Bayesian Statistics to Improve Reference Values in Metrology in Chemistry: an ERANET-Plus Project

The European Framework

In May 2007, EURAMET e.V., the European Association of Institutes of Metrology, submitted the proposal for an ERANET Plus (European Research Area Network) funded European Metrology Research Program (EMRP) to the European Commission. The three pillars of the EMRP are fundamental metrology, innovation and quality of life, and technologies of the future with the aim of demonstrating that coordination of European Metrology is sustainable. As a part of this program, a horizontal targeted program of the EMRP has been devised on metrology underpinning “health”. Indeed, not only are health advances a policy priority in Europe but also an area where metrology can make a significant impact swiftly.

Starting from the observation that data and uncertainty analysis in the field of laboratory medicine and clinical chemistry are not yet well established, European chemical experts have developed Joint Research Projects (JRPs). Among them the JRP entitled "Tracebioactivity" has been selected so that the European Commission will provide its financial support, and it is not incidentally that this JRP includes a work package "Modelling and data analysis: - New methods and algorithms employing a Bayesian approach to compute reference values from interlaboratory comparisons in laboratory medicine” which will be conducted by PTB (Germany), LNE (France) and SP (Sweden) by 2008.

An important aspect underlying this project is the application to Certified Reference Materials (CRM), so that this paper aims to shed light on how Bayesian statistics can improve the reliability of the certified values of reference materials. Before going any further let us first have a brief look at the crucial role of CRMs.

Certified Reference Materials: the tool of the chemical analyst

Accreditation of laboratories for chemical analyses is an essential quality management tool and a means to obtain confidence and comparability of results. Due to the introduction of the international standard ISO/CEI 17025, field laboratories have expressed their need of specific traceability schemes to ensure the reliability of chemical measurements. Actually these traceability schemes are mainly based on the use of CRMs. These CRMs are usually produced by National Metrology Institutes (NMIs) and provided to laboratories for different specific purposes of Quality Control and Quality Assurance. In addition, NMIs have implemented so called primary methods of measurements which ensure the traceability of measurements to the SI and that are also used by NMIs to assign the certified value to materials.

Value-assignment can also be based on data from an interlaboratory study. Consensus values are based on the results of all participating laboratories or only on selected laboratories. According to the development or improvement of analytical techniques by manufacturers and laboratories, NMIs are used to certify their reference materials from time to time in order to obtain more reliable and accurate certified values through a re-certification process.

Re-certification: situation

Re-certification of a certified reference material can be performed either through interlaboratory comparisons or by a single laboratory, typically a NMI carrying out a primary method of measurement. In both cases old and new data are at hand. Old data are past measurements used to compute the past certified values. When they were computed, the past certified values represented our best knowledge of the true value of the CRM. Nevertheless, by now this valuable information have been left out when computing each new certified value.

Challenge: To take into account the most data at hand to get more accurate certified values and associated uncertainties and thus to improve traceability.

Answer:: Bayesian statistics have a widespread approach to science and technology. Recently researchers have been trying to apply it to metrology and, more specifically to metrology in chemistry.

Bayesian Approach

Bayesian approach is a learning method based on the Bayes’ formula, designed to update our knowledge of a quantity through the combination of prior information about this quantity and new observations. As a result, our knowledge, called posterior, about the targeted quantity and its associated uncertainty is improved. More specifically Bayesian estimates give a complete description of the uncertainty given observations and prior knowledge. Here prior means before observations are made, so that prior information is elaborated from expert judgements and past results. This prior information is modelled by a probability density function (pdf) which is meant to best represent our degree of belief about the possible values of the quantity before new observations.

So what’s new with Bayesian approach? Unlike classical frequentist statistics which works with the sole new observations modelled by the likelihood, Bayesian approach allows, via the introduction of a prior, to take into account the most information at hand within a specific statistical modelling framework to obtain more accurate estimates of parameters.

Bayes’ Formula : from prior to posterior

Let θ be the random variable modelling the targeted quantity and $x = (x_1, \ldots, x_n)$ be n new measurements. Let us define the following notations for the different notions introduced above:

- The prior knowledge about θ is modelled by the pdf $\pi(\theta)$.
- The posterior updated knowledge is modelled by the pdf $\pi(\theta|x)$.
- The data are still modelled by the likelihood $p(x|\theta)$.

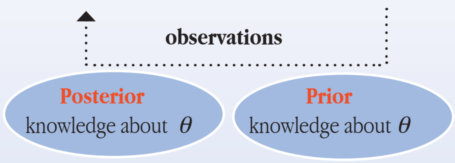
The Bayes’ formula is then given by:

$$\pi(\theta|x) = \frac{p(x|\theta)\pi(\theta)}{p(x)}$$

where $p(x)=\int p(x|\theta)\pi(\theta)d\theta$ is the normalizing constant so that the left member integrate to unity and be a density.

Using the symbol of proportionality $\propto$, it appears then more clearly that the posterior density is proportional to the prior times the likelihood:

$$\pi(\theta|x) \quad \propto \quad p(x|\theta) \quad \times \quad \pi(\theta)$$



Estimates of the mean and the variance of the quantity are then given by the mean and the variance of the posterior distribution $\pi(\theta|x)$ and are usually computed numerically.

