. Students quickly realize that predicting the effect of variability is more difficult in this example, and they welcome the spreadsheet. We show them the results obtained with the values used in Figures 2 and 4. The fact that it is possible to obtain a better product with worse, and thus cheaper, components becomes evident.

The time we spend to present these examples is about 20 minutes. We leave the spreadsheet available to students so they can play with it. At this point we do not justify the results -students still lack the needed concepts and tools- but we let them know that they will find out the justification, through exercises, later on. The remaining time of the second week is devoted to random variables, probability distributions and their properties.

The third week is dedicated to specific probability distributions, among them the normal distribution. One of the suggested exercises is, for the simple circuit, calculating what the nominal value of the resistance should be in order to minimize the proportion of defective circuits if the tolerances are $10 \pm 3A$. It is a rather original exercise that, in spite of being relatively easy, has a surprising result.

In the fourth week, the two main topics are linear combination of random variables and sampling distribution of $\bar{X}$ and S^2. The product or the ratio of random variables is not covered; we tell students that when needed, a good solution is to expand the function in Taylor series. Exercises proposed in this week are, among others:

- Calculating the nominal value of R for $E(I) = 10A$ in the simple circuit.
- Deriving a general expression for the relation between R_1 and R_3 in the Wheatstone bridge that maximizes the measurement precision and applying it to the introductory example conditions ($\mu_1 = 5000$, $\sigma_1 = \sigma_3$, and resistances to be measured of around 20Ω). Section A.2 of the Appendix provides a solution.

 ACCEPTED MANUSCRIPT

- Calculating $E(\tilde{R}_x)$. It is easy to see (section A.3 on the Appendix) that the parameters of the R_3 distribution introduce a positive bias in the measurement of R_x.

Students are offered the final result of all proposed exercises, but not the way to solve them that is provided only in some cases of especial difficulty. The exercises described above are among the ones solved, with a level of detail similar to the one used in the appendix of this article.

We believe that these examples can be used to illustrate variability related issues in many different ways and be adapted to different situations and levels. We have used these examples in more advanced courses dealing with more specific issues, such as industrial statistics, quality improvement and design of experiments. They are useful to introduce important concepts, such as robustness, and to show that an appropriate parameter design can make quality features insensitive to the variability of components' characteristics. Of course, students need to have a technical background.

 ACCEPTED MANUSCRIPT

APPENDIX: THEORETICAL JUSTIFICATIONS

A.1 Circuit with only a battery and a resistor

The value needed to obtain an average intensity of $10A$ can easily be calculated expanding $1/R$ into a Taylor series around $E(R)$ up to the second-order derivatives:

$$\frac{1}{R}=\frac{1}{E(R)}-\frac{R-E(R)}{[E(R)]^2}+\frac{[R-E(R)]^2}{[E(R)]^3}+\cdots$$

$$E\left(\frac{V}{R}\right)=V\cdot E\left(\frac{1}{R}\right)\cong\frac{V}{E(R)}+\frac{V\cdot\ \mathrm{Var}(R)}{[E(R)]^3}$$

Letting $V=100V$ and $\mathrm{Var}(R)=1\Omega^2$ in the above expression, it is easy to show that 10.1Ω is the value of $E(R)$ which makes $E(V/R)=10A$.

The approximation is very good because the term corresponding to the third order derivative is zero (the numerator is the skewness of a normal distribution: $E\{[R-E(R)]^3\}=0$). And the term corresponding to the fourth order derivative gives values of the order of thousandths. Another way to check the goodness of this approximation is by simulating values of $R\sim N(10.1;1)$ and checking that the average value of $100/R$ is equal to 10.00.

