International Atomic Energy Agency

INDC(CCP)-343
Distr.: L

INDC INTERNATIONAL NUCLEAR DATA COMMITTEE

METHOD FOR EVALUATING NON-CONSISTENT DATA USING TWO STATISTICAL CRITERIA

Yu.F. Yaborov

(Translation of a paper published in Yadernye Konstanty (Nuclear Constants 2, 1989) - Original Report in Russian was distributed as INDC(CCP)-303/G)

Translated by the IAEA

November 1991

IAEA NUCLEAR DATA SECTION, WAGRAMERSTRASSE 5, A-1400 VIENNA

INDC(CCP)-343
Distr.: L

METHOD FOR EVALUATING NON-CONSISTENT DATA USING TWO STATISTICAL CRITERIA

Yu.F. Yaborov

(Translation of a paper published in Yadernye Konstanty (Nuclear Constants 2, 1989) - Original Report in Russian was distributed as INDC(CCP)-303/G)

Translated by the IAEA

November 1991

Reproduced by the IAEA in Austria
November 1991

91-05708

METHOD FOR EVALUATING NON-CONSISTENT DATA
USING TWO STATISTICAL CRITERIA

Yu. F. Yaborov

In certain data evaluation tasks we have to evaluate a series of n measurements of unequal accuracy represented by the values $\{x_i\}$ and the corresponding error levels $\{\sigma_i\}$ ($i = 1, \dots n$). The weighted mean $\bar{x}_\sigma$ [1] is an unbiased efficient estimator for $\{x_i\}$

$$\bar{x}_\sigma = \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \Big/ \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \, . \tag{1}$$

Using the internal and external consistency criteria proposed by Birge [2], two types of error level may be calculated: the internal error level σ_I and the external error level σ_E:

$$\sigma_I = \left[\sum_{i=1}^{n} 1/\sigma_i^2 \right]^{-\frac{1}{2}}, \tag{2}$$

$$\sigma_E = \frac{\left[\sum_{i=1}^{n} (x_i - \bar{x}_\sigma)^2 / \sigma_i^2 \right]^{\frac{1}{2}}}{\left[(n-1) \cdot \sum_{i=1}^{n} 1/\sigma_i^2 \right]^{\frac{1}{2}}} \, . \tag{3}$$

The error level for the mean-weighted value $\bar{x}_\sigma$ will be σ and it can be determined as follows [3]:

$$\sigma = \begin{cases} t_{n-1};\ \alpha \cdot \sigma_E\,, & \text{if} \quad \sigma_E > \sigma_I \\ t_{n-1};\ \alpha \cdot \frac{1}{2}\,(\sigma_E + \sigma_I), & \text{if} \quad \sigma_E \leqslant \sigma_I\,, \end{cases} \tag{4}$$

where $t_{n-1}; \alpha$ are the Student coefficients for the significance level α given, for example, in Ref. [4]. The data to be processed are consistent if

$$\frac{(n-1) \cdot \sigma_E^2}{\sigma_I^2} = \sum_{i=1}^{n} \frac{(x_i - \bar{x}_\sigma)^2}{\sigma_i^2} = \chi^2_{calc}(n) \leqslant \chi^2_\alpha(n-1), \tag{5}$$

where $\kappa^2_{calc}(n)$ is the calculated value and $\kappa^2_{\alpha}(n-1)$ is the tabular value for κ^2 which is the distribution for the significance level α.

If the data are consistent, there is no difficulty in processing them to obtain the required evaluations. In practice, we are often dealing with systems of non-consistent data for which no single processing algorithm can be used. This is a pressing problem and it is this which we examine in this paper.

Non-consistency of data can be caused by the presence of k mismatched data with underestimated error levels $\{\sigma_i\}$ (j=1, ... k), and/or by values $\{x_j\}$ which differ sharply from the values for the majority of the data $\{x_i\}$. In other words, the majority (n-k) of the data which belong to the general set $N(\mu, \xi^2)$ are augmented by the k data ($k < n/2$) which belong to the sets $N(\mu+\lambda, \xi^2)$ or $N(\mu, \kappa^2 \cdot \xi^2)$ [5].

This paper looks at methods of identifying mismatched data and finding possible causes of their being mismatched by using a statistical criterion which highlights abnormal extreme values (AEV), and it suggests processing methods for systems of non-consistent data.

