

Bayesian approach to uncertainty evaluation: is it always working?

Alessandro Ferrero¹, Simona Salicone^{1,a}, and Harsha Vardhana Jetti²

¹ Politecnico di Milano, DEIB - Dipartimento di Elettronica, Informazione e Bioingegneria, P.za Leonardo da Vinci 32, 20133 Milano, Italy;

² Politecnico di Milano, DENG - Dipartimento di Energia, Via Lambruschini 4, 20156 Milano, Italy

Abstract. Since the GUM has been published, measurement uncertainty has been defined in terms of the standard deviation of the probability distribution of the values that can be reasonably attributed to the measurand, and it has been evaluated using statistical or probabilistic methods. A debate has always been alive, among the metrologists, on whether a frequentist approach or a Bayesian approach should be followed to evaluate uncertainty. The Bayesian approach, based on some available a-priori knowledge about the measurand seems to prevail, nowadays. This paper starts from the consideration that the Bayesian approach is based on the well-known Bayes theorem that, as all mathematical theorems, is valid only to the extent the assumptions made to prove it are valid. The main question, when following the Bayesian approach, is hence whether these assumptions are satisfied in the practical cases, especially when the a-priori information is combined with the information coming from the measurement data to refine uncertainty evaluation. This paper will take into account some case studies to analyze when the Bayesian approach can be usefully and reliably employed by discussing the amount and pertinence of the available a-priori knowledge.

1 Introduction

A general consensus has grown in the scientific and technical community that a measurement result cannot be expressed by a single numerical value and a measurement unit.

According to the GUM [1, 2], a measurement result must be represented by a probability density function (pdf), which represents the distribution of the values that can be reasonably attributed to the measurand. Then, from this pdf, it is possible to retrieve

1. a measured value and a standard uncertainty, which are given, respectively, by the mean value and standard deviation of the pdf;
2. a confidence interval, about the measured value, that is an interval of possible measurement values along with the associated confidence level, or coverage probability.

The above information can be retrieved following two different approaches [1]: a type A evaluation of measurement uncertainty, which is generally considered to correspond to

^aCorresponding author: simona.salicone@polimi.it

the *frequentistic approach* and a type *B* evaluation of measurement uncertainty, which is generally considered to correspond to the *Bayesian approach*.

It is well known that, in the frequentistic approach, measurement uncertainty is obtained by processing a statistically significant amount of experimental data. On the other hand, in the Bayesian approach, measurement uncertainty is obtained starting from the *a priori* available knowledge, and the evaluation on the relevance of the available information is left to the *scientific judgment* of the operator.

These different points of view (Bayesian and frequentistic) have been widely discussed in the literature and they appear to be incompatible when seen under a purely mathematical and philosophical perspective [3]. However, when a more pragmatic, metrological approach is adopted, devoid of all philosophical disputes of the last three centuries, they appear to be the two sides of the same coin [4]: the instrument's manufacturer, who has the complete series of experimental data about the behavior of each manufactured instrument, can process them in a frequentistic way to find the value of measurement uncertainty associated to the measured values¹; on the other hand, the instrument's user, who cannot directly access the manufacturer's data, is somehow forced to adopt the Bayesian approach [4].

Regardless to the way measurement uncertainty is evaluated, its correctness depends on the correctness of the available information, and how well it is processed to evaluate the different contributions to uncertainty and propagate them through the measurement process [1, 5].

Ensuring that the available information is reliable is a critical key point whenever measurement results are employed as input elements in decision-making processes, such as conformity assessment. This is probably the most frequent use of measurement results, especially in industrial applications, as well as in healthcare, environment protection and many other important applications.

According to [6], no decision can be taken without the knowledge of the measurement uncertainty value associated to the obtained measured value. Therefore, in this respect, it could be thought that the GUM [1] provides the main and most important guidelines to be followed. However, [6] also states that the approach followed by the GUM to assign a pdf to a measured value is only a simplification, which applies when no *a-priori* information about the measurand is available. The more general approach is represented by the application of Bayes' theorem.

It follows that Bayes' theorem is expected to provide a better result than that provided by the commonly employed GUM approach. The aim of this paper is to further discuss this statement and the robustness of the proposed approach with respect to the correctness of the *a priori* assumptions. Simulation and experimental results will be proposed and analyzed to this purpose.

The next section recalls Bayes' theorem and tries to highlight its advantages and disadvantages when it is applied in metrology.

2 Bayes' theorem in metrology

Let us consider a random variable X , mathematically represented by its probability density function $p_X(x)$, and a random variable Y , mathematically represented by its probability density function $p_Y(y)$, where x and y represent single realizations of X and Y respectively.

