J. Indian Inst. Sci., Jan.–Feb., 1992, **72**, 27-33.

Reflections on uncertainty and errors

C. G. JESUDASON
Department of Chemistry, University of Malaya, Kuala Lumpur 59100, West Malaysia

Received on September 20, 1990; Revised on May 6, 1991

Abstract

It is argued that errors in experiments wholely attributable to measurements are based largely on views of classical physics. The statistical interpretation of systems requires a generalization of error concepts such as are proposed here for Gaussian and other experimental variable distributions. It is concluded that the acceptability of theories is conditioned by uncertainties in measurements, and consensus is required to set values for error limits.

Key words: Experimental error analysis, measurements, fluctuations, uncertainty.

1. Introduction

Prior to the introduction and prevalence of statistical interpretation of systems, classical phenomenological equations of physics could be written $Q(\mathbf{P}, d^n\mathbf{P}/dt^n, t) = 0$ where the physical quantities represented by variables $\mathbf{P}$ (and their derivatives) for any observable Q could take only one value at a particular time t, and that any fluctuation from the 'actual' value was due to random errors, which "cause successive readings to spread about the true value of the quantity"[1]. However, if it can be shown that the 'true value' of a system quantity is not always unique, then it follows that the usual interpretation of error estimates[1] as provided in reports would require further qualification and elaboration.

Section 2 addresses the non-uniqueness of the true value of the physical quantity according to recent theories which therefore necessitates further qualification in error estimates, and also argues that theory determines the nature of error and the meaning of the variable that is being measured. The extensions and generalizations required are presented in Section 4, which follows from the conventional representation briefly reviewed in Section 3. Section 5 concludes with a suggestion for consensus in deciding on the limits for error estimates.

2. Theoretical significance of mean values and dispersions

The mean value of a quantity, even if it may be derived to any arbitrary degree of accuracy from experiment, may not theoretically represent the 'true value' for the

system necessarily. Statistically considered, a whole range of values, usually clustered about the mean are all equally legitimate, occurring with varying probability and may be predicted by theory. In quantum mechanics, this corresponds to the situation $\bar{a} = \langle \Psi, \hat{A}\Psi \rangle$, $\Delta A^2 = \langle (\hat{A} - \bar{a})\Psi, (\hat{A} - \bar{a})\Psi \rangle \neq 0$ where $\hat{A}$ is the operator for a system in state Ψ, which is well defined, and $\bar{a}$ the mean of observable A. If quantum mechanics is not deemed suitable, then even classical statistical mechanics also has constructed systems with a range of values. For instance, in the canonical ensemble, the probability $P(E)$ of observing a particular system in energy state E is $P(E) = P(\bar{E})\exp - (E - \bar{E})^2/2kT^2C_v$, (henceforth all barred quantities are averages) where C_v is the heat capacity at constant volume, T the temperature and k Boltzmann's constant. Even the size or material constituent of a system need not remain invariant, as illustrated by the grand canonical ensemble, where the variance σ_N^2 of the particle density is $\sigma_N^2 = \bar{N}^2kTK/v$, where K is the isothermal compressibility in volume v, and $\bar{N}$ the mean particle number. In the above interpretations, nature has been extended from the Platonic description of invariant ideality to that of the Aristotelian 'species', which is the Latinized $\epsilon\iota\delta o\varsigma$ of Plato, where the range of motion is contained within the species type[2]. Another well-known example of this view of nature is in Brownian motion theory.

We denote a quantity A which may exist as a range of values as range positive and symbolically enclose it in curly brackets, *i.e.*, $\{A\}$. However, there are instances where current theory predicts (and in fact demands) exact measurements, *i.e.*, $\Delta A^2 = 0$ in quantum mechanical language. Then, instead of an equally valid range $\{a\}$, we have a single value for the observable a, (which is termed a range-negative variable) and which may be experimentally presented as $a = \bar{a} \pm 2\ \sigma_{er}$ where $\sigma_{er} \to 0$ in principle if 'systematic' errors are eliminated. It is thus clear that whether or not a variable is range positive is theory dependent Further, it is misleading to write the error as $2\sigma_{er}$ for range-positive variables, which is only reserved for range-negative variables, since, in the absence of any external experimental error sources, there would still exist the dispersion which would be measured by the experimentally derived $2\sigma_{er}$ factor, where $2\sigma_{er}$ now refers not to the experimental error estimates, but to the intrinsic dispersal of the system. Hence, there must exist a convention that distinguishes the two cases, as is suggested in Section 4 for range-positive and range-negative variables.