Reference [6] suggests that, by successively eliminating each k-th result, the test statistic κ^2 for the remaining (n-1) data could be calculated. If the condition of consistency is fulfilled with no k-th result:

$$æ^2_{calc}(n-1) = \sum_{\substack{i=1 \\ i \neq k}}^{n} (x_i - \bar{x}_{\sigma},k)^2/\sigma_i^2 \leqslant æ^2_{\alpha}(n-2), \tag{6}$$

where $\bar{x}_{\sigma,k}$ is the weighted mean without the k-th result. In this instance, it is precisely the k-th result which is mismatched. In Ref. [7], those k results are considered to be mismatched which, upon successive individual elimination, produced a match among the (n-1) remaining results in accordance with expression (6).

We propose the successive and simultaneous exclusion of 2, 3 and, generally, k data ($k<n/2$) with a view to obtaining a matched system from the

(n-k) data for the lowest possible value of k, i.e. to find those data the exclusion of which results in the fulfilment of the following condition

$$\varkappa^2_{calc}(n-k) = \sum_{i=1}^{n-k} \frac{(x_i - \bar{x}_{\sigma}, n-k)^2}{\sigma_i^2} \leq \varkappa^2_{\alpha}(n-k-1), \tag{7}$$

where

$$\bar{x}_{\sigma}, n-k = \sum_{i=1}^{n-k} \frac{x_i}{\sigma_i^2} \Bigg/ \sum_{i=1}^{n-k} \frac{1}{\sigma_i^2}$$

Reference [8] points out that the consistency criterion κ^2 must be complemented by other criteria since it is an approximation.

We subject the system of k mismatched data to the following analysis using a parametric criterion for identifying AEVs. A variational series is constructed from the values $\{x_i\}$: $x_1 \leq x_2 \cdots x_{n-1} \leq x_n$. If there are mismatched data within the variational series, the mismatch is due to underestimation of the error level. If the mismatched data are minimum or maximum values, the Rosner procedure [9] for identifying AEVs may be applied to these extreme data $\{x_j\}$; the advantages and superiority of this procedure are demonstrated in Ref. [10]. If some of the mismatched data from $\{x_j\}$ are AEVs, according to the Rosner procedure, for the level of significance adopted, i.e. they belong to the other general set where $\mu' = \mu + \lambda$, they are excluded from subsequent processing. The remaining $l(1 \leq k < n/2)$ unexcluded mismatched data are included in the subsequent processing. We then extend the error level $\{\sigma_l\}$ $(1 \leq k < n/2)$ of these l data in order to achieve the fitting condition given in expression (5). Reference [11] obtains matched data from a mismatched system by rejecting some data and extending the error levels of other data. However, objective uniqueness was not achieved since, out of the 22 variants considered, the one consistent system was chosen on an arbitrary and subjective basis. Reference [12] puts forward a formalized algorithm for extending all n error levels of a system of non-consistent data without

$T_{1/2}$ ^{239}Pu data processing results

Method / Value	Evaluation of source data $n = 8$	Evaluation from Ref. [15] minus (X_1, σ_1) $n = 7$	Evaluation minus (X_1, σ_1) $n = 7$	Evaluation minus (X_1, σ_1) and (X_8, σ_8) $n = 6$	Extension of all σ values by 1.69 times as per Ref. [7] $n = 8$	Extension of σ_1 and σ_8 only by 2.17 times $n = 8$	Minimum σ_E for σ_1=679 σ_8=18.3 $n = 8$	Maximum σ_E for $\sigma_1 = \sigma_8 = 40,2$ $n = 8$
$\overline{X}$	24106,2	24119	24118,7	24111,1	24106,2	24106,2	24106,2	24106,2
$\overline{X}_\sigma$	24114,4		24120,6	24113,0	24114,4	24113,3	24117,7	24112,1
$\chi^2_{calc}(n)$	40,2		18,3	7,0	14,1	14,1	14,1	14,1
$\chi^2_{0,05}(n-1)$	14,1		12,6	11,1	14,1	14,1	14,1	14,1
σ_I	5,3		5,4	5,9	8,9	5,7	5,6	5,75
σ_E	12,6		9,45	6,9	12,6	8,1	7,9	8,15
$t_{n-1;\,0,05}$	2,365		2,447	2,571	2,365	2,365	2,365	2,365
σ	29,9	26	23,1	17,9	29,9	19,2	18,7	19,3

identifying mismatched results by minimizing the functional

$$S=\sum_{i=1}^{n}(1-S_i^2/\sigma_i^2)^2=min, \qquad (9)$$

where σ_i is the original error levels and S_i is the matched error levels. Many different error level extension algorithms have been and could be put forward. Reference [13], on the basis of the maximum likelihood principle, specifies the conditions which extension algorithms should satisfy. In particular, the joint probability distribution $Q(\sigma|S)$ should be in a normalizable form so that it can be interpreted as a probability distribution function. If we accept the principle that all the non-excluded mismatched data can be viewed as being equally reliable, then identical relative changes in the initial error levels have the same probability, i.e. the simplest solution is an algorithm which extends $\{\sigma_l\}$ by an identical number of times.