If the possible mutual dependence of X and Y is also known, it is possible to define the joint probability density function $p_{X,Y}(x, y)$ of the two variables X and Y .

¹The complete series of historical data can be also seen as *a priori* knowledge and, in this case, the two approaches are identical.

In particular, it can be written:

$$p_{X,Y}(x, y) = p_X(x) \cdot p_{Y|X}(y|x) \quad (1)$$

where $p_{Y|X}(y|x)$ is the conditional probability of random variable Y , given X .

It is obvious that the conditional probability $p_{Y|X}(y|x)$ of event y given x depends on the particular value assumed by the random variable X . When the random variables X and Y are independent of each other, the conditional probability $p_{Y|X}(y|x)$ of event y given x does not depend on the value x assumed by X . Therefore, it is:

$$p_{Y|X}(y|x) = p_Y(y) \quad (2)$$

Under this assumption, it follows:

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y) \quad (3)$$

Equation (1) can be also rewritten as:

$$p_{X,Y}(x, y) = p_{X|Y}(x|y) \cdot p_Y(y) \quad (4)$$

where $p_{X|Y}(x|y)$ is the conditional probability of random variable X , given Y . $p_{X|Y}(x|y)$ depends on the particular value assumed by random variable Y , except when X and Y are independent variables. In this case:

$$p_{X|Y}(x|y) = p_X(x) \quad (5)$$

and (4) becomes the same as (3).

In the more general situation of dependent variables, because of (1) and (4), it is:

$$p_{X,Y}(x, y) = p_X(x) \cdot p_{Y|X}(y|x) = p_{X|Y}(x|y) \cdot p_Y(y) \quad (6)$$

that can be rewritten as:

$$p_{X|Y}(x|y) = \frac{p_X(x) \cdot p_{Y|X}(y|x)}{p_Y(y)} \quad (7)$$

which is the formulation of Bayes' theorem.

From the mathematical point of view, it is also:

$$p_Y(y) = \int_x p_{X,Y}(x, y) dx = \int_x p_X(x) \cdot p_{Y|X}(y|x) dx \quad (8)$$

where $p_Y(y)$ is called the marginal probability distribution.

Thanks to (8), it follows that $p_{X|Y}(x|y)$ can be obtained by simply knowing $p_X(x)$ and $p_{Y|X}(y|x)$.

Eq. (7) is very important in metrology. Indeed, let us suppose that random variable X represents the measurand and random variable Y represents the measurement result. The obtained distribution of the measured values is $p_{Y|X}(y|x)$, since it is the distribution of the measured values, given the particular measurand's value.

However, the very aim of a measurement process is not that of knowing the distribution of the measured values ($p_{Y|X}(y|x)$), but knowing the distribution of values that can reasonably be attributed to the measurand, given the measured value. In mathematical terms, this means that the aim of a measurement process is that of evaluating $p_{X|Y}(x|y)$.

This can be obtained by applying Bayes' theorem (7). In fact, according to [6], Bayes' theorem can be effectively employed to obtain the desired distribution if reliable *a-priori* information ($p_X(x)$) is available about a measurand.

In particular, the following interpretation is given to the quantities in (7):

1. $p_X(x)$ is the pdf expressing the *a-priori* information about the measurand;
2. $p_{Y|X}(y|x)$ is the pdf representing the distribution of measured values provided by the measurement process;
3. $p_Y(y)$ is the marginal pdf, given by (8);
4. $p_{X|Y}(x|y)$ is the pdf associated to the measurand, which combines the *a-priori* information with the information provided by the distribution of measured values.

The main advantage in the use of Bayes' theorem in metrology is that it yields the pdf associated to the measurand (*posterior*), by combining two different kinds of information: the available information about the measurand itself (*prior*) and the information provided by the measured values.

However, all previous considerations have been derived under a purely mathematical perspective, that does not doubt about the correctness of all considered quantities. In other words, the implicit assumption behind Bayes' theorem is that there is full belief on both the *a-priori* knowledge and the distribution of measured values. This means full knowledge about the possible distribution of the measurand values and zero doubts about the correctness in the evaluation of the uncertainty associated to the measured value.

It can be readily perceived that this is not the situation in practical applications. Even if the measurand variability is known, because, for instance, the production process is known, it may always drift from the normal operating conditions. Similarly, the measurement process may deviate from its expected operating conditions, thus making the evaluated uncertainty incorrect. It is then important to understand the consequences of a lack of total trust on the *a-priori* knowledge or the measured values.