A.2 The Wheatstone Bridge parameters design

Of course the measurement precision is measured by $V(\tilde{R}_x)$. From (3) we have:

$$V(\tilde{R}_x)=\left(\frac{R_x\mu_3}{\mu_1}\right)^2 V\left(\frac{R_1}{R_3}\right)+\left(\frac{\mu_3}{\mu_1}\right)^2\sigma_2^2 \quad (4)$$

Since only R_1 and R_3 are random variables, the only difficulty is finding $V(R_1/R_3)$. Again, an easy way to do it is to expand R_1/R_3 into a Taylor series around $E(R_1)$ and $E(R_3)$; up to the first-order derivatives, the result is:

$$\frac{R_1}{R_3} = \frac{E(R_1)}{E(R_3)} + \frac{1}{E(R_3)}[R_1 - E(R_1)] - \frac{E(R_1)}{[E(R_3)]^2}[R_3 - E(R_3)] + \cdots$$

Using μ and σ^2 to represent the expectation and variance we get:

$$V\left(\frac{R_1}{R_3}\right) = \frac{1}{\mu_3^2}\sigma_1^2 - \frac{\mu_1^2}{\mu_3^4}\sigma_3^2$$

Hence substituting in (4) we have:

$$V(\tilde{R}_x) = R_x^2\left[\frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_3^2}{\mu_3^2}\right] + \frac{\mu_3^2}{\mu_1^2}\sigma_2^2$$

We can see that μ_1 always appears in the denominator, and therefore we want it to be as large as possible in order to minimize $V(\tilde{R}_x)$. After having set a value for μ_1, we can find the value of μ_3 that minimizes $V(\tilde{R}_x)$ in the following way:

$$\frac{\partial V(\tilde{R}_x)}{\partial \mu_3} = R_x^2\sigma_3^2(-2)\frac{1}{\mu_3^3} + \frac{\sigma_2^2}{\mu_1^2}(2)\mu_3$$

Setting the derivative equal to zero yields:

$$\frac{\sigma_2^2\mu_3}{\mu_1^2} = \frac{R_x^2\sigma_3^2}{\mu_3^3}$$

And we get that:

$$\mu_3 = \sqrt{\frac{\mu_1 R_x \sigma_3}{\sigma_2}} .$$

To be able to measure resistances with values of around $20\ \Omega$ if $\mu_1 = 5000\ \Omega$, we should choose $\mu_3 = 316\Omega$, considering that σ_2 and σ_3 are equal.

A.3 The Wheatstone Bridge calibration

This exercise provides a nice way to "visualize" bias. Also, students can check that this bias can be compensated by a proper calibration.

From (3), we calculate:

$$E(\tilde{R}_x) = R_x \frac{\mu_3}{\mu_1} E\left(\frac{R_1}{R_3}\right)$$

In this case we also expand R_1/R_3 into a Taylor series. As we did before, we consider that the values of R_1 and R_3 are independent, and using the terms up to the second-order derivatives, we have:

$$E(\tilde{R}_x) = R_x \frac{\mu_3}{\mu_1} \left(\frac{\mu_1}{\mu_3} + \frac{\mu_1 \sigma_3^2}{\mu_3^3}\right)$$

$$= R_x \left(1 + \frac{\sigma_3^2}{\mu_3^2}\right)$$

It is then clear that the parameters of the R_3 distribution introduce a positive bias in the measurement of R_x.

REFERENCES

Bisgaard, S. (1991), “Teaching Statistics to Engineers”, *The American Statistician*, 45(4), pp. 274-283.

Box, G. and Fung, C.A. (1994), “Is Your Robust Design Procedure Robust?”, *Quality Engineering*, 6, pp. 503-514.

Box, G. (1992), “Teaching Engineers Experimental Design with a Paper Helicopter”, *Quality Engineering*, 4(3), pp. 453-459

Garfield, J. and Ben-Zvi, D. (2005), “A Framework for Teaching and Assessing Reasoning about Variability” *Statistics Education Research Journal*, 4(1), pp. 92-99.

Lee, H. (2007), “Chocolate Chip Cookies as a Teaching Aid”. *The American Statistician*, 61(4), pp. 351-355.

Taguchi, G. and Wu, Y. (1979), *Introduction to Off-line Quality Control*, Tokio: C. Japan QC Assoc.

