THE EXPRESSION OF THE MODEL UNCERTAINTY IN MEASUREMENTS

An Application of the Ellipsoidal Nested Sampling

C. Palmisano[1] and G. Mana[2]

1: UNITO - Università di Torino, Dipartimento di Fisica, Via P. Giuria, 1 - 10125 Torino, Italy palmisan@to.infn.it
2: INRIM - Istituto Nazionale di Ricerca Metrologica, Str. delle Cacce, 91 - 10135 Torino, Italy g.mana@inrim.it

EMRP
European Metrology Research Programme
Programme of EURAMET
The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union

This work was jointly funded by the EMRP participating countries within the EURAMET and the European Union. (JRP NEW04-Uncertainty).

Abstract

The measurand value, the conclusions, and the decisions inferred from measurements may depend on the models used to explain and to analyze the results. In this paper, the problems of identifying the most appropriate model and of assessing the model contribution to the uncertainty are formulated and solved in terms of Bayesian model selection and model averaging. The computational cost of this approach increases with the dimensionality of the problem. Therefore, a numerical strategy to integrate over the nuisance parameters and to compute and to sample the measurand post-data distribution is also outlined.

Introduction

According to the prescriptions of the Guide to the Expression of Uncertainty in Measurement [JCfGiM08], data analysis is carried out by selecting a data model and by proceeding as if it had actually generated the data.
This approach ignores the model uncertainty and can lead to underestimates of the uncertainty, to overconfident inferences, and to decisions that are more risky than one thinks they are. Model uncertainty arises in data analysis – outlier identification, inconsistent-data analysis, reference value determinations in the comparison of measurement capabilities, inverse problems, nonparametric function estimation, only to mention a few. Unanswered questions are: How accurately does a model explain the data and what is the impact of the model uncertainty on the measurand estimate and the inferences that we draw from the measurement results? Given an uncertain data model and a measurand estimate based on it, can the total uncertainty of the measurand value be assessed and controlled?
Advances in computing technology have allowed for the consideration of Bayesian methods as ways to incorporate the model uncertainty into the data analysis and the uncertainty budget. Accordingly, we can consider different models, each indexed by one or more parameter, where the Bayesian model selection and model averaging provides the probabilistic framework to simultaneously treat both the model and data uncertainties. Eventually, computational advances foreshadow opportunities for the analysis of the model space, but the impact of these methods on metrology is still to be fully exploited.
This paper lays out the essential ideas of this approach, illustrates some key aspects of its application to metrology problems, and highlights some of the major developments.

Bayesian Data Analysis

In order to explain a set of measurement data $\boldsymbol{x}$, we can consider a number of possible models – say, $A, B, \dots M, \dots$ – where each model is parameterized by the measurand Θ and a set of nuisance parameters Θ_M. Given $\boldsymbol{x}$ and the likelihood $L(\theta, \theta_M, M|\boldsymbol{x}) = P(\boldsymbol{x}|\theta, \theta_M, M)$, the Bayesian approach to the data analysis proceeds by assigning a prior probability distribution $\pi(\theta, \theta_M|M)$ to the parameters of each model and a prior probability $\Pi(M)$ to each model. Next, by using the product rule of the probability calculus, we can write the joint distribution of the data, parameters, and models,

$$P(\boldsymbol{x}, \theta, \theta_M, M) = L(\theta, \theta_M, M|\boldsymbol{x})\pi(\theta, \theta_M|M)\Pi(M)\,. \qquad (1)$$

Through conditioning and marginalization, the joint distribution $P(\boldsymbol{x}, \theta, \theta_M, M)$ can be used to obtain the post-data distributions of interest. For instance, by conditioning (1) on the $\boldsymbol{x}$ and M, and introducing the notation $Z(\boldsymbol{x}|M) = P(\boldsymbol{x}|M)$, the post-data probability distribution of the parameters is

$$P(\theta, \theta_M|\boldsymbol{x}, M) = \frac{L(\theta, \theta_M, M|\boldsymbol{x})\pi(\theta, \theta_M|M)}{Z(\boldsymbol{x}|M)}, \qquad (2)$$

where the the normalizing factor

$$Z(\boldsymbol{x}|M) = \int_{\Theta\oplus\Theta_M} L(\theta, \theta_M, M|\boldsymbol{x})\pi(\theta, \theta_M|M)\,\mathrm{d}\theta\,\mathrm{d}\theta_M \qquad (3)$$