This paper considers, with both simulations and experimental data, the effect of the application of Bayes' theorem in different possible situations:

1. when one can assign total belief on both the *a-priori* knowledge and the measured values;
2. when no total belief can be assigned to the *a-priori* knowledge $p_X(x)$; this means that a deviation might occur in the measurement model, or the measurement model does not represent exactly the measurement process;
3. when no total belief can be assigned to the assumed distribution of measured values; this means that the measured values might deviate from the expected distribution due, for instance, to a drift in the employed measurement instrument.

This paper also considers the risk in taking a wrong decision (false acceptance or false rejection of a product) when the measurement results are employed in conformity assessment [6]. A comparison is made, under different assumption, in the obtained risks values:

- when only the distribution of measured values is taken into account without applying the Bayes' theorem to estimate the measurand value;
- when Bayes' theorem is applied, that is when the *a posteriori* values are considered as the measurement result.

3 Simulations

A typical example of conformity assessment is considered to discuss the robustness of Bayes' approach: the resistance of the manufactured resistors is measured to check whether it is inside the tolerance limits or not. The *a-priori* information is represented by the expected dispersion of the resistance values of the manufactured resistors due to variations in the production process.

A nominal value $R_{nv} = 15 \Omega$ is assumed for the resistors, and the given production tolerance is $\pm 0.5\%$, that is $\pm 75 \text{ m}\Omega$. The dispersion of values due to the production process is supposed to be represented by a normal probability distribution, where the $\pm 3\sigma$ coverage interval is assumed to be the same as the tolerance interval. Therefore, in the considered example, the *prior* distribution is given by:

$$p_X(x) = \mathcal{N}(R_{nv}, \sigma_{nv}^2) \quad (9)$$

where $R_{nv} = 15 \Omega$ and $\sigma_{nv} = 25 \text{ m}\Omega$.

It is then assumed that the resistance is measured by a Fluke 8845A, 6.5 digit precision multimeter². According to the manufacturer specifications, the following applies:

- the measured 15Ω value falls in the 100Ω range;
- in the 100Ω range, an accuracy interval

$$\pm 0.003\% \text{ of reading} \pm 0.003\% \text{ of range}$$

is provided, which corresponds to:

$$\pm R_m \cdot 3 \cdot 10^{-5} \pm 3 \cdot 10^{-3} \Omega$$

where R_m is the resistor measured value.

The distribution of the possible measured values is assumed to be a normal distribution, with mean value equal to the measured value R_m and standard deviation given by:

$$\sigma_m = 0.001\% \text{ of reading} \pm 1 \text{ m}\Omega = \pm R_m \cdot 1 \cdot 10^{-5} \pm 1 \cdot 10^{-3} \Omega$$

If, for instance, $R_m = R_{nv}$, then it is $\sigma_m = 1.15 \text{ m}\Omega$.

10000 values have been considered, in each simulation, for the measured values, and the three following cases have been considered.

3.1 Case I: no deviation in the instrument or the process

Case I considers the situation in which neither the instrument nor the process is deviating.

According to the assumptions given above in Sec. 3, an *prior* normal pdf with mean value R_{nv} and standard deviation σ_{nv} is associated to each resistor. Then, since the process is not deviating, the "simulated true values" R_{STV} of the 100000 resistors have been simulated by randomly generating 100000 values from this normal pdf.

On the other hand, since the instrument is not deviating, to simulate the measurement of each resistor (R_m), a random value is generated from a normal pdf centered on the corresponding "simulated true value" R_{STV} and with standard deviation σ_m . Hence, a normal pdf centered at R_m with the corresponding standard deviation σ_m is associated to every measured value.

²The behavior of this instrument was simulated, because this same instrument was used in the experimental validation process.

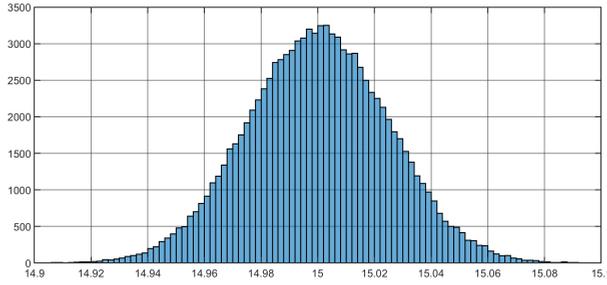


Figure 1. Histogram associated to the obtained 100000 *a posteriori* values, in case I.