From the above, it is clear that whether a variable is range positive or negative is theory dependent, and a clash of theory would ascribe contrary variable designations, depending on the theory choice. An example in classical theory is the temperature parameter T, which essentially measures the energy exchange propensity in a system and is thus range negative. In such instances, there exists a theoretical possibility of an absolute measurement for, as indicated, theory dictates the nature of certainty. Recently, it has been claimed by workers like Lavenda[3] and Schlögl[4] that "if the energy is fixed, a definite temperature cannot be assigned to the system. . . .and it must be supposed that the temperature of the system undergoes fluctuations"[3]. However, a micro-canonical ensemble is precisely of this form, where the Lagrange multiplier $\beta(=1/kT)$ in maximizing the entropy at equilibrium (the

maximum entropy method (or MAX)[5]) is the fixed or invariant temperature parameter. He adds "on account of the duality between measurements of the temperature and energy of a system in contact with a heat bath, we may consider β itself as a realized value of a random process equipped with an *a priori* density $\omega(\beta)$. . . ."[3]. However, β is determined as an invariant parameter relevant for *all* systems through the Maxwellian Zeroth Law through *thermal contact*; if there is no net energy exchange between systems A and B as time $t \to \infty$, then $T_A = T_B$ by definition, irrespective of the variance of the energies ΔE_A^2 (or ΔE_B^2). For statistical systems, one can always develop an energy exchange metric $\| M \|$ such that the condition $\| M \| \to 0$ as $t \to \infty$ corresponding to an invariant, unique temperature, as has been constructed[6]. The physical measurement of temperature is highly correlated to the Zeroth law, and any other theory not in consonance with it is not describing the same phenomenon; indeed, Lavenda states that the maximum entropy formalism is "incompatible with the statistical inference approach"[3]. Zeroth law considerations alone are sufficient to cast severe doubts over other temperature concepts. We now illustrate contrary variable designations due to different theory choices and the impact of this on error estimates by examining the temperature variable.

Let the temperature T be measurable from some observable in the system where T is a linear function of the energy absorbed E by the (canonical ensemble) system to first order. The Lavenda theory (LT) suggests that each of these T values represents a temperature whose average $\bar{T} = \int_0^\infty T(E)P(E)dE$ where $P(E)$ is the probability function whereas MAX demands that $\bar{T} = \bar{T}(E_0) = T_0 = T(E)$ is the only temperature. Expanding $T = T_0 + \sum_{i=1}^{\infty} \alpha_i \ (E - E_0)^i$, where $\alpha_i = 1/i! \ \partial^i T / \partial E^i|_{E_0}$, setting $\bar{T} = T_0$ demands $\sum_{i=1}^{\infty} \alpha_i \sigma_E^i = 0$ where $\sigma_E^i = \int_0^\infty (E - E_0)^i P(E)dE$, the ith moment. According to MAX, $\sigma_E^2 = kT^2C_v$ for a canonical ensemble, so that $\langle (T - T_0)^2 \rangle = \sigma_T^2 = kT^2C_v\alpha^2$ where $\alpha = dT/dE|_{E_0}$. Hence, the Lavenda-type theories would predict variance σ_T^2 for the observation of temperatures with mean value $\bar{T}$, whilst MAX would consider $\bar{T}$ only to represent the temperature of a system in contact with a heat bath. Hence, measurement which yields a spread of values would be interpreted as error or noise contributions according to MAX, whereas LT would consider the spread to be a verification of the theory. Further, for MAX, the experimental report would present the spread contribution as error ($\pm 2 \ \sigma_E$ for Gaussian processes usually), and for Gaussian processes, the 'mean of means' estimate (M.M.E.), as elaborated and recommended in most texts[1] (and discussed below) would be used for all range-negative variables, whereas there does not exist any convention at present to present data for range-positive variables which clearly distinguishes error from the intrinsic spread of true values.

3. Discussion of normal presentation (M.M.E.)

In range-negative variable measurements, it is assumed that each successive measurement A_i of variable A constitutes an independent random stochastic process such

that if variance V and mean E of each random variable A_i in a sequence is given by $V(A_i) = \sigma^2$ and $E(A_i) = \mu$, then the expectation E of $\bar{A} = \sum_{i=1}^{n} A_i/n$ is $E(\bar{A}) = \mu$ and $V(\bar{A}) = \sigma^2/n$. In this form of interpretation $V(\bar{A}) \to 0$ as $n \to \infty$, so that a mean value, which is also the 'true value' is attained to any desired accuracy within the limitations of the probabilistic model. Thus, if we assume that the mean distribution is Gaussian, the '95% confidence' convention (corresponding to 1·96 or approximately two standard derivations) allows us to write

$$A = \bar{A} = A_{av} \pm 2\sigma_{est} / \sqrt{n} \tag{1}$$

where Bessel's formula[7], for instance, gives the estimate σ_{est} for σ as $\sigma^2_{est} = ns^2/(n\text{-}1)$, where the computed standard deviation (s.d.) s is $s = [1/n \sum_{i=1}^{n} (X_i - \bar{X})^2]^{1/2}$ for a series of n successive measurements. A_{av} may be computed as $A_{av} = \sum_{i=1}^{n} X_i/n$. Equation (1) is the M.M.E. where the uncertainty or error for the mean is $2\sigma_{est}/\sqrt{n}$, and may be termed the standard model. The Xs are the data points deemed independent random variables representing quantity A.

