

Using *MINITAB* software for teaching measurement uncertainty

Andrea Zanobini

Department of Information Engineering, University of Florence (Italy)
Via Santa Marta, 3 – 50139 Florence, Italy

Email: andrea.zanobini@unifi.it

Abstract. The concept of measurement uncertainty should be regarded not only related to the concept of doubt about the validity of the measurement result, but also to the quantization of this concept. In this sense the measurement uncertainty is that parameter associated with the result characterizing the dispersion of the values that could reasonably be assigned to the measurand (or more properly to its representation through a model). This parameter may be for example a multiple of the standard deviation but especially, and more importantly, the half width of an interval with a predetermined level of confidence or trust. In these terms in this paper I attempt, with the help of *MINITAB* software, to analyze this parameter; with simple and quick operations to evaluate the mean, the standard deviation and the confidence interval and by the use of several plotted graphs.

1. Introduction: the statistical procedures in educational research

The approach to learning statistical procedures presented in this paper will be useful through the behavioural sciences. Furthermore, and more importantly, educational research is an especially suitable source of examples for an introduction to statistical methods because of the multidisciplinary nature of research in education. An investigator studying students, teachers, and schools is sometimes a sociologist, frequently a psychologist, and at times anthropologist, psychiatrist, historian, economist, or architect.

The recent explosive expansion of activity in educational research, girded strongly by public and private financial support, is another justification for attention to this field. One of the chief concern is that there is now more money available to conduct educational research than there are people prepared to design, execute, report, and utilize such investigations. This need has not been met despite the fact that the new support has attracted to educational problems many competent scholars from many disciplines who previously had neither the money nor inclination to devote themselves to the study of educational problems.

Increased statistical sophistication will help reduce another difficulty in current educational research. The fact is that rather statistics model for guiding inquiry have dominated research activity in education. Statistical analysis plays a dominant role in two major types of educational research activities: *surveys* and *experiments*. Although the student may not clearly see the distinction between them, he will see that there are many varieties of surveys and experiments. Some surveys, for example, involve simple counts on a single variable, while others comprise measures of many variables and their inter-relationships. However these data will serve as practice data so that the student can “play scientist” as he begins to master his new statistical tools.



2. An educational experiment

An educational experiment occurs when a scientist assigns subjects at random to two or more treatment groups, imposes a different educational treatment on each group, and then measures the groups to see whether the different treatments have produced important differences among the groups on one or more outcome variables.

After the period in which the group of subjects have received different educational treatments, the scientist measure all subjects on one or more outcome variables (frequently these are achievement tests), and analyzes the scores to decide whether the groups differ in outcomes.

We can best define *statistics* by considering the problems the scientist faces in performing this analysis of outcome measurements. First, he has to characterize each group by some summary of the scores (*summary statistics*) for the subjects in the group. Computing an average score for the group is an example of such a summary. Statistics is in part the science of summarizing masses of measurements to obtain useful summary *descriptions of trends in the data*, such as characterization of groups and of differences among groups. Second, the scientist has to decide whether the differences among the groups on the descriptive statistics could be due simply to the failure of his initial randomization procedure to equate the groups prior to the different treatments. He knows that if he had given precisely the same treatment to all the groups, they would nevertheless differ in outcomes by chance alone, as a result of accidents in his randomizing procedure. The other part of statistics is the science of making *inferences about the significance of trends in the data*, such as observed differences among groups. In short, statistics is the science of making *descriptions and inferences* from measurements.

3. A brief introduction to the MINITAB software for a statistical analysis

The statistical analysis with *MINITAB* software, which often relies on a good background knowledge of the phenomenon to be analyzed, requires a series of steps and the main are:

- Data Analysis with graphs
- Perform statistical analysis and procedures
- Evaluation of the quality of a measure

Here I give a brief introduction to the use of this program by inviting the reader to download the trial software [6] and to carry out the .pdf file "*Meet Minitab*" a self-training mini-course that allows a basic use of all the main functions. When we open the application, the graphical user interface - GUI - Graphical User Interface - will present two windows.

The first is related to the current session and shows the results of the analysis in text format. It is also possible to enter commands in the form of instructions rather than using the menu bar.

The data window, very similar in appearance to a spreadsheet, contains the open worksheet. Looking at that sheet you can see the classic table structure of spreadsheets with the data arranged in columns, data that can be of three types: numeric, text, date. At any time you can change the properties of these data from the menu bar.

3.1 An example: the evaluation of the quality of a measure connected to its uncertainty

For quality of a measure we refer to the degree to which products or services meet the needs of customers. The ultimate goal of every manufacturer is to reduce the rate of defective products and, therefore, their variability; they produce manufacturing products that are compliant much as possible with the specifications [4].

