Bayesian inference of a negative quantity from positive measurement results

D. Calonico, F. Levi, L. Lorini and G. Mana
INRIM - Istituto Nazionale di Ricerca Metrologica
Str. delle Cacce 91, 10135 Torino Italy

Abstract **In this paper Bayesian analysis is applied to a Cesium fountain frequency standard to estimate the density shift correction and its uncertainty. Cesium fountain frequency-standards realize the second in the International System of Units with a relative uncertainty approaching 10^{-16}. Among the main contributions to the accuracy budget, cold collisions play an important role because of the so called atomic density shift of the reference atomic transition. The Bayes theorem allows the a priori knowledge of the sign of the collisional correction that is definitively negative to be rigourously embedded into the analysis. In the H-maser frequency measurements using INRIM Cesium fountain as reference, we used the Bayesian analysis and we report a reduction on the final combined uncertainty by more than 25% that demonstrates the Bayesian analysis as a relevant tool in primary frequency-metrology.**

I. INTRODUCTION

Atomic fountains depend on laser cooling of caesium atoms down to a temperature of 1 μK or an even lower value; this, together with the use of a single microwave cavity to implement the Ramsey separated field spectroscopy, allowed the accuracy of primary frequency standards to be improved by more than one order of magnitude.

Today, fountains realize the second with a relative accuracy ranging from 3×10^{-16} to 10×10^{-16} [1-7] and, in the last ten years, allowed the uncertainty of the Atomic International Timescale (TAI) unit to be reduced from 1×10^{-14} to 5×10^{-16}. Since 2003, INRIM operates the Cs fountain ITCs-F1 and evaluates on a regular basis the TAI unit with a type B uncertainty of 5×10^{-16} and an overall uncertainty less than 1×10^{-15} [8]. Since the beginning of 2009 a second fountain, the cryogenic fountain IT-CsF2 is under operation and it is expected to reach an accuracy close to 1×10^{-16} [9]. In Cs fountain clocks, cold collisions occur between the ultra-cold atoms with consequent perturbation of the atomic energy levels and shift of the atomic reference frequency. This collisional shift, is proportional to the atom density and it must be carefully evaluated to correct the clock frequency. Since this shift is a major component of the uncertainty budget, it has been the subject of many theoretical and experimental studies [10, 11]. The density shift depends on the collisional dynamics during the atom ballistic flight, which is related to the way the atom cloud is prepared before launch. If the atomic cloud is prepared with direct molasses capture or molasses expansion after magneto-optical-trap (MOT) capture [12], the collision energy is high during the whole ballistic flight and then the collisional shift has a negligible sensitivity on the atom temperature. In this situation, if the atomic sample is a quantum mixture of the hyperfine eigenstates $|F = 3;m_F = 0>$ and $|F = 4;m_F = 0>$, as occurs in the fountain operation, the collisional shift has also a negligible dependence on the mixture ratio and it is linearly dependent on density through a negative collisional coefficient [10]. On the other hand, when a MOT-captured atom cloud is launched, its high initial density limits the collision energy and the collisional shift becomes strongly dependent on the temperature and the mixture ratio [11].

This paper illustrates the use of the Bayes theorem to infer the collisional correction of Cesium fountain clocks from the results of a sequence of measurements. Bayesian analysis is acquiring increasing importance in metrology; a tutorial guide can be found in Sivia's book [13] and a comprehensive and authoritative account is the treatise by Jaynes [14].

The paper illustrates in a very simple way the differences between the orthodox and Bayesian analyses, differences that are usually unnoticed.

First of all, we used the Bayesian inference [15] to estimate the collisional coefficient, consistently with the constraint of a negative value. This estimate can be used to extrapolate the fountain frequency to the zero-density value. In the present paper we illustrate how to perform the extrapolation, given two or more frequency measurements at different atomic densities, irrespectively of the value of the collisional coefficient value, but consistently with its negative sign. From a mathematical viewpoint, the problem is to find the best-fit line through two or more points, with the constraint that the regression coefficient is negative. The Bayesian approach is important because the collisional shift has the same magnitude as the measurement noise. Therefore, despite the fact that the shift is a linear function of the atom density having a negative regression coefficient, measurement results having the low-density frequency lower than the high-density one are relatively common. A linear extrapolation to zero density from these data is clearly meaningless. On the contrary, a Bayesian inference makes it possible to deal with this situation, thus avoiding physical absurdity and, consequently, reducing the extrapolation uncertainty.