4. Extension to range-positive variable measurements

In the case of a spectrum $\{Y\}$ of range-positive measurements, the error in measurement is usually the random scatter of points due to the environment for *each* system value Y. In quantum mechanical language, what is actually being measured M' is the sum of the environmental and probe contributions δM and the pure system value M_0, *i.e.*, $M' = \langle \Psi_t, (\hat{M}_0 + \delta\hat{M})\Psi_t \rangle$ from which $M_0 = \langle \Psi_0, \hat{M}_0\Psi_0 \rangle$ is to be inferred Hence, suitable perturbation models must exist to eliminate the ∂M environmental contributions and to map the total wavefunction Ψ_t to the state function Ψ_0. $\Psi_t \to \Psi_0$, as in perturbation theory. We may generalize on the standard (M.M.E.) model as follows, for the simplest case.

4.1. *With Gaussian assumption*

If the Y system value pure random variable giving rise to the range $\{Y\}$ is interfered with an environmental random variable X which leads to only *one* observable outcome $U = Y + X$, then if Y and X are Gaussian variables with variances σ_y^2 and σ_x^2 and means μ_y and μ_x, respectively, then it is well known that the variable[7] U (corresponding to the observable) is also Gaussian with mean $\mu_u = \mu_x + \mu_y$ and variance $\sigma_u^2 = \sigma_y^2 + \sigma_x^2$. In this case, the true value of Y is a *range* $\bar{Y} \pm 2\sigma_y$ with error factor $\pm 2\ \sigma_x$ if $\mu_x = 0$ and if $\sigma_y^2 \ll \sigma_x^2$. The 95% confidence convention is retained for the range $\{Y\}$. Since the U observable is Gaussian, then we may determine $\bar{U} = \bar{Y}$ to any degree of accuracy following the method of M.M.E., if $\mu_x = 0$. The value of σ_x^2 may be estimated experimentally and is the variance of the noise contribution Δ^2 Exp (*i.e.*, $\sigma_x^2 = \Delta^2$ Exp) when the system is not probed. Thus, the Gaussian probability distribution function for Y may be determined by subtraction from the experimental variance σ_u^2 to determine σ_y^2 for linear systems. We shall call this procedure the statistical variable estimate (S.V.E.) for range-positive variables.

If theory demands a spectrum of values for variable A, then we write the S.V.E. result (for stationary, symmetrical Gaussian distributions) following the convention below

$$\{A\} = \bar{A} \pm 2\,\sigma_A \pm 2\,\Delta\,\mathrm{Exp} \tag{2}$$

where we have retained the 95% confidence criteria for range $\{A\}$. Δ Exp is the estimated error due not to the 'pure' spectrum of range $\bar{A} \pm 2\sigma_A$ but to the assumed experimental errors, and $2\sigma_A$ is derived from subtracting the experimentally derived spectrum for U from Δ Exp attributable to 'interference' during measurement of the 'pure system', *i.e.*, $\sigma_u^2 - \Delta^2\mathrm{Exp} = \sigma_A^2$. The convention is particularly suitable if $|\sigma_A| \gg \Delta$ Exp. The mean value for variable $\bar{A}$ ($=\bar{U}$) may be computed to any degree of accuracy even if A exists theoretically as range $\{A\}$. For large number of trials n, we may write $\sigma_u^2 \simeq \sigma_{est}$, and since for M.M.E $\sigma_{cr} \simeq \sigma_u^2/\sqrt{n}$, A may be determined to any degree of accuracy in principle. Since $\sigma_{est}/\sqrt{n} \ll \sigma_A$, the $\sigma_{est}/\sqrt{n}$ term may be absorbed by the Δ Exp term, where the $\bar{A}$ of (M.M.E.) is the same as for (S.V.E.) as written in eqn (2). Conventions relating to the relative magnitudes of these quantities before error terms may be neglected must be stated prior to the usage of expressions such as (1) and (2) The assumption made here is that the mean of the noise contribution $\mu_y = 0$, and may be adjusted accordingly if it is not the case.