MINITAB offers many methods to assess the quality of a measure qualitatively and quantitatively, definitely its uncertainty: control charts, statistical tools, quality planning, process capability and reliability.

Control charts are also useful for monitoring the stability of the process in time and to detect the presence of anomalies in the same process.

For example in **Figure 1**, easily obtained with *MINITAB* software, we can see:

- Centerline at the average value of the statistic
- Upper limit control (UCL) at 3σ above the center line
- Lower limit control (LCL) at 3σ below the center line

In this case, the student is asked to perform the exercises on "*Meet Minitab*" to understand the utility of control charts in a measurement process also to evaluate its measurement uncertainty.

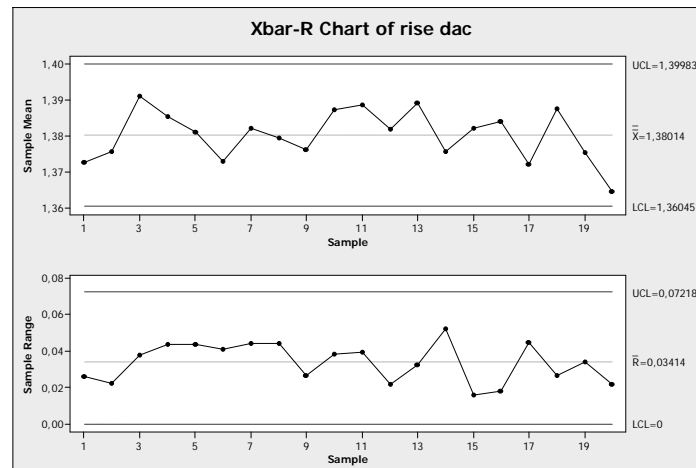


Figure 1. A *MINITAB* control chart plot for the readings of the rise times of the step response in a digital oscilloscope. How this plot is it connected to the measurement uncertainty?

4. Uncertainty evaluation with the help of statistical theory

The objective of a measurement is to determine the value of the measurand, that is the specific quantity subject to measurement. In general, no measurement or test is perfect and the imperfections give rise to errors in the result. Consequently, the result of a measurement is only an approximation to the value of the measurand and is complete only when accompanied by a statement of the uncertainty of the approximation [1]. The International Vocabulary of Basic and General Terms in Metrology defines uncertainty as “a parameter associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand” [2].

Taking into considerations all these sources of variability it is possible to elevate the quality level of the whole measurement process. It is important to reduce the uncertainty in the information and to enhance the detail of the measurand.

First of all it is important to point out that the measurement process cannot generate a rational number as result; in fact, some information connected to its validity and to the quality of the whole measurement process are necessary. Both these numbers should be referred to their measurement unit. Now it should be clear that a measurement process should be able to transform an unknown measurand in a measurement result, with all the information necessary to state its quality level.

In this process the statistical theory plays a fundamental role but some questions have to be taken into account in the parameters estimation: what about expected value, variance and standard deviation by means of sample statistics?

4.1 Measurement, probability and random variables

Concerning the uncertainty it is possible to identify two main contributes: the first is originated by random effects (*GUM - type A*); the other one depends on systematical effects (*GUM - type B*) [2]. In particular:

- The first contribute is essentially related to the stochastic variations of the influence quantities which will give different observation obtained in *repeatability* conditions. This means that the same person can collect these observations, in an independent way, with the same procedure, the same instrument, in the same conditions of use and in a restricted time interval.
- The second contribute is essentially related to those effects that are identical every time the measure is repeated and it depends, for example, on instrument accuracy and resolution, reference samples, and others.

If these two contributes are present, it is often possible to correct them and, eventually, to delete many of these effects. So the measurement can be treated like a *random variable*, M , distributed within the measurement interval. This random variable determines the measurement process and so the quality of the test results. In the paragraph 6, I will introduce the *confidence level* to attribute to every single event, associated to M in $S = \{m_{\min} \leq M \leq m_{\max}\}$, the space domain of the results.

Obviously the maximum confidence level, equal to one, can be assigned when M belongs to S ; vice versa the minimum confidence level, equal to zero, is assumed when the values of M do not belong to the space S of all the possible measurement results. Consequently, it is logical to assign a real positive number, in the range from zero to one, to the confidence level defined in a subinterval $\{m_a \leq M \leq m_b\}$ of S ; this number is named *probability* [5].

To each M is associated a probability distribution, or a function of random events which represents the probability that a measure belongs to one of the possible subintervals in which it is possible to divide the space S of all the possible measurement results.