As an alternative to extending the error levels by an identical number of times, we can use an error level extension algorithm which yields a minimum value for σ_E^2. From expressions (2) and (5) it is clear that the minimum value of the function for the l variables $\{\sigma_l\}$ must be found for the matched data

$$\sigma_E^2=\frac{\varkappa_\alpha^2(n-1)}{(n-1)}\cdot\sigma_I^2=\frac{\varkappa_\alpha^2(n-1)}{(n-1)}\cdot\left[\sum_{i=1}^{n-l}\frac{1}{\sigma_i^2}+\sum_{j=n-l+1}^{n}\frac{1}{\sigma_j^2}\right]. \qquad (10)$$

Simple analysis shows that the σ_I^2 minimum occurs near the boundaries of the domain of the $\{\sigma_l\}$ arguments, and the maximum value occurs when l arguments assume identical values. This minimum value can be found in the following way: first of all, the smallest error level σ_{ml} is extended to the value S_{ml} so that the following condition is fulfilled

$$\varkappa_{calc}^2(n-l+1)\leqslant\varkappa_\alpha^2(n-l). \qquad (11)$$

Then we extend consecutively in rising order σ_{m2}, ... σ_{ml} ($\sigma_{ml} \leq \sigma_{m2} \leq \ldots \sigma_{ml}$) to obtain expression (5).

It should be noted that the extension algorithm which produces the least dispersion of σ_E^2 extends the error levels by different numbers of times, i.e. it works on the subjective assumption that mismatched data are of varying reliability without having analysed the physical and methodological way in which they were obtained. Thus, the formalized algorithms, such as expression (9) for instance, will also be subjective.

The nature and advantages of the methodology proposed in this paper for evaluating mismatched data is demonstrated in three specific examples. In Ref. [14], the half life of ^{234}U is evaluated on the basis of seven results: 2.439(24); 2.450(8); 2.458(12); 2.259(7); 2.270(30); 2.475(16); 2.520(8). The evaluated value is $T_{1/2} = 2.250(20) \cdot 10^5$ years. By calculation we obtain $\kappa^2_{calc}(7) = 50.3 > 12.6 = \kappa^2_{0.05}(6)$. The mismatched value is x_7, which is also the abnormal extreme value, since κ^2_{calc} (minus x_7) = 2.69 < 11.1 = $\kappa^2_{0.05}(5)$; $\lambda(7) = 2.1 > 2.0 = \lambda^{crit.}_{Rosner;7}(7)$. By excluding x_7 from further processing we obtain a consistent system which yields a value of $T_{1/2} = 2.457(4) \cdot 10^5$ years, including an increase in the error level σ_E by the amount of the Student coefficient in accordance with expression (4). Clearly, the value of $T_{1/2}$ increased by only 0.3% whereas the error level fell by a factor of 4.

Reference [15] evaluates $T_{1/2}$ for ^{239}Pu on the basis of eight results: 24 019(21); 24 089(23); 24 101(10); 24 102(20); 24 112(16); 24 124.2(136); 24 138.6(137); 24 164(14). x_1 was excluded and the arithmetic mean was taken as the evaluation with the standard deviation $T_{1/2}$ = 24 119(26) years. By calculation we find that x_1 and x_8 are mismatched. However, neither x_1 nor x_8 are abnormal extreme values: $\lambda_8(8) < \lambda_1(8) = 2.05 < 2.1 = \lambda_1^{crit.}(8)$ and $\lambda_{1.8}(8) = 1.8 < 2.0 = \lambda^{crit.}_{1.8}(8)$. The table compares the results obtained using various methods for extending σ_1 and σ_8 and from these results we can draw the following conclusion: if we extend the error levels of only the mismatched data, the error level of the evaluation σ is noticeably lower than when we use the algorithm which extends all the error levels. Extension of the mismatched error levels by an identical number of times on the assumption that the mismatched data are equally reliable is preferable, since inequal extension is subjective and is not compensated, in our opinion, by the small reduction in σ. The value for $T_{1/2}$ ^{239}Pu obtained by us by extending σ_1 and σ_8 by 2.17 times is $T_{1/2}$ = 24 113.3(192). Obviously, this $T_{1/2}$ value is 0.02% lower and the error level is 1.35 times lower than the corresponding values in Ref. [15]; they are 0.001% higher and 1.07 times higher respectively than the $T_{1/2}$ and σ values obtained by excluding the mismatched data x_1 and x_8. Reference [16] evaluates $T_{1/2}$ ^{239}Pu using as a basis the evaluation given in Ref. [15], which it accepts; six extra $T_{1/2}$ values are given which were not included in the processing. If we perform an evaluation based on all 14 results the mismatched data are, once again, x_1 = 24 019(21), which is not an abnormal extreme value, and x_{13} = 24 264(14). Extending σ_1 and σ_{13} by 2.4 times we obtain an evaluated value for $T_{1/2}$ ^{239}Pu = 24 119.9(137). The value obtained in this third example is 0.004% higher than the value obtained in Ref [16] and the error level is 1.9 times lower. All the evaluations in this paper have a confidence level of P = 0.95.