978-1-4244-3510-4/09/$25.00 ©2009 IEEE

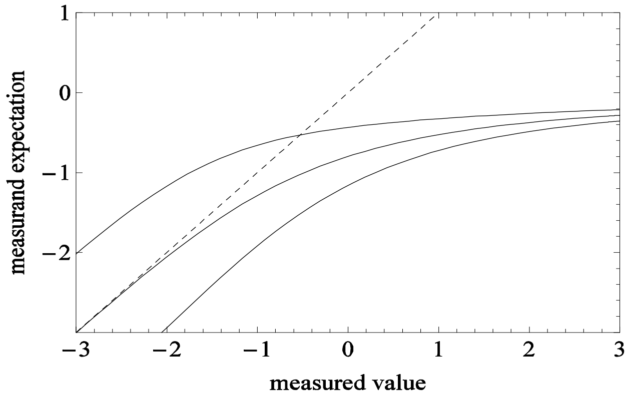

Figure 1. *Expected value of a negative quantity given a measurement value sampled from an unbiased Gaussian distribution with unit variance. Lower and upper lines indicate the plus/minus one standard deviation interval. The dashed line is the orthodox estimate of the*

II. BAYESIAN INFERENCE OF A NEGATIVE QUANTITY FROM POSITIVE MEASUREMENT RESULTS

When data are used to estimate the measurand, the orthodox analysis concerns the probability density of the estimate [16]. Instead, the Bayesian approach concerns the probability density of the measurand, given the single data set generated by measurement and prior possible information. If the data are unbiased normal variables, in the absence of prior information, the same Gaussian function describing the estimate distribution is also the probability density of the measurand. Since this function is invariant when the estimate is exchanged for the measurand, the orthodox and the Bayesian inferences, though conceptually different, are numerically the same. This explains the limited awareness of the risk of using the solution of one problem as the solution of the other.

Had the data been sampled from a distribution not having the above symmetry, the difference between the probability densities of the estimate and of the measurand could not be unnoticed. An interesting case is that of a positive measurement result when the measurand is negative. Though the measurement result is an unbiased estimate of the measurand, this inference looks quite strange. As exemplified in figure 1, if the measurand approaches zero when compared with the measurement uncertainty, there is nothing unusual about obtaining a positive value. Assuming a Gaussian sampling distribution, there is almost a 50%probability that the measurement result is positive and about a 16% probability that the 68% classical confidence interval (the result plus/minus the standard uncertainty) lies in the non-physical region. Though the confidence interval is instrumental only in the assessment of the estimate and not of the measurand [17], this view is not shared by most metrologists. They generally consider a Gaussian function centred on the estimated value and discuss whether the measurand lies within the confidence interval. In the above situation, the consequent paradoxical confidence interval in the non-physical region is explained by observing that, according to a seminal 1937 paper by Neyman [17], the confidence interval is so defined as to keep strictly to a statement about the probability density of the measurement result without considering the probability density of the measurand. According to this statement, independently of the measurand value, only 68% of the intervals obtained from a set of measurements will contain the measurand.

The Bayesian infererence is introduced using a very simple case: the outcome of a single measurement of a quantity which is negative definite. Let be y_1 a possible value of y. If the measurement value y_m is normally distributed about y_0 with standard deviation s the probability to get $y_m = y_1$ is:

$$P_{y_m}(y_m = y_1 \mid y = y_0) = \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[\frac{-(y_1 - y_0)^2}{2\sigma_1^2}\right] \quad (1)$$

Given the following condition on y:

$$P_y(y = y_0) = \vartheta(-y_0) \quad (2)$$

The Bayes theorem prescribes that the post-data probability density of the possible y values, given the measurement result $y_m = y_1$, is the tail of (1) lying in the $y_0 < 0$ interval

$$P_y(y = y_0 \mid y_m = y_1) = \frac{P_{y_m}(y_m = y_1 \mid y = y_0)P_y(y = y_0)}{P_{y_m}(y_m = y_1)} \quad (3)$$