5. Gauss/Laplace or Normal distribution

The probability distribution of the normal random variable M is expressed by:

$$f_M(m) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(m-\mu)^2}{2\sigma^2}} \quad -\infty < x < +\infty$$

Note that $f_M(m)$ contains two statistical parameters: μ is the expected value and σ denotes the standard deviation. Such distribution is denoted as *Gauss/Laplace or Normal distribution*, its plot, easily obtained with *MINITAB* software again, is shown in **Figure 2** and its notation is $M \sim N(\mu, \sigma^2)$ [3].

The graph is symmetrical with respect to μ , with higher concentration of measurement values near μ . It also appears that the density is the same for the two symmetrical points of μ ; the points on the graph at which correspond $\mu - \sigma$ and $\mu + \sigma$ in the abscissa, represent the points of inflexion of the curve.

For different values of the mean μ and same σ , the plot changes in the position with respect to the abscissa. Instead, with different σ values and same μ , the graph changes in shape maintaining the symmetrical condition; in other words, the higher σ is, the higher is the dispersion of measures close to μ , and vice versa.

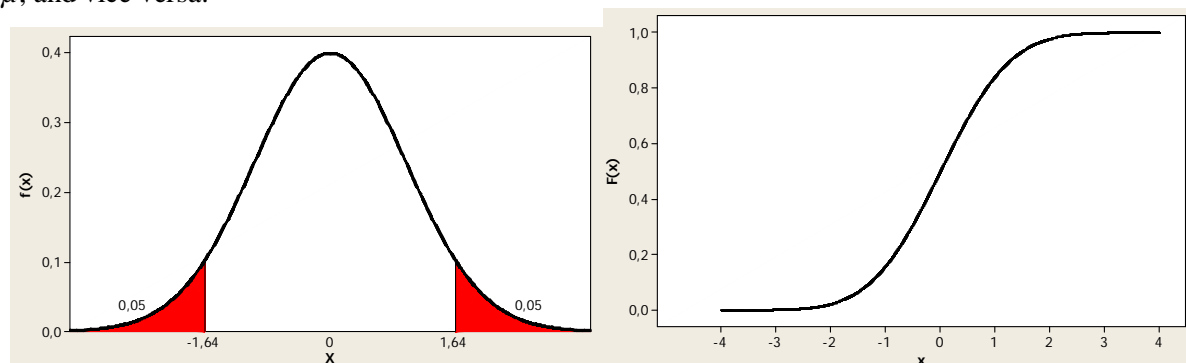


Figure 2 Probability density function (PDF) for Gauss/Laplace distribution $M \sim N(\mu, \sigma^2)$ and Cumulative distribution function for Standardized Gaussian $Z \sim N(0, 1)$.

The distribution function is:

$$P\{m_a \leq M \leq m_b\} = F_M(m_b) - F_M(m_a) = \\ = \frac{1}{\sigma\sqrt{2\pi}} \int_{m_a}^{m_b} e^{-\frac{(m-\mu)^2}{2\sigma^2}} dm = \frac{1}{\sqrt{2\pi}} \int_{\frac{m_a-\mu}{\sigma}}^{\frac{m_b-\mu}{\sigma}} e^{-\frac{m^2}{2}} dm = Z\left(\frac{m_b-\mu}{\sigma}\right) - Z\left(\frac{m_a-\mu}{\sigma}\right)$$

where $Z=(M-\mu)/\sigma$ denotes the standardized random variable [3].

Considering k the *coverage factor*, in **Table 1** the probability, in percent, that M is in the closed interval $\mu \pm k\sigma$ is shown in function of k .

$P\{\mu - k\sigma \leq M \leq \mu + k\sigma\}$	k
68,27	1
90	1,645
95	1,960
95,45	2
99	2,576
99,73	3

Table 1: Different probability values, in percent, in function of the coverage factor k .

6. Confidence interval of measurement results

The measure M , can be defined, as introduced in the previous paragraph, by the following expression:

$$P\{|M - E\{M\}| \leq k u_M\} = P\{E\{M\} - k u_M \leq M \leq E\{M\} + k u_M\} = p$$

This is the probability that the measure M is comprised within an interval that is a function of its expected value and of the standard uncertainty u_M multiplied by the *coverage factor* k . Obviously the *confidence level* p should be greater than possible, better if near to one. The interval:

$$E\{M\} - k u_M \leq M \leq E\{M\} + k u_M$$

is called *confidence interval*; it represents the interval in which the probability to find a great number of possible values of M , within this interval, is near to one.