Thus, this method for processing mismatched data using two statistical criteria and extending the error levels of unexcluded mismatched results by an identical number of times results in practically no change in the evaluation, but it does significantly reduce its error level.

In conclusion, it should be noted that this methodology should only be applied after the data system has been subjected to expert evaluation, including the mismatched data for which various methods can be used such as the exclusion of insufficiently justified results and a re-evaluation of their error levels, and the selection of one highly accurate and convincing result as an evaluation. We are thus entirely in agreement with the evaluation given in Refs [14, 15, 16] and the data from those papers has been used to illustrate the method.

This methodology has a wide field of application, as we are constantly coming up against the problem of having to process systems of inconsistent data in various fields of science and technology. Apart from the above-mentioned areas, we have the example of Ref. [17] where the data are matched by extending all the (σ_i) values by an identical number of times.

REFERENCES

[1] BROWNLEY, K.A.[*], Statistical theory and methodology in science and technology, Nauka, Moscow (1977) [in Russian].

[2] BIRGE, R.T., Phys. Rev. 40 2 (1932) 207.

[3] AGEKYAN, T.A., Fundamentals of the theory of errors for astronomers and physicists, Nauka, Moscow (1968) [in Russian].

[4] Direct measurements with multiple observations, processing methods for observation results - Main principles, State All-Union Standard 8.207-76. GSI, Moscow (1986) [in Russian].

[5] FERGUSON, T.S., On the rejection of outliers, Proc. Fourth Berkely Symp. Math. Stat. Probability, Berkely, Vol. 1 (1961) 253.

[6] ID'E, V., DRAYARD, D., JAMES, F.[*], et al., Statistical methods in experimental physics, Atomizdat, Moscow (1976) [in Russian].

[7] GRIGORYAN, Yu.Y., SOKOLOVSKY, L.L., CHUKREEV, F.E., The evaluation of nuclear data, Report INDC(USSR)-75 LN, Vienna (1976).

[8] HIMMELBLAU, D.[*], Use of statistical methods to analyse processes, Mir, Moscow (1973) [in Russian].

[9]] ROSNER, B., Technometrics 17 2 (1975) 221.

[10] YABOROV, Yu.F., Analysis of some parametric criteria for identifying abnormal extreme values using sensitivity equations, III All-Union Conf. on the Problems of Obtaining and Using Data on the Characteristics of Materials and Substances in the Economy, Gosstandart SSSR, Moscow (1987) 189 [in Russian].

[11] Precision measurement and fundamental constants, USA: Nation. Bur. Stand., N343, Spec. Publ. (1971).

[12] TUNINSKIJ, V.S., KHOLIN, V.M., Metrologiya 3 (1975) 3 [in Russian].

[13] COHEN, E.[*], "Matching algorithm", Quantum Metrology and Fundamental Constants, Mir, Moscow (1981) [in Russian].

[14] ELLIS-AKAVALI, Y.A., Nucl. Data Sheets 40 4 (1983) 523.

[15] STROHM, W.W., Intern. J. Appl. Radiat. Isotop. 29 3 (1978) 481.

[16] SCHMORAK, M.R., Nucl. Data Sheets 40 1 (1983) 1.

[17] GIZATULLINA, V.G., KOZLOV, A.D., Methodological pointers: Evaluation of the reliability of data on the characteristics of petroleum and petroleum products, MI 656-84, Izd-vo Standartov, Moscow (1985) in Russian].

[*] Translator's Note: Spellings of foreign names transliterated back from the Slavonic are not necessarily reliable.