is the evidence of the data given the model M and the integration is carried out over the parameter space $\Theta \oplus \Theta_M$.
The marginalization of (1) over the model parameter and the conditioning on the data yields the updated probability for the model M provided by the data,

$$\mathrm{Prob}(M|\boldsymbol{x}) = \frac{Z(\boldsymbol{x}|M)\Pi(M)}{\sum_M Z(\boldsymbol{x}|M)\Pi(M)}, \qquad (4)$$

where the notation Prob has been introduced in order to indicate discrete probability distributions.
According to (1), firstly, the model M is sampled from $\Pi(M)$; then, the model parameters θ, θ_M are sampled from $\pi(\theta, \theta_M|M)$; eventually, the data $\boldsymbol{x}$ are sampled from $L(\theta, \theta_M, M|\boldsymbol{x})$. Within the framework of this explanation of the data, $\mathrm{Prob}(M|\boldsymbol{x})$ is the conditional probability that M was the model sampled in the first step.
The pre-data distribution $\pi(\theta, \theta_M|M)$ and the prior probabilities $\Pi(M)$ synthesize the model uncertainty before the measurement is carried out; subsequently, the updated probabilities $\mathrm{Prob}(M|\boldsymbol{x})$ provides the uncertainty after the data $\boldsymbol{x}$ have been observed. The simplest way to select a model is to choose the most probable. However, when no single model stands out, the expression of the uncertainty may require to report a set of models along with their probabilities.

Ellipsoidal Nested Sampling

The evaluation of the integral (3) becomes impractical as soon as the parameter space has more than very few dimensions. Among the algorithms for carrying out these integrations numerically, we are investigating a nested sampling technique relating the likelihood values to the prior volume [Ski04]-[Ski06].
Firstly, p likelihood samples $L_{\theta_1}, \dots, L_{\theta_p}$ are sampled in $\Theta \oplus \Theta_M$ according to $\pi(\theta, \theta_M|M)$. Next, the smallest, indicated as L_1, is removed and replaced by a new sample, L_{new}, subject to the constraint $L_{\text{new}} > L_1$. The prior volume enclosed by the iso-likelihood surface $L = L_1$ is estimated as $p/(p + 1)$, where $p/(p + 1)$ is the mean value of the largest of p uniform samples in $[0, 1]$, the total volume of $\Theta \oplus \Theta_M$ being normalised to 1. The discharge of the lowest likelihood L_n, sampling of a replacement constrained to $L_{\text{new}} > L_n$, and shrinking of the prior volume of the associated iso-likelihood surface to $V_n = p^n/(p + 1)^n$ are repeated until the contribution to (3) of the surviving likelihood samples – $L_{\max}p^n/(p + 1)^n$, where $L_{\max}$ is the maximum sample – is less than some pre-defined value.
By using the sequence of the discarded likelihoods $0 < L_1 < L_2 \dots < L_N$, and the differences $\Delta V_n = V_{n-1} - V_n$ of the associated prior volumes $V_0 = 1 > V_1 > V_2 > \dots V_N > 0$, Eq. (3) can be approximated as (rectangle method)

$$Z(\boldsymbol{x}|M) \approx \sum_{n=1}^{N} L_n \Delta V_n + L_{\max}\left(\frac{p}{p+1}\right)^N. \qquad (5)$$

The challenge in implementing (5) is sampling within the iso-likelihood surfaces $L > L_n$. The Monte Carlo Markov Chain algorithm proposed in [Ski04]-[Ski06] may be inefficient. Improvements have been proposed in [MPL06], [SBH07], [FH08], and [FHB09], which we are presently investigating for application to metrology.
Ellipsoidal sampling [MPL06] replaces the iso-likelihood surface $L = L_n$ by a hyper-ellipsoid given by the covariance matrix of the living samples and centered in their mean value, and the $L_{\text{new}} > L_n$ is sampled within the intersection of the domain of integration and this ellipsoidal boundary [AD14].

A realization by *MATHEMATICA* of 100 independent calculations of the integral $\int_{30}^{45}\left(\int_{28}^{40} f(x, y)\,\mathrm{d}y\right)\mathrm{d}x$, where $f(x, y) = \left(-(x - 35)^2 + 9\right)\mathtt{UnitBox}\left(\frac{x-35}{6}\right) \times \left(-(y - 35)^2 + 9\right)\mathtt{UnitBox}\left(\frac{y-35}{6}\right)$, Fig. 1, with 200 sampled points has generated the histogram in Fig. 2, having a mean value $= 1324.9$ (to be compared with the exact value $= 1296$), and a standard deviation $= 131$.