When the probability density function of M is known, it is possible to evaluate the confidence level p by the following expression:

$$p = \int_{E\{M\} - k u_M}^{E\{M\} + k u_M} f_M(m) dm$$

It is now possible to deduce the measurement result as the **uncertainty interval**, connected to the measurand with confidence level equal to p .

It is possible to write:

$$P\{m_\alpha \leq M \leq m_{p+\alpha}\} = \int_{m_\alpha}^{m_{p+\alpha}} f_M(m) dm = F(m_{p+\alpha}) - F(m_\alpha) = p$$

where α is a value, comprised between zero and one, defined by $0 \leq p \leq p + \alpha \leq 1$. Finally:

$$F(m_\alpha) = P\{M \leq m_\alpha\} = \alpha$$

The following **Figure 3**, easily obtained with MINITAB software again, represents an example of this interesting situation.

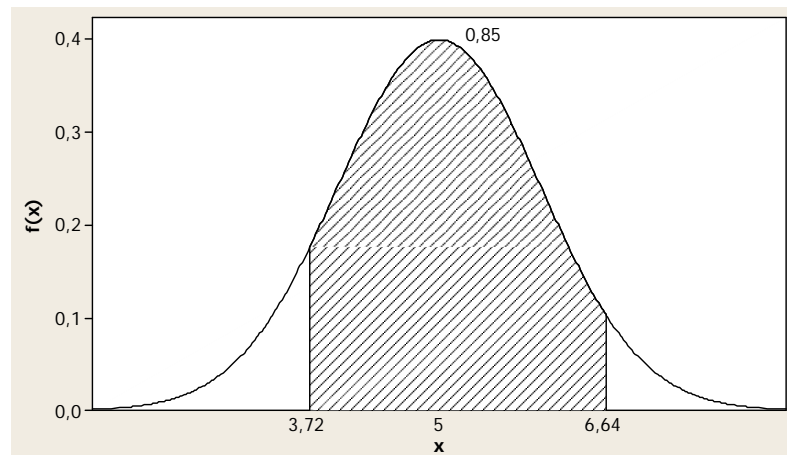


Figure 3 . PDF of a normal distribution with expected value equal to 5 and standard deviation equal to 1. The three regions under the curve, from the left, correspond respectively to $\alpha = 0,1$; $p = 0,85$ and $(1 - p - \alpha) = 0,05$.

The choice of $\alpha=(1-p)/2$ generates, a **symmetric uncertainty interval**, with p as the uncertainty level, in the sense that the three sectors showed in Figure 3 seem identical.

If the probability distribution of M is symmetrical, in the sense that its PDF is also symmetrical, this means that the distribution is symmetrically centered around the expected value of M , and the borders of the narrower interval, with confidence level equal to p , are equidistant.

In this case, introducing again the standard uncertainty of M , $u_M = \sqrt{\text{Var}\{M\}}$, the uncertainty interval can be written as $E\{M\} \pm ku_M$ where k is the coverage factor opportunely chosen to realize that $E\{M\} - ku_M$ and $E\{M\} + ku_M$ are respectively the so called $[(1-p)/2]$ -quantile and $[(1+p)/2]$ -quantile of the cumulative distribution function $F(m)$.

In case of asymmetrical distribution, all things change but, conceptually, the problem remains the same and in this case the value of α in should be different by $(1-p)/2$ in order to have the minimum amplitude of $(m_{p+\alpha} - m_\alpha)$ and the uncertainty interval is the narrowest interval with a given confidence level p .

7. Conclusions

In this paper the use of *MINITAB* software is presented as a very useful tool to introduce selected examples of behavioral sciences and education connected to “real live” data for any student practice. With the help of this software the students understand to turn the data analysis of statistics over to a computer in order to free his intelligence for the study of human problems of theory, design and interpretation in a challenging theory of knowledge.

8. References

- [1] WEB BIPM: <http://www.bipm.org/>
- [2] International Organization for Standardization, ISO/IEC Guide 98-3:2008- Uncertainty of measurement-Part 3: Guide to the expression of uncertainty in measurement, ISO, 2008.
WEB GUM: <http://www.bipm.org/en/publications/guides/gum.html>
- [3] William Navidi, “*Probability and Statistics for engineering and sciences*”, Mc Graw Hill, 2006.
- [4] Douglas C. Montgomery, “*Statistical Quality Control*”, Mc Graw Hill, January 2006.
- [5] Athanasios Papoulis, “*Probability, R.V. and Stochastic Processes*”, Mc Graw Hill, 2002.
- [6] WEB MINITAB: <http://www.minitab.com>