Which, in this specific case becomes

$$P_y(y = y_0 \mid y_m = y_1) = \frac{2\exp\left[-(y_1 - y_0)^2/2\right]\vartheta(-y_0)}{\sqrt{2\pi}\,\mathrm{erfc}\left(y_1/\sqrt{2}\right)} \quad (4)$$

Standard techniques for probability theory are used to find numerical estimations of different parameters. To this purpose the loss function $L(\hat{y} - y)$ is introduced. This function maps onto a number the error associated with the cost of a wrong estimate. Estimation of $\hat{y}$ is:

$$\hat{y} = \arg_{\hat{y}} \min E\left[L(\hat{y} - y)\right] \quad (5)$$

Different expression for the loss function provides different estimators through (5)

$$\begin{aligned} L(\hat{y} - y) &= (\hat{y} - y)^2 \rightarrow mean \\ L(\hat{y} - y) &= |\hat{y} - y| \rightarrow median \end{aligned} \quad (6)$$

Confidence intervals are expressed by integrating (4) to obtain the relevant cumulative distribution function.

III. BAYESIAN INFERENCE OF A REGRESSION LINE

In the usual fountain operation the collisional coefficient is calculated as the slope of a regression line which fits frequency data taken at different operational densities.

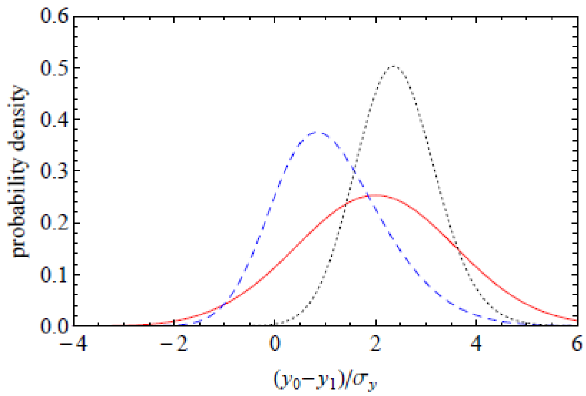

Figure 2. *Probability density of the intercept value; solid (red) is (y2 - y1)=σy = -4, dashed (blue) is (y2 - y1)=σy = 0, dotted (black) is (y2 - y1)=σy =+4.*

Bayesian inference is applied here to the calculation of the regression line.

Given N measures y_i at the density x_i we assume that the data are normally distributed about a regression line with standard deviation only along y. The joint probability density to find y_1.. y_N, given x_1... x_N is

$$P_y(y_1,...y_N \mid ax_i + b, \sigma_i) = \prod_{i=1}^{N} N(y_i \mid ax_i + b, \sigma_i^2) \quad (7)$$

Where N is a normal distribution with mean ax_i+b and standard deviation σ_i.

Applying the Bayes theorem to (7) with the condition that a (slope of the regression line) is negative the following expression is obtained

$$P_{ab}(a_0, y_0 \mid x_i, y_i) \sim \exp\left[-\begin{pmatrix} a_0 - \hat{a} & y_0 - \hat{y}_0 \end{pmatrix} \frac{C_{ab}^{-1}}{2} \begin{pmatrix} a_0 - \hat{a} \\ y_0 - \hat{y}_0 \end{pmatrix}\right] \vartheta(a_0) \quad (8)$$

where $\hat{a}$, $\hat{y}_0$, and C_{ab} are the least-squares estimates and covariance matrix of a and b.

The meaning of expression (8) is clearer in a simple case, where only two measurements *(1,0)* and *(x, y)* are available with $x>1$ and $a<0$. Then (8) simplifies to:

$$P_b(b = y_0 \mid x, y) \sim 2\exp\left[\frac{-(y_0 - \hat{y}_0)^2}{2\sigma_0^2}\right] \operatorname{erfc}\left[\frac{-(y_0 - \tilde{y}_0)^2}{\sqrt{2}\tilde{\sigma}_0}\right] \quad (9)$$

Where $\hat{y}_0 = -y/(x-1)$ and $\sigma_0^2 = (1+x^2)/(x-1)^2$ are the least-squares estimate and thevariance of b (i.e., the intercept of the best-fit line through the data), $\tilde{y}_0 = -xy/(x+1)$, $\sigma_0^2 = (1 + x^2)/(x + 1)^2$

Fig 3. reports the Expected mean value and standard deviation of the zero-density frequency, given $\{x_1=1;\ y_1\}$ and $\{x_2=3;\ y_2\}$. The coloured area indicates the extrapolation uncertainty. The straight lines (asymptotic limits) are the $-(y_2-y_1)/2$ extrapolation and the $(y_1+y_2)/2$ sample mean. Dashed line is the expected frequency when the same {1; y_1} and {3; y_2} pairs is observed three times.