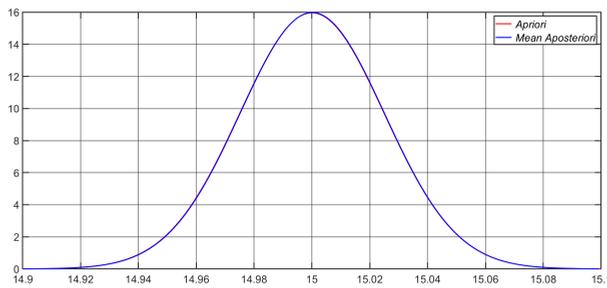


Figure 2. *A priori* pdf associated to the resistors (red line) and *a posteriori* pdf (blue line), when case I is considered.

Bayes’ theorem is then applied, according to the measurement data and the *a priori* knowledge.

It is well-known that the mean value of the *posterior* pdf, according to the given measurement data y and the *a priori* knowledge is given by:

$$\mu_{posterior} = \frac{\mu_{a\ priori} \cdot \sigma_y^2 + y \cdot \sigma_{a\ priori}^2}{\sigma_y^2 + \sigma_{a\ priori}^2} \tag{10}$$

Considering the proposed example, the quantities in (10) take the following values:

- $\mu_{a\ priori} = R_{nv} = 15 \ \Omega$;
- $\sigma_{a\ priori} = \sigma_{nv} = 25 \ \text{m}\Omega$;
- $y = R_m$;
- $\sigma_y = \sigma_m$

Fig. 1 shows the histogram associated to the 100000 obtained *a posteriori* mean values $\mu_{posterior}$, given by (10). These values represent the *a posteriori* values associated to the 100000 resistors ($R_{a\ posteriori}$).

It can be readily checked that the obtained histogram approximates quite well a normal pdf, as expected since normal pdfs have been assumed. The mean and standard deviation associated to the 100000 obtained *a posteriori* means are evaluated and the normal pdf, drawn in blue line in Fig. 2, is obtained. This pdf is compared with the *a priori* pdf (red line in Fig. 2).

It can be seen that the two pdfs are in perfect agreement. This result was expected, since no deviation in the measuring instrument or the process is considered in this case I.

It is also possible to perform a risk analysis, according to $R_{a\text{ posteriori}}$. In fact, for every obtained value, it is possible to verify whether it falls inside the tolerance limit ($R_{\text{nv}} \pm 75 \text{ m}\Omega$) or not. If the value falls inside the limit, the resistor is within the tolerances and should be accepted; on the other hand, if the value falls outside the limit, the resistor is supposed to exceed the tolerances and should be rejected.

Of course, due to measurement uncertainty, the actual resistance value can differ from the measured one. Therefore, if this last value is considered in the comparison with the tolerance limits, there is always a risk that a bad resistor is erroneously accepted, or that a good resistor is erroneously rejected. In order to state if a decision is correct or wrong, for every resistor, we also compared the corresponding “simulated true value” R_{STV} with the tolerance limit ($R_{\text{nv}} \pm 75 \text{ m}\Omega$) and verified if the two values ($R_{a\text{ posteriori}}$ and R_{STV}) lead to the same decision. Of course, the correct decision is the one obtained when R_{STV} , which is the *true value* of the resistor, is considered.

The percentage risk of taking a wrong decision is defined as:

$$Risk_{\text{total}} = \frac{\text{Total wrong decisions}}{\text{Total decisions}} \cdot 100 \quad (11)$$

When only false acceptances are considered, the following risk can be defined, which represents the risk that bad resistors are erroneously accepted:

$$Risk_{\text{f. a.}} = \frac{\text{Total false acceptances}}{\text{Total decisions}} \cdot 100 \quad (12)$$

On the other hand, when only false rejections are considered, the following risk can be defined, which represents the risk that good resistors are erroneously rejected:

$$Risk_{\text{f. r.}} = \frac{\text{Total false rejections}}{\text{Total decisions}} \cdot 100 \quad (13)$$

Of course, it is:

$$Risk_{\text{total}} = Risk_{\text{f. a.}} + Risk_{\text{f. r.}}$$

Considering the values provided by Eq. (10) in this case I, the following percentage risk values have been obtained:

$$Risk_{\text{total}} = 0.038\% \quad Risk_{\text{f. a.}} = 0.02\% \quad Risk_{\text{f. r.}} = 0.018\%$$

The total risk can be compared with the corresponding total risk value evaluated without applying the Bayes theorem, when the “measured values” R_m are taken into account and compared with the tolerance interval $R_{\text{nv}} \pm 75\text{m}\Omega$:

$$Risk_{\text{total}} = 0.038\%$$

The obtained values are perfectly compatible, as expected since no deviation in the instrument or the process is supposed in this case I. Therefore, in this case I in which no deviation is present in the process or the instrument, the application of Bayes’ theorem does not modify the results of the risk analysis.