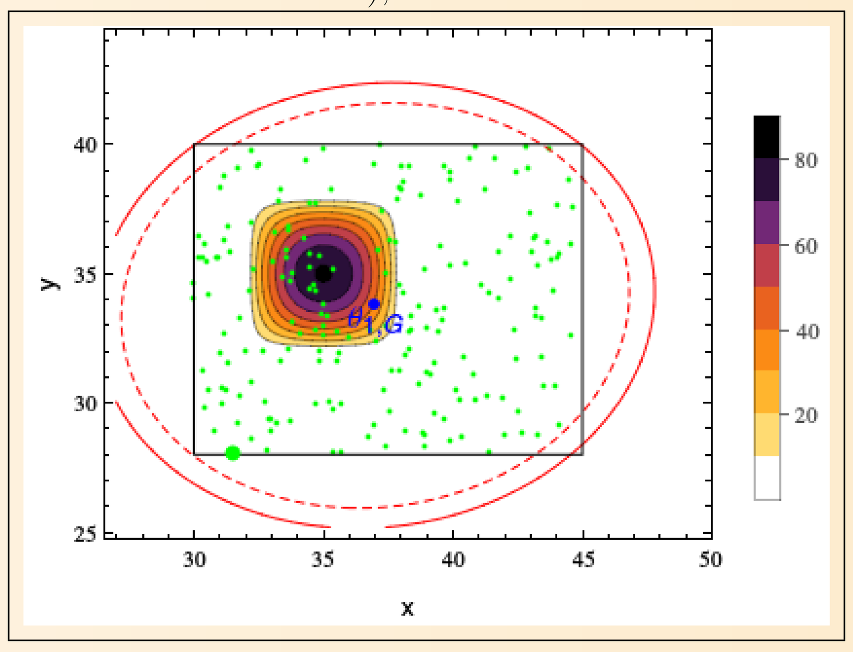

Figure 1: Contour plot of the function $f(x, y)$; are also shown 200 sampled points (green), a couple of ellipsoids (red curves) containing the sampled points, the mean value $\theta_{1,G}$ (blue) and the minimum point (green disk).

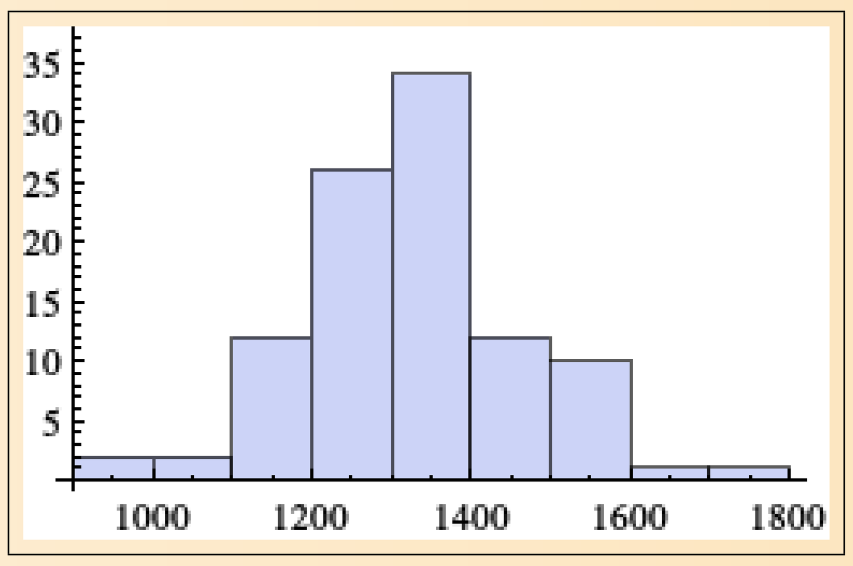

Figure 2: Histogram obtained by 100 independent calculations of $\int_{30}^{45}\left(\int_{28}^{40} f(x, y)\,\mathrm{d}y\right)\mathrm{d}x$; exact value $= 1296$, mean value $= 1324.9$, standard deviation $= 131$.