4.2. *Without Gaussian assumption*

The spectra for (stationary) systems without symmetrical Gaussian distributions must be presented as variable quantities (preferably graphically with the estimated errors indicated). What is required is the use of conditional probabilities to estimate the error components. Mathematically, the actual probability distribution for the observable variable X, $f_u(X)$ is

$$f_u(X) = \int_{-\infty}^{\infty} f(X \mid Y)\, g(Y)\, \mathrm{d}Y \tag{3a}$$

where $f(X \mid Y)$ is the conditional probability density of experimentally determined variable X given the pure system value Y and g the probability density for distribution Y. The determination of the probability distribution for the pure system is *via* the posterior determination, *i.e.*,

$$g(Y) = \int_{-\infty}^{\infty} f(Y \mid X) f_u(X) \mathrm{d}X. \tag{3b}$$

The 'pure spectrum' may then be superimposed on the experimental spectrum to determine the extent of errors in the ideal case or both the probability distribution functions may be written down (pure and experimental) with their means and standard deviations quoted. The method of conditioning is used to determine the above functions $g(Y)$ and $f_u(X)$[8].

4.3. *Errors and theory*

Lastly, write the predicted spectra as f_t, the observed as f_0 and the experimental

error contributions deemed separate from the pure systems as Σf_i. Then, we define a discrepancy factor as Δ where

$$| f_0 \ominus f_t \ominus \sum_i f_i | = \Delta \tag{4}$$

and where $\ominus$ means subtracting the theoretical error estimates in f_t (*e.g.*, as in (1) above) and $\sum_i f_i$ where each f_i is a randomness factor due to the environment (or measuring process). If $f_o = A \pm F$, $f_t = C \pm D$, $\sum_i f_i = \pm E$, then (4) means $| A - C | + | F - D | + 2E$. The $\pm F$ term is the experimentally determined fluctuational spread derived by subtraction from the estimated $\sum_i f_i$ due to the environment. If Δ is small, then we have a reasonable theory if the derived spectrum f_0 is similar to the theoretical spectrum f_t in appearance; the factor that causes a theory to be acceptable or not is actually Σf_i, the estimated error of measurement due to the environment or interference; if Σf_i is very large, f_t may be accepted since $| f_o \ominus \sum_i f_i |$ is large, and may accommodate f_0 *i.e.*, $| f_o \ominus \sum_i f_i | \gg | f_t |$. If $| \sum_i f_i | \to 0$ and $| f_t | \to\!\!\!\!\times\!\!\!\!\leftarrow | f_o \ominus \sum_i f_i |$, then the theory of the pure system is modified until a matching occurs, which generally increases the credibility of a theory. Often, especially in the undergraduate laboratories, the meaning of error is misconstrued as meaning merely the difference between the theoretical spectrum and the experimental one; however, the factor due to the environment $\sum_i f_i$ must be incorporated or else there might arise the situation of optimistic report of small errors which subsequently proved to be large[9].

Let the random variable belonging to the pure system be ξ_i and an environmental interfering variable be ξ_j. Then one must always stipulate that the covariance of ξ_i and ξ_j, $Cov\ (\xi_i, \xi_j) = 0$ or is small as an ideal, so that these systematic errors are eliminated. An 'improvement' of a measuring technique, guided by a current theory, is largely concerned with ensuring a small covariance so that $\Delta\ \text{Exp} \to 0$; the search for such conditions would also cause further theories to be developed since observations on cause and effect would increase the propensity towards newer formulations. At this juncture, human factors normally excluded from scientific discourse must be admitted to describe the nature of interpretation.

5. Conclusion

From the above, it is clear that for range-positive variables, the current conventions using the M.M.E. are not adequate in providing a format of representation which distinguishes fluctuations from 'erroneous' environmental effects. Even so, Section 4 argues that the limits for what is considered to be a significant value for an experimental variable is still subjected to a community consensus, or, failing that, to error margins defined and stated when presenting data.

Acknowledgement

A discussion with Ong Seng Huat, Department of Mathematics, University of Malaya, is acknowledged.

References

1 SQUIRES, G L — *Practical physics*, Second edn, 1976, p 9, McGraw-Hill

2 CONFORD, F M — *Before and after Socrates*, 1972, pp 78-79, Cambridge University Press

3 LAVENDA, B H — On the phenomenological basis of statistical thermodynamics, *J Phys Chem Solids*, 1988, **49**, 685–693 (especially Sec II after eqn (9), Sec III, first and second paragraphs, and Sec IV, p 689)

4 SCHLOGL, F — Thermodynamic uncertainty relation, *J Phys Chem Solids*, 1988, **49**, 679–683

5 MCQUARRIE, D A — *Statistical thermodynamics*, 1973, pp 37 and 56, Harper & Row

6 JESUDASON, C G — Recoverable systems A quantum statistical thermodynamic study, *Indian J Pure Appl Phys*, 1991, **29**, 163–182

7 TOPPING, J — *Errors of observation and their treatment*, 1960, p 63, Chapman and Hall, Unwin Pubs

8 SCHEAFFER, R L AND MENDENHALL, W — *Introduction to probability Theory and applications*, 1975, p 203, Duxbury Press

9 — See discussion of Millikan oil-drop experiment, Ref 1, p 11