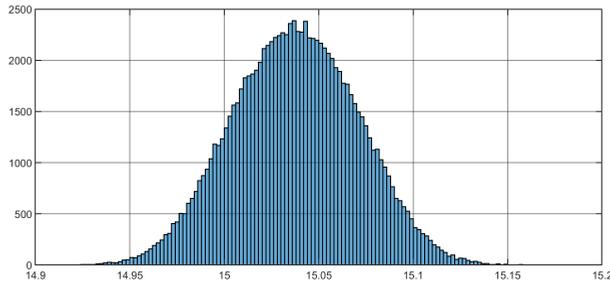


Figure 3. Histogram associated to the obtained 100000 *a posteriori* values, in case II.

3.2 Case II: deviation in the process

Case II considers the situation in which the process is deviating.

In this second case, to consider the process' deviation, the “simulated true values” cannot be extracted by the *a priori* pdf. Therefore, 100000 normal pdfs has been considered with:

- mean value

$$R_{nv} + (0.75 \cdot k \mu\Omega)$$

where $k = 1 \dots 100000$;

- standard deviation σ_{nv}

In other words, the same shape and standard deviation of the *a priori* pdf are considered, but the mean value is not centered on the resistors' nominal value R_{nv} , but it is shifted by value $0.75 \cdot k \mu\Omega$. The value $0.75 \mu\Omega$ has been chosen so that, at the end of the 100000 iterations, the process is deviated by exactly 75 m Ω which is the tolerance limit (see Sect. 3).

Hence, the 100000 “simulated true values” R_{STV} are obtained by randomly extracting one value from each of the above pdfs.

On the other hand, since the instrument is not deviating, to simulate the measured value of each resistor (R_m), a random value is generated from a normal pdf centered on the corresponding “simulated true value” R_{STV} and with standard deviation σ_m . Hence, a normal pdf centered on R_m with the corresponding standard deviation σ_m is associated to every measured value.

Bayes' theorem is then applied, according to the measurement data and the *a priori* knowledge, as given by Eq. (10). In the application of this formula, with respect to previous case I, the *a priori* knowledge is always the same, but the measurement values R_m are different, as explained above. The histogram in Fig. 3 is obtained, associated to the obtained resistors' values $R_{a \text{ posteriori}}$.

Also in this case II, the histogram approximates quite well a normal distribution and the pdf in blue line in Fig. 4 represents the corresponding pdf. In this same figure, the *a priori* pdf is also reported (red line). It can be seen that the blue line is shifted in the direction in which the process is deviating. Moreover, it has a higher variance than the *a priori*.

As in case I, it is also possible to perform a risk analysis, according to the obtained values $R_{a \text{ posteriori}}$ (that is, when Bayes' theorem is applied) and according to the measured values R_m (that is, without applying Bayes' theorem). The following values are obtained, as far as the total risk is concerned:

$$Risk_{total} = 0.64\%$$

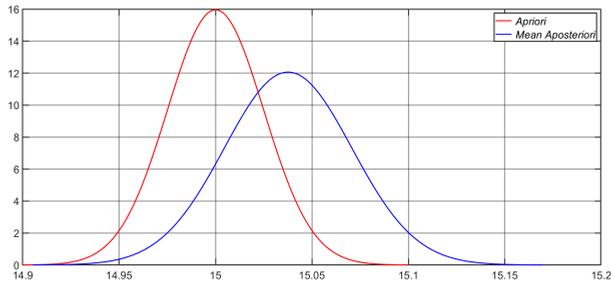


Figure 4. *A priori* pdf associated to the resistors (red line) and a *posteriori* pdf (blue line), when case II is considered.

when $R_{a\text{ posteriori}}$ is considered;

$$Risk_{\text{total}} = 0.59\%$$

when R_m is considered.

It can be noted that the risk increases a little bit after the use of Bayes’ theorem. Therefore, the application of Bayes’ theorem, in this case, increases the risk of wrong decision.

It is worth noting that the amount of risk is strictly dependent on the accuracy of the instrument. The greater the accuracy of the instrument with respect to the *a priori*, the lesser the correction due to Bayes’ theorem. Therefore, the lesser the difference in the risks with and without the application of Bayes’ theorem. Therefore, with less accurate instruments, the risk of taking wrong decisions after the application of Bayes’ theorem could become significant.

This is maybe a trivial conclusion, since very accurate instruments do not require, in principle, corrections. On the other hand, this conclusion must be taken into careful consideration if the correction obtained by the application of Bayes’ theorem is used to counterbalance possible drifts in the measurement process. If this is the case, the belief in the stability of the manufacturing process must be higher than that in the stability of the measurement process not to have unpleasant surprises.