Example

In order to explain the the "experimental" data shown in Fig. 3, we will consider two different models: "Lorentzian + offset" and "Gaussian + offset". According to these models the mean values n of the counts are

$$n(ch) = \frac{A}{\pi} \times \frac{\Gamma}{(ch - ch_0)^2 + \Gamma^2} + \alpha + \beta \times ch, \qquad (6)$$

$$n(ch) = \frac{A}{\sqrt{2\pi}\,\sigma} \times \mathrm{e}^{-\frac{(ch-ch_0)^2}{2\sigma^2}} + \alpha + \beta \times ch\,. \qquad (7)$$

In both case the single-channel counts are ruled by a Poisson distribution

$$\mathrm{Prob(count)} = \frac{n(ch)^{\mathrm{count}}}{\mathrm{count!}}\mathrm{e}^{-n(ch)}\,. \qquad (8)$$

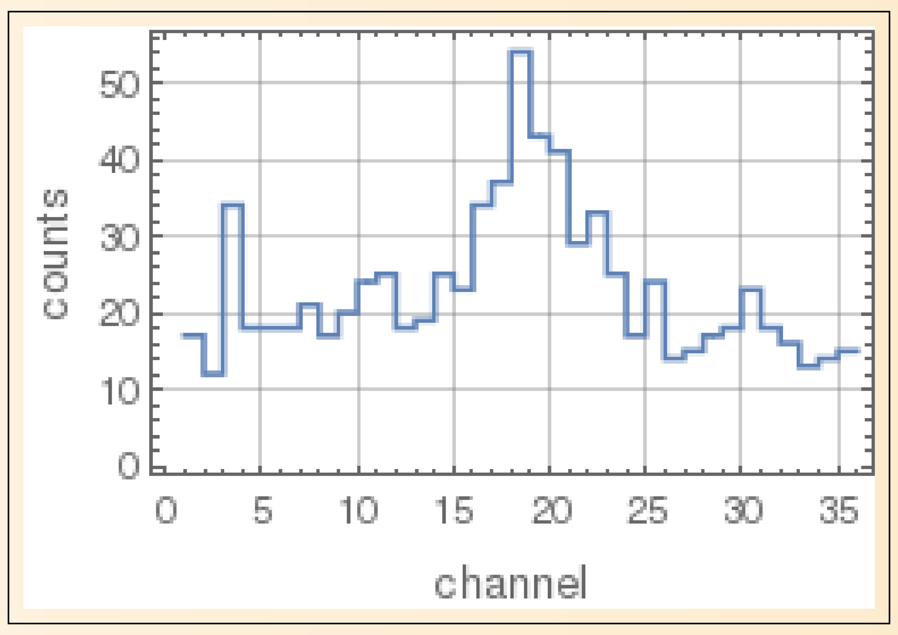

Figure 3: Example of plot "counts *vs* channel".

As in our example there are 36 channels, the likelihhod in Eq. (1) is given by the product of 36 Poisson distributions parameterized by A, Γ or $\sigma, ch_0, \alpha, \beta$ according to Eqs. (6-8). For the parameter A we choose the pre-data distribution $\pi(A) = 1/A$, and for the others 4 independent normal distributions. A realization by *MATHEMATICA* of 200 independent calculations of the 5-dim integral in Eq. (3) for both models has generated the histograms in Figs. 4 and 5.