Let's give an easy example. Just consider that the prior pdf of the quantity θ and the pdfs of the observations $x_1,\ldots,x_n$ are both Gaussian as described below:

$$x_i|\theta \sim N(\theta,\sigma^2)$$
$$\theta \sim N(\mu,\tau^2)$$

In that case the posterior distribution of θ is:

$\theta|x_1,\ldots,x_n \sim N(\theta_1,\phi_1)$

Where: (P)

$$\theta_1 = \frac{\sigma^2/n}{\sigma^2/n+\tau^2}\mu + \frac{\tau^2}{\sigma^2/n+\tau^2}\bar{x} \text{ and } \phi_1^{-1} = \frac{1}{\tau^2} + \frac{1}{\sigma^2/n}$$

Notice that:

- The posterior mean θ_1 is the weighted mean of the prior mean μ and of the sample mean of the observations $\bar{x}$ where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{x}|\theta \sim N\left(\theta, \frac{\sigma^2}{n}\right)$
- The posterior inverse variance ϕ_1^{-1} is the sum of the inverse variance of the prior $\frac{1}{\tau^2}$ and of the inverse variance of the sample variance of the observations $\frac{1}{\sigma^2/n}$.

which shows that Bayesian approach is a compromise between the prior and the observations.

Bayesian Approach in Metrology in Chemistry

Undoubtedly Bayesian approach can be applied to the recertification process of CRMs. In this case the recertified value and its associated uncertainty are the updated values of the past certified values and their past associated uncertainty through Bayesian analysis. In particular the recertification process through intercomparisons is illustrated in the Figure 1.

In the case where all the distributions are Gaussian, the example above may be considered as a special case. Indeed if one supposes that:

- θ is the certified property
- the x_i $i=1\ldots n$ are the results of a sole laboratory (case N=1)
- the x_i are issued from the same Gaussian distribution (mean θ, variance σ^2)

then the posterior distribution of θ is given by expression (P).

This particular model can thus be applied when recertification is performed by one laboratory only.

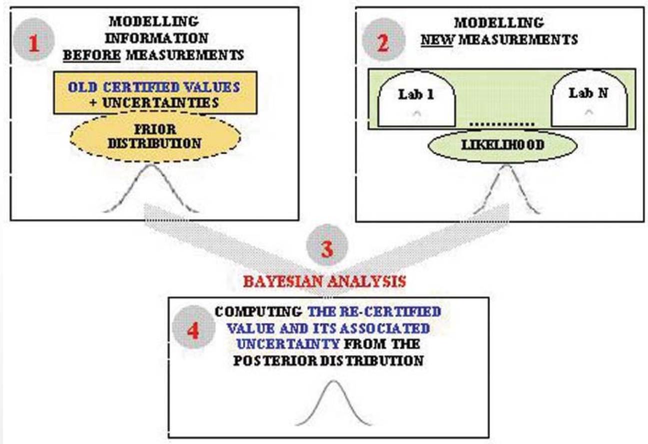


***Figure 1:** Recertification of a certified property of a CRM in a Bayesian statistical framework*

When N laboratories are involved the expression of the posterior distribution is more complicated because the expression of the likelihood is then heavier. Indeed, now let x_{ij} be the j^{th} measurement of the i^{th} laboratory where each laboratory makes n_i measurements that is j is between 1 and n_i for i between 1 and N. In addition, suppose that the variances of the measurements in each laboratory are different, say σ_i^2 for laboratory i. The model then becomes:

$$x_{ij}|\theta \sim N(\theta,\sigma_i^2)$$
$$\theta \sim N(\mu,\tau^2)$$

The posterior distribution of θ is then given by:

$$\theta|x_{ij} \sim N(\theta_N,\phi_N) \quad i=1\ldots N \quad j=1\ldots n_i$$

Where:

$$\theta_N = \frac{\sum_{i=1}^{N}\frac{1}{\sigma_i^2/n_i}\bar{x}_i + \frac{1}{\tau^2}\mu}{\sum_{i=1}^{N}\frac{1}{\sigma_i^2/n_i} + \frac{1}{\tau^2}} \text{ and } \phi_N^{-1} = \sum_{i=1}^{N}\frac{1}{\sigma_i^2/n_i} + \frac{1}{\tau^2}$$

We observe that the posterior estimates of the mean and the inverse variance are still a weighted mean and the sum of the inverse of the variances respectively. In this respect it is worth noticing that this time, not only do we obtain a weighted mean between the prior and the observations but also the term standing for the observations is a weighted mean of the means of the laboratories. Thus all the information available has been taken into account and weighted according to the new measurements and prior knowledge. And that's precisely this idea of compromise that makes Bayesian approach so attractive.

Improving Reliability in Recertification: providing confidence on measurements

In conclusion, working on new measurements together with previous reliable measurements gives more accurate and reliable updated values. That way the new certified value and its associated uncertainty is consistent with the previous certified values and their uncertainties. In addition, when updating the certified value, metrologists increase the traceability of CRMs by improving the knowledge of the best estimate of the true value of the CRM.

Finally many decisions in the clinical sector are based on results of content measurements. That's the reason why the JRP project will first investigate the requirements of the European clinical sector for improved mathematical tools and then intends to make the demonstration that Bayesian probability theory is really the most promising approach to provide appropriate solutions for the improvement of reliability and comparability of measurements.

Séverine Demeyer
LNE-CNAM, France

Dr. Philippe Charlet
LNE, France