3.3 Case III: Deviation in the instrument

Case III considers the situation in which the instrument is deviating.

In order to simulate the true values of the resistors, since the process is not deviating, the same considerations as in case I apply. Therefore, 100000 extractions are taken from the *a priori* pdf. These are the “simulated true values” R_{STV} .

On the other hand, the measurements R_m must consider that the instrument is deviating. Therefore, 100000 normal pdfs has been considered with:

- mean value

$$R_{STV} + (0.75 \cdot k \mu\Omega)$$

where $k = 1 \dots 100000$;

- standard deviation σ_m

In other words, all pdfs have the same shape and standard deviation, but their mean value drifts. The value $0.75 \mu\Omega$ has been chosen so that, for the last 100000th resistor, the measured value is deviated by exactly $75 \text{ m}\Omega$, which is the resistor tolerance limit (see Sect. 3).

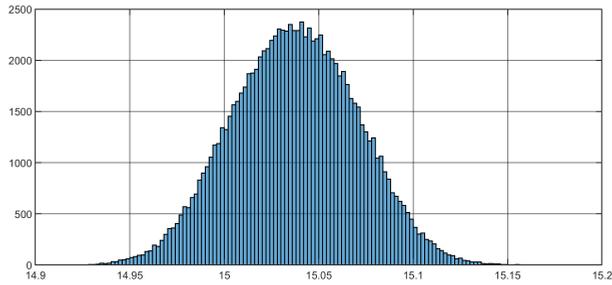


Figure 5. Histogram associated to the obtained 100000 *a posteriori* values, in case III.

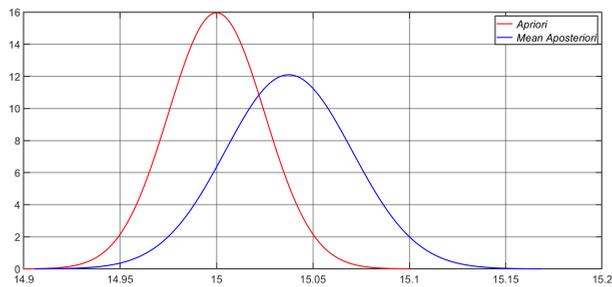


Figure 6. *A priori* pdf associated to the resistors (red line) and *a posteriori* pdf (blue line), when case III is considered.

Hence, the 100000 measurement results R_m are obtained by randomly extracting one value from each of the above pdfs. Then, for every measured value, a normal pdf with mean value R_m and standard deviation σ_m is assumed.

Bayes' theorem is then applied, according to the measurement data and the *a priori* knowledge, as given by Eq. (10). As already stated above, the *a priori* knowledge is always the same, but the measurement values R_m are different. The histogram in Fig. 5 is obtained, associated to the obtained resistors' values $R_{a\text{ posteriori}}$.

Also in this case III, the histogram approximates quite well a normal distribution and the pdf in blue line in Fig. 6 represents the corresponding pdf. In the same figure, the *a priori* pdf is also reported (red line). It can be seen that the blue line is shifted in the direction in which the instrument is deviating. Moreover, it has a higher variance than the *a priori*.

As in the previous cases, a risk analysis has been performed, and the risks of wrong decisions, obtained when the measurements R_m and when the *a posteriori* values $R_{a\text{ posteriori}}$ are considered, are compared with each other. In particular:

$$Risk_{\text{total}} = 13.14\%$$

when $R_{a\text{ posteriori}}$ are considered;

$$Risk_{\text{total}} = 13.26\%$$

when R_m are considered.

The difference between the obtained values for the total risk is small, even in this case III. However it can be noted that, in this case, the application of the Bayes' theorem yields better results.

Once again, it can be underlined that the risk depends on the accuracy of the instrument. The greater the accuracy of the instrument with respect to the *a priori*, the lesser the correction due to Bayes's theorem and, consequently, the lesser the difference in the risks with and without the application of Bayes' theorem. Therefore, with a less accurate instrument, the improvement due to the application of Bayes' theorem could be significant.

It is also interesting to observe that:

$$Risk_{total} = 13.14\% \text{ and } Risk_{f_r} = 13.02\%$$

thus meaning that, among all the wrong decisions, most of them correspond to false rejection, because of the wrong measurements due to the instrument's deviation.

4 Experimental results

In order to verify whether the simulation results are experimentally confirmed, so that the method based on bayes' theorem can be applied in practice, 80 resistors have been used with resistance value, given by the manufacturer specifications: $15 \pm 0.15 \Omega$.