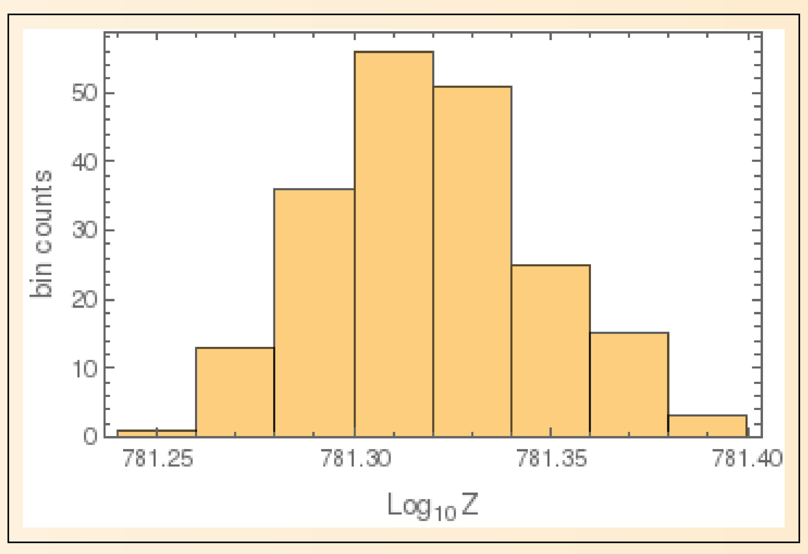

Figure 4: Histogram obtained by 200 independent calculations of the evidence $Z(\boldsymbol{x}|M)$, where M indicates the model "Lorentzian + offset". Mean value $Z = 2.08 \times 10^{781}$, standard deviation $= 1.4 \times 10^{780}$.

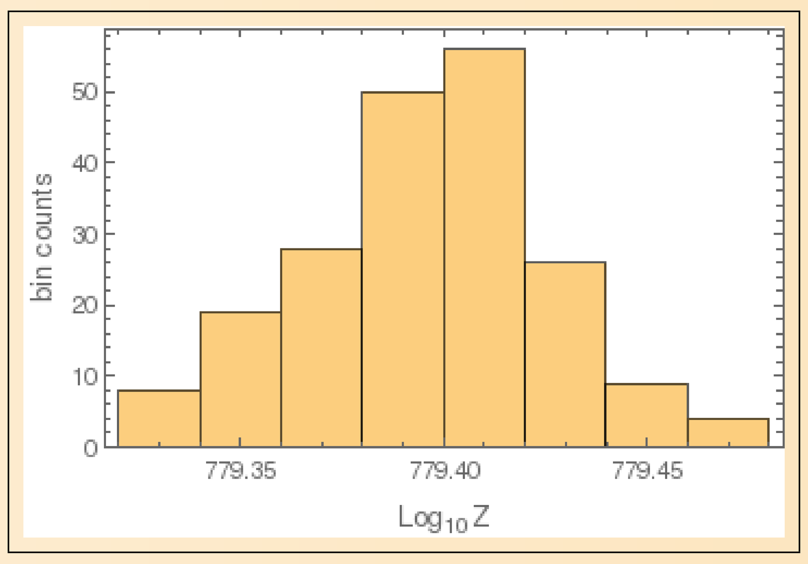

Figure 5: Histogram obtained by 200 independent calculations of the evidence $Z(\boldsymbol{x}|M)$, where M indicates the model "Gaussian + offset". Mean value $Z = 2.50 \times 10^{779}$, standard deviation $= 1.8 \times 10^{778}$.

If we assign the same prior probability $\Pi(M) = 0.5$ to both models, then the application of Eq. (4) reduces to a simple comparison of the huge values for Z calculated by use of the ellipsoidal nested sampling. In this example $Z(\text{Lorentzian + offset}) > Z(\text{Gaussian + offset})$.

Conclusions

The appeal of the Bayesian approach to the expression of the model uncertainty in measurements is the quantification of both the sampling and model contributions to the post-data uncertainty. Whether to synthesize this uncertainty by model selection or model averaging depends on decision theory considerations.

References

[AD14] R. Allison and J. Dunkley. Comparison of sampling techniques for Bayesian parameter estimation. *Monthly Notices of the Royal Astronomical Society*, 437(4):3918–3928, 2014.

[FH08] F. Feroz and M.P. Hobson. Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. *Monthly Notices of the Royal Astronomical Society*, 384(2):449–463, 2008.

[FHB09] F. Feroz, M.P. Hobson, and M. Bridges. MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Not. R. Astron. Soc.*, 398:1601–1614, 2009.

[JCfGiM08] Joint Committee for Guides in Metrology. JCGM 100: Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement. Technical report, BIPM-Sèvres, 2008.

[MPL06] P. Mukherjee, D. Parkinson, and A.R. Liddle. A Nested Sampling Algorithm for Cosmological Model Selection. *The Astrophysical Journal Letters*, 638(2):L51, 2006.

[SBH07] J.R. Shaw, M. Bridges, and M.P. Hobson. Efficient Bayesian inference for multimodal problems in cosmology. *Monthly Notices of the Royal Astronomical Society*, 378(4):1365–1370, 2007.

[Ski04] J. Skilling. Nested Sampling. *AIP Conference Proceedings*, 735(1):395–405, 2004.

[Ski06] J. Skilling. Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1(4):833–860, 2006.