Analytical expressions for y_0 are available in asymptotic cases

$$y << 0 \qquad y_1 - y/(x-1) \qquad y >> 0 \qquad y_1 + y_2 = y \quad (10)$$

If $y>>0$ the expression is similar to the expected one from an orthodox solution. If $y<<0$ the y is the sample mean between y_1 and y_2.

IV. BAYESIAN INFERENCE APPLIED TO FOUNTAIN DATA

Here below the Bayesian inference is applied to real data coming from IT-CsF1 PFS at INRIM. Dataset is representative of a single run among many performed to provide calibration of TAI unit to BIPM. An H-maser is used as a frequency flywheel for differential measurements (and as a transfer oscillator to TAI). Its frequency drift is clear from the picture below, where the dataset is reported with respect the measurement epoch.

Zero-density extrapolation is accomplished alternating high (red) density and low density (blue) sessions. Density leverage is about 4 and the duration of low density session is 2.5 times longer to obtain data with similar uncertainties

At INRIM, we varied the atom density through the loading time of the magneto-optical trap. To evaluate the collisional shift, the fountain is operated alternately at high and low atomic density; the 70 ms and 300 ms loading times provide respectively the low and high density configurations. The ratio between the number of detected atoms in the two configurations ranges from three to four. Since in the low density configuration

stability is poor and resolution is limited by the atom shot noise, the fountain is operated alternating about 21000 s in the low density configuration and about 6000 s in the high density one. The hydrogen maser frequency is then extrapolated to zero density by

$$\hat{y}_0 = y_1 - y/(x-1) \quad (11)$$

where $\hat{y}_0$ is the sought zero density frequency, y_1 and y_2 are the frequency in low and high density conditions, $y = y_2-y_1$, and $x=x_2/x_1$ is the ratio between the number of atoms detected at the high and low densities. The extrapolation is carried out for each $(y_1;\ y_2)$ pairs. The total duration of each run is 27000 s; this ensure a good rejection of systematic effects (fluctuations of the hydrogen maser frequency, magneto-optical trap, and atom detection efficiency) which could bias the frequency extrapolation. The extrapolation uncertainty associated with (3) is where, for the sake of simplicity, we assumed $|y_2-y_1|\sigma_x<<\sigma_y$. The actual experimental practice in the various laboratories differs in the way the atom clouds of different densities are prepared and in the durations of the fountain

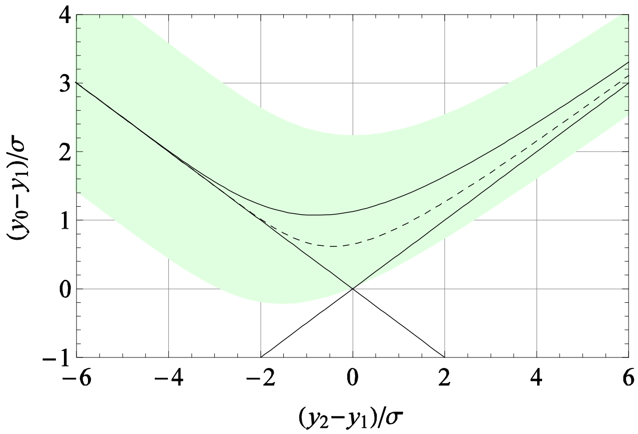

Figure 3. *Expected value and standard deviation of the zero-density frequency, given {x1 = 1; y1} and {x2 = 3; y2}. The coloured area indicates the extrapolation uncertainty. The straight lines (asymptotic limits) are the -(y2 - y1)=2 extrapolation and the (y1 + y2)=2 sample mean. Dashed line is the expected frequency when the same {1; y1} and {3; y2} pairs is observed three times.*

operation at low and high density. The extrapolations techniques of the fountains of the Physikalish Technische Bundesanstalt (PTB-CsF1), of the National Physical Laboratory (NPL-CsF1), and of the National Institute of Standards and Technology (NIST-F1) differ only by the driving parameter and the time scheduling of the differential measurements. At SYRTE, a rapid adiabatic passage technique is used, which ensures very accurate evaluation of the collisional coefficient because the atom density is precisely set at one value and at its half.