First of all, the values of the resistors have been measured with a reference multimeter: Fluke 8508A with 8.5 digits.

The obtained measurements have been considered the conventionally true values of the resistors. The mean of all conventionally true values has been found to be 14.94Ω . This value is indeed within the manufacturer tolerance interval but, ideally, it was supposed to be 15Ω . So, according to the three case studies described above, we are in the case in which "the process is deviating".

Moreover, for the sake of experiment, we also made a different assumption. Since the measured true values of the resistors were all inside the tolerance interval given by the manufacturer, in order to simulate some outliers, we assumed a stricter manufacturer specification: $15 \pm 0.075 \Omega$. Also under this different assumption, the mean of the true values (14.94Ω) falls within the tolerance interval.

Therefore, it is supposed that the *a priori* knowledge is represented by a normal pdf with mean value 15Ω and standard deviation $25 \text{ m}\Omega$.

At this point, we measured the resistance values of the 80 resistors with the precision multimeter Fluke 8845. The obtained values are the resistors' measured values. Taking into account these measured values and the *a priori* knowledge, Bayes' theorem has been applied to get the *a posteriori* mean values according to (10), where:

- y is the measured value of the specific resistor;
- σ_y is the uncertainty associated to the resistor's measured value and, according to the Fluke 8845 precision multimeter specifications, it is calculated as:

$$\sigma_y = \frac{1}{3} \cdot \frac{0.003 \cdot y + 0.003 \cdot 100}{100} \quad (14)$$

where it is assumed that the total error of the instrument, given by the manufacturer's accuracy specifications, corresponds to the $\pm 3\sigma$ interval;

- $\mu_{a \text{ priori}} = 15 \Omega$, that is the nominal value given by the manufacturer;
- $\sigma_{a \text{ priori}} = 25 \text{ m}\Omega$, as explained above.

The distribution of the *a posteriori* mean values is shown in Fig. 7 in blu line, while the red line shows, for a comparison, the distribution of the *a priori* knowledge.

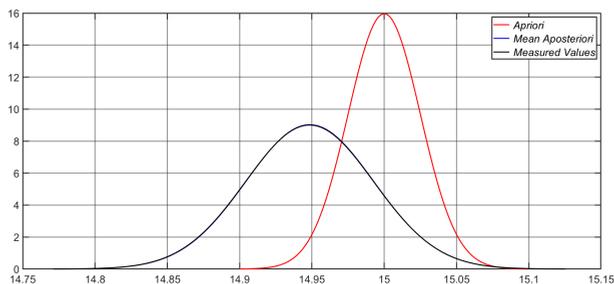


Figure 7. *A priori* pdf (red line) and *a posteriori* pdf (blue line) in the first experimental study, where the process is deviating.

The obtained *a posteriori* mean values have been stored and compared with the tolerance interval $15 \pm 0.075 \Omega$ in order to state whether the resistor is compliant with the specifications or not. This is what happens in the industrial world (in a resistor manufacturing company, in this case), where a decision needs to be made (accept or reject) whether the product is compliant or not with the specifications.

A risk analysis has been also made according to (11), by considering, firstly, the measured values and, subsequently, the *a posteriori* values. The following results have been obtained:

$$Risk_{total} = 1.27\%$$

when the *a posteriori* means are considered;

$$Risk_{total} = 0\%$$

when the measured values are considered.

This means that the application of Bayes' theorem worsens the results of the risk analysis, in this case. It is worthwhile to observe that the entire 1.27% risk is entirely due to false acceptances. So, actually, consumer risk has been increased because of the application of Bayes' theorem.

After this first experimental validation, which considers the case of a “deviation in the process”, a “deviation in the instrument” has been simulated, by adding an increasing offset to the resistors.

In particular, we increased the offset by $5 \text{ m}\Omega$ every 5 resistors starting with an offset of $5 \text{ m}\Omega$ for the first 5 resistors. According to this assumption, a final offset of about $80 \text{ m}\Omega$ is added at the end of the measurement procedure, which is a little above 0.075Ω , the 3σ interval of the considered *a priori* pdf.

The offset has been added to the resistors' measured values taken with the Fluke 8845A multimeter, as stated above, and the obtained values have been considered as the new measured values. These values have been used in Bayes' formula (10) to obtain the new *a posteriori* mean values.

The distribution of the obtained *a posteriori* values is shown in Fig. 8 in blue line, while the red line represents the pdf associated to the *a priori* knowledge and the black line represents the actual pdf, according to the resistors' true values.

