Nuclear Instruments and Methods in Physics Research A 604 (2009) 701–706

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Parametric analysis of discrepant sets of data

O. Helene

Instituto de Física, Universidade de São Paulo, CP 66318, CEP 05314-970, São Paulo, SP, Brazil

ARTICLE INFO

Article history:
Received 2 December 2008
Received in revised form
5 March 2009
Accepted 9 March 2009
Available online 18 March 2009

Keywords:
Data analysis
Discrepant data
Half-lives

ABSTRACT

In this paper a new parametric method to deal with discrepant experimental results is developed. The method is based on the fit of a probability density function to the data. This paper also compares the characteristics of different methods used to deduce recommended values and uncertainties from a discrepant set of experimental data. The methods are applied to the ^{137}Cs and ^{90}Sr published half-lives and special emphasis is given to the deduced confidence intervals. The obtained results are analyzed considering two fundamental properties expected from an experimental result: the probability content of confidence intervals and the statistical consistency between different recommended values. The recommended values and uncertainties for the ^{137}Cs and ^{90}Sr half-lives are 10,984 (24) days and 10,523 (70) days, respectively.

1. Introduction

We expect to obtain from a measurement an estimate, x, of a quantity's true value, x_0, and its uncertainty, σ_x. Unless otherwise stated, the true value of the measured quantity, x_0, lies within the interval $x\pm\sigma_x$ with a probability of about 68%:

$$P(x_0 \in x \pm \sigma_x) \approx 68\%. \quad (1)$$

If the probability function is Gaussian and the uncertainty σ_x is known exactly, then the approximation becomes exact. This fundamental property of x_0, x, and σ_x is used for testing hypotheses and theories and planning new experiments. An experimental result and its uncertainty must also satisfy the fundamental property given by Eq. (1) regardless of whether they are used in fundamental or applied physics (as nuclear medicine, environmental protection, safeguards, etc.), or in weighted averages, or in error propagations. Thus, underestimating or overestimating the confidence interval σ_x is equally undesirable; in the first case we could be rejecting valid theories and hypotheses; in the second case we may be accepting poor and unrealistic ones. Also, incorrect decisions could be taken in cases where a wrong uncertainty is propagated or is used in weighted averages.

However, often in experimental physics it is very difficult to deduce a good estimate of a standard deviation due to undetected or unknown uncertainties. Thus, we often encounter published data $x_i\pm\sigma_i$, $i=1, 2, \ldots, m$, which correspond to various uncorrelated measurements of the same quantity such that

$$|x_i - x_j| \geqslant \sqrt{\sigma_i^2+\sigma_j^2}, \quad j\neq i \quad (2)$$

for several pairs i, j, suggesting that the data are discrepant. In such cases we should not apply the most common statistical methods, such as the weighted average; instead we should use special procedures to deduce a recommended value and its uncertainty that satisfy Eq. (1).

Many procedures have been proposed to deal with discrepant data sets. In this paper we will analyze the recommended values and uncertainties deduced using some of them, and propose a new method to calculate recommended values and their uncertainties based on the fit of a probability density function (pdf) to the published data.

2. Comparing various methods: consistency condition

MacMahon et al. [1] applied various procedures to deduce recommended values for two typical discrepant sets of data: the half-lives of ^{137}Cs and ^{90}Sr. The main goal of Ref. [1] was to study the change of the recommended values as the size of the set of data grew over time.

In this section we extend the work of Ref. [1] and study the change of the recommended values and their uncertainties as the number of data in the data set increases. The published data considered in this paper, reproduced in Tables 1 and 2, are exactly the same as those used in Ref. [1] so that both studies can be compared.

The question is whether such recommended values and their uncertainties do satisfy the condition given by Eq. (1). To study this, let us consider $x_i\pm\sigma_i$, $i=1, 2, \ldots, m$, a (discrepant) set of data

E-mail address: otaviano@if.usp.br

0168-9002/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2009.03.021

Table 1
Published ^{137}Cs half-life data (and uncertainties), in days, taken from Ref. [1].

9715 (146)	10,665 (110)	11,023 (37)	10,967.8 (4.5)
10,957 (146)	11,220 (47)	11,020.8 (4.1)	10,940.8 (6.9)
11,103 (146)	10,921 (183)	11,034 (29)	11,018.3 (9.5)
10,994 (256)	11,268 (256)	10,906 (33)	10,970 (20)
10,840 (18)	11,191 (157)	11,009 (11)	

Table 2
Published ^{90}Sr half-life (days) data from Ref. [1].

10,120 (150)	10,410 (330)	10,588 (91)	10,495 (4)
10,700 (580)	10,636 (88)	10,665 (37)	10,557 (11)
10,230 (150)	10,282 (12)	10,561 (14)	

ordered by publication date $t_i < t_{i+1}$, and the recommended values $x_{rec\,j} \pm \sigma_{rec\,j}$ also ordered in the same way. That is, the jth recommended value and its uncertainty were deduced taking into account the published data $i = 1, 2, \ldots, j$.