H- maser drift complicates the data analysis. To preserve the most general approach, the data are considered normally distributed about the plane $y_i=ax_i+b+ct_i$ where (y_i, x_i, t_i) are the frequency, the density and the epoch defining each measurement and c is the H-maser drift

Then the data are analyzed using both the statistical orthodox and the Bayesian inference. For the orthodox inference, data have been treated with the least-square method and a regression plane $y=\tilde{a}x+\tilde{b}+\tilde{c}t$ is estimated, where $\tilde{a},\tilde{b},\tilde{c}$ are the least-square estimation for a b and c. For the Bayesian inference (9) has been modified as follows

$$P_{abc}(a_0,y_0,\tilde{c}\,|\,x_i,y_i)\sim\prod_{i=1}^{N}N(y_i\,|\,a_0x_i+y_0+\tilde{c}t_i,\sigma_i^2)\times \qquad (12)$$
$$\times N(\tilde{c}\,|\,c_0,\sigma_{c_0})\vartheta(a_0)$$

And the expectation values for a y_0 and c have been calculated. It is worth noting that a pre-data estimation for the maser drift, can be used (the maser drift is very stable on long times) and that y_0 is the frequency value extrapolated at zero density and at the epoch zero (the mid-point of the evaluation period)

In Table 1 a comparison between orthodox and Bayesian inferences is proposed. Data are normalized with respect to σ_υ = 3.9 x10^{-15} (the standard deviation of a single measurement) When extrapolating the frequency of a caesium-fountain

TABLE I. COMPARISON:ORTHODOX AND BAYESIAN DATA ANALYSIS

	Classical Analysis	***Bayesian Analysis***
H-maser drift	(0.43±0.03)σ_υ/day	(0.44±0.02) σ_υ/day
Collisional coefficient	(-0.14±0.10)$\sigma_\upsilon/\rho_{low}$	(-0.18±0.08) $\sigma_\upsilon/\rho_{low}$
Zero density freq	(0.18±0.25) σ_υ	(0.24±0.18) σ_υ

clock to zero atom-density, the Bayes theorem makes it possible to take account of a negative correlation between the density and the frequency at the very beginning of data analysis, in a way ensuring consistency of the extrapolation. In the specific example here considered we have considered the H-maser evaluation by INRIM Cs fountain during a standard operation; type A uncertainty contribution (H-maser + atomic density shift) is reduced by 29%, from 9,8 x10^{-16} to 7,0 x10^{-16}, demonstrating the Baysian analysis as a useful technique for frequency metrology.