A risk analysis has been done, according to both the new measured values and the *a posteriori* values. The following results are obtained:

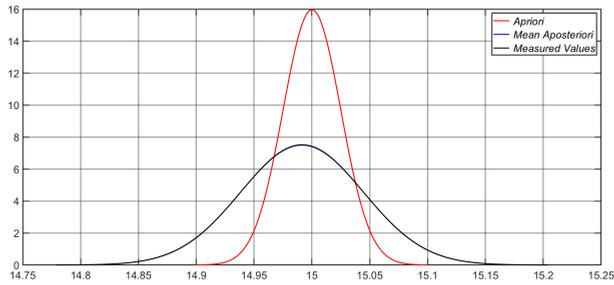


Figure 8. *Prior* pdf (red line), *posterior* pdf (blue line) and *measured values* pdf (black line), in the second experimental study, where both the process and the instrument are deviating.

$$Risk_{total} = 26.58\%$$

when the *a posteriori* means are considered;

$$Risk_{total} = 26.58\%$$

when the measured values are considered.

This result (equal total risk) should not be surprising: indeed, in this case, the employed instrument is quite accurate and the variance of the pdf expressing the distribution of the possible measured values is significantly lower than that of the *prior* pdf. Therefore, the correction provided by the application of the Bayes' theorem is negligible and does not go in favor of a risk reduction.

Moreover, if the partial risks $Risk_{f.a.}$ and $Risk_{f.r.}$ are analyzed, we get:

$$Risk_{f.a.} = 22.78\% \text{ and } Risk_{f.r.} = 3.8\%$$

when the *a posteriori* mean values are considered; and

$$Risk_{f.a.} = 22.78\% \text{ and } Risk_{f.r.} = 3.8\%$$

when the measured values are considered. So, also the partial risks do not differ. It is important to underline that, out of the total 26.58%, 22.78% of the risk is due to false acceptances, while only 3.8% is due to false rejection. This happens because the instrument is deviating in the opposite direction with respect to the process' deviation. So, the bias caused by the process deviation is compensated by the bias caused by the instrument deviation. Therefore, in this case Bayes' theorem does not improve the measurement result and does not reduce the risk of wrong decision.

5 Conclusions

The application of Bayes' theorem in metrology has been discussed, on the basis of theoretical considerations and a simple practical example, analyzed both in terms of simulation results and experimentally.

The practical example confirmed the theoretical analysis: the validity of the results provided by the application of the Bayes' theorem highly depends on the validity of the assumptions made. Since the main hypothesis of Bayes' theorem is the full credibility of the *a priori*

knowledge, that is the validity of the random variable expressing it, whenever this fails, the final result may fail too.

Despite this conclusion may sound quite pessimistic and doubtful of the actual utility of Bayes' theorem in metrology, it should sound only as a warning to check all assumptions before trusting the results blindly, as the best practice in measurement suggests.

The worst situation, because the most difficult to detect, is the use of untrustworthy *a-priori* knowledge that may happen, for instance, when the available information on the production process is considered, but the process itself is deviating from its normal operating conditions.

However, a deviation in the production process should manifest itself in several “disagreements” between the expected data and those actually measured in several points along the production line. In the *big-data* and *Internet-of-Things* era, all those data should be available for different kinds of post-processing activities. Therefore, a diligent comparison of the production data at the different points along the production line must show a drift in the production process, thus allowing the metrologist to implement the needed corrections.

Once more, this is a step forward in the direction of the best practice in measurement, since it exploits as much “relevant information” as possible in providing a measurement result.

References

- [1] JCGM 100:2008, *Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement, (GUM 1995 with minor corrections)*, Joint Committee for Guides in Metrology (2008), 1
- [2] JCGM 101:2008, *Evaluation of measurement data – Supplement 1 to the Guide to the expression of uncertainty in measurement – Propagation of distributions using a Monte Carlo method*, Joint Committee for Guides in Metrology (2008)
- [3] N. Giaquinto, L. Fabbiano, A. Trotta, G. Vacca, *About the Frequentist and the Bayesian Approach to Uncertainty*, in *Imeko TC4* (Benevento, Italy,, 2014)
- [4] G.B. Rossi, F. Crenna, *Probabilistic measurement modeling may overcome the opposition between the bayesian and the frequentistic view*, in *IMEKO TC7* (St. Petersburg, Russia, 2019)
- [5] A. Ferrero, S. Salicone, *IEEE Trans. Instrum. Meas.* **67**, 607 (2018)
- [6] JCGM 106:2012, *Evaluation of measurement data – The role of measurement uncertainty in conformity assesment*, Joint Committee for Guides in Metrology (2012)