The recommended values $x_{rec\,j}$ must agree within their quoted uncertainties, $\sigma_{rec\,j}$. This is a necessary condition for the recommended values to obey the fundamental property given by Eq. (1). To study the agreement between two recommended values $x_{rec\,i} \pm \sigma_{rec\,i}$ and $x_{rec\,j} \pm \sigma_{rec\,j}$ we should consider that they are correlated, since they were deduced using common published results. If the published data were normally distributed, had the same standard deviation, and orthodox statistical procedures were used to deduce the recommended values, then the correlation coefficient between $x_{rec\,i}$ and $x_{rec\,j}$, $i<j$, is given by $\rho_{ij} = \sqrt{i/j}$. Thus, two recommended values agree when

$$|x_{rec\ i} - x_{rec\ j}| \sim\gg \sqrt{\sigma^2_{rec\ i} + \sigma^2_{rec\ j} - 2\sigma_{rec\ i}\sigma_{rec\ j}\sqrt{i/j}}, \quad i<j \tag{3}$$

where $\sim\gg$ means "not much larger than". Thus, a figure of merit to measure the agreement between two recommended values could be

$$Q^2_{ij} = \frac{(x_{rec\ i} - x_{rec\ j})^2}{\sigma^2_{rec\ i} + \sigma^2_{rec\ j} - 2\sigma_{rec\ i}\sigma_{rec\ j}\sqrt{i/j}} \tag{4}$$

which is a chi square (χ^2)-like statistic with one degree of freedom (the expected values of χ^2 with one degree of freedom is 1).

In what follows, whether some proposed methods developed to deal with discrepant sets of data satisfy the consistency condition given by Eq. (3) will be investigated. The recommended values and their uncertainties for the ^{137}Cs and ^{90}Sr half-lives taken from Ref. [1] are shown in Figs. 1 and 2, respectively. The recommended values are displayed in chronological order obtained from the first m published data indicated on the abscissas, $m \geqslant 5$. (Ref. [1] shows also the results where fewer data points were used for deducing recommended values. However, since two parameters are deduced from the published data, i.e., the recommended value and its uncertainty, the analysis has been restricted to cases for which at least five published data points have been included. However, the conclusions given below would not change had the analysis considered a smaller number of experimental results.) The procedures used in Ref. [1] to deduce recommended values and their uncertainties are the "Limitation of Relative Statistical Weights (LRSW)" [2], the "Normalized Residuals (NR)" [3], the "Rajeval method (Rajeval)" [4], the "Median method (Md)", with the standard deviation calculated as recommended in Ref. [5], the "Bootstrap method (Boot)" [6,7], and the "Extended Bootstrap method (ExtB)" [8]. (Ref. [1] gives a concise description of these methods.) Figs. (1) and (2) also show recommended values obtained with the "Maximum Likelihood (ML) Method", developed in this work and presented in the next section.

As can be seen from Figs. 1 and 2, the NR and the Rajeval methods clearly underestimate the uncertainties and do not satisfy the consistency condition (Eq. (3)). Consequently, the fundamental property established in Eq. (1) cannot be satisfied by the recommended values produced by these methods.

The second and the last recommended values given by the ExtB method both in the case of ^{137}Cs and ^{90}Sr half-lives hardly agree at all. Their figures of merit given by Eq. (4) are 5.7 and 6.0, respectively.

The figure of merit obtained when comparing the less consistent recommended values in the case of the LRSW method (the second and the last values, respectively, in both cases) are 2.9 and 3.2. These values are not too large compared to the expected value 1.0; however, they can be an indication that the standard deviations are underestimated.

The standard deviation of the median, as proposed by Müller [5], is calculated assuming a Gaussian distribution for the data, which is not always true in the case of discrepant data. The figure of merit obtained when comparing the second and the last recommended half-life values for ^{90}Sr is very large ($Q^2 = 12$). Thus Md possibly underestimates the standard deviation of the recommended value in some cases.

The recommended values calculated using BM obey the consistency condition in both the cases of ^{137}Cs and ^{90}Sr half-lives. Also, when applied to the $\omega(782)$ full-width, and the K^0_S and the neutron half-lives, the BM gives consistent estimates [7].

3. Fitting a probability density function to a discrepant data set

In this section, we present a new parametric method to deduce a recommended value and its uncertainty from a discrepant set of data. The pdf

$$f(x) = N\frac{1}{\left(1 + (x - x_0)^2/a_0 n_0\right)^{(n_0+1)/2}} \tag{5}$$

where N is the normalization factor given by

$$N = \frac{\Gamma((n_0+1)/2)}{\sqrt{n_0 a_0 \pi}\,\Gamma(n_0/2)} \tag{6}$$

will be fitted to the discrepant sets of data analyzed above. The parameters to be fitted are the true value of the measured quantity (x_0), a scale parameter related to the data spread (a_0), and a parameter related to the shape of the function (n_0). Thus, by fitting all three parameters to a discrepant set of data we obtain the recommended value and its uncertainty and also the pdf of the data.

The function in Eq. (5) is a *Student-t*-like distribution that has some interesting properties. When $n_0 \leqslant 2$, the standard deviation of x diverges, as in the case of a Breit–Wigner or the Lorentzian ($n_0 = 1$) distributions. (The justification to consider a pdf with an infinite standard deviation when dealing with experimental data is given in the Appendix.) When n_0 increases, the pdf of Eq. (5) tends to a Gaussian distribution.

The parameters x_0, a_0, and n_0 were fitted using the Maximum Likelihood estimation method. This method estimates true values of the parameters by maximizing the likelihood function, which is the joint pdf of the experimental observation x_i, $i = 1, 2, \ldots, m$:

$$L(x_0, a_0, n_0) = \prod_{i=1}^{m} f(x_i). \tag{7}$$