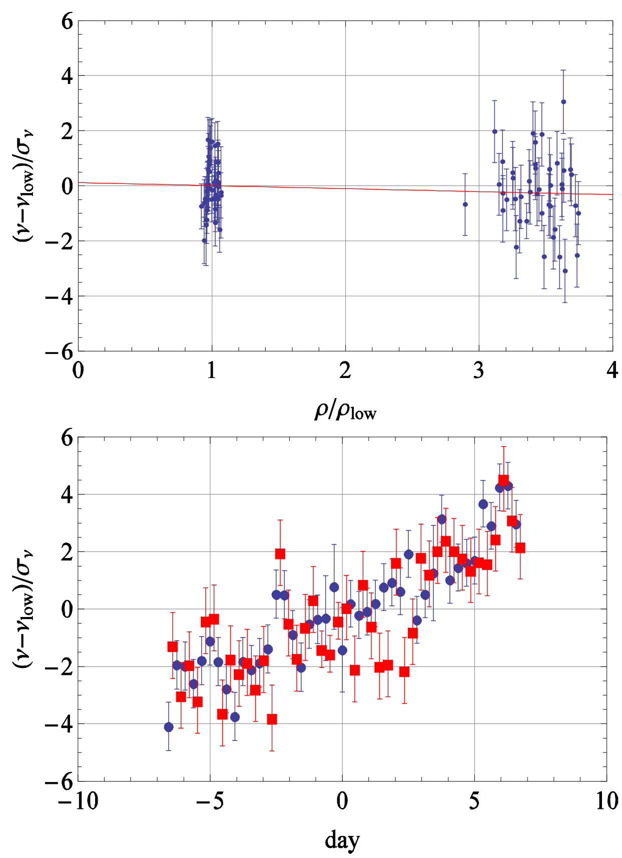

Figure 4. *Results of low (blue dots) and high density (red squares) frequency measurements. Density values scaled by the mean low density, ρlow.Frequency values shifted by the mean low-density frequency, νlow, and scaled by the mean measurement uncertainty σν. Top: data vs measurement time; Bottom: data vs density; the line is the intersection of the best-fit plane with t = 0.*

REFERENCES

[1] F. Levi, D. Calonico, L. Lorini and A. Godone, "IEN-CsF1 primary frequency standard at INRIM: accuracy evaluation and TAI calibrations", Metrologia, 43, pp. 545–55, 2006

[2] T. P. Heavner, S.R. Jefferts, E.A. Donley, J.H. Shirley, T.E. Parker "NIST-F1: recent improvements and accuracy evaluations", Metrologia, 42, pp. 411-422, 2005

[3] C. Vian, P. Rosenbusch, H. Marion, et al. "BNM-SYRTE fountains: Recent results", IEEE Trans on Instr and Meas, 54, pp. 833-836, 2005

[4] S. Weyers, U. Hubner, R. Schroder, et al., "Uncertainty evaluation of the atomic caesium fountain CSF1 of the PTB", Metrologia, 38, pp. 343-352, 2001

[5] K. Szymaniec, W. Chalupczak,PB Whibberley, S. N. Lea, D. Henderson, "Evaluation of the primary frequency standard NPL-CsF1", Metrologia, 42, pp. 49-57, 2005

[6] M. Kumagai, H. Ito , M. Kajita, and M. Hosokawa, "Evaluation of caesium atomic fountain NICT-CsF1", Metrologia, 45, pp. 139-148, 2008

[7] K. Takayuki, Y. Fukuyama, Y. Koga, K. Abe, "Preliminary evaluation of the Cs atomic fountain frequency standard at NMIJ/AIST", IEEE Trans. IM, 53, pp. 466-471, 2004

[8] A Bauch et al., "Comparison between frequency standards in Europe and the USA at the 10-15 uncertainty level", Metrologia, 43, pp. 109-120, 2006

[9] F. Levi at al. "The Cryogenic Fountain ITCsF2", these proceedings

[10] P J Leo, P S Julienne, F H Mies and C J Williams, "Collisional frequency shifts in 133Cs fountain clocks", Phys. Rev. Lett., 86, pp. 3743–6, 2001

[11] K Szymaniec et al., "Cancellation of thecollisional frequency shift in caesium fountain clocks", Phys.Rev. Lett., 98, 153002, 2007

[12] F Levi, L Lorini, D Calonico, A Godone, "Systematic shift uncertainty evaluation of IEN CSF1 primary frequency standard", IEEE Trans. on Instr. and Meas., 52, pp. 267-271, 2003

[13] D S Sivia and J Skilling "Data Analysis: a Bayesian Tutorial", Oxford, Oxford University Press, 2007

[14] E T Jaynes, "Probability Theory: the Logic of Science", Cambridge, Cambridge University Press, 2003

[15] D Calonico, F Levi, L Lorini and G Mana "Bayesian inference of a negative quantity from positive measurement results", Metrologia, 46, pp. 267, 2008

[16] J Neyman, "Outline of a theory of statistical estimation based on the classical theory of probability", Phil. Trans. R. Soc. Lond. A, 236, pp. 333–80, 1937

[17] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 1995 Guide to the Expression of Uncertainty in Measurement, Geneva, Swizerland: International Organization for Standardization, Metrologia, 46, pp. 267–271, 2009

[18] R T Cox "Probability, Frequency, and Reasonable Expectation" Am. Jour. Phys., 14, pp.113, 1946