

Avaliação e Expressão de Incerteza em Medição

(Evaluation and expression of uncertainty in measurement)

José Henrique Vuolo

Instituto de Física, Universidade de São Paulo

Caixa Postal 66318, CEP 05315-970, São Paulo, SP Brasil

Recebido em 5 de maio, 1998

Neste artigo, são resumidas a terminologia e as regras gerais relativas à avaliação e expressão de incertezas de medição, conforme recomendações de expressivas organizações internacionais (BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML). Embora estas recomendações sejam relativamente recentes, parece importante sejam divulgadas e introduzidas nos cursos experimentais para ciências exatas. Também são discutidas algumas expressões (em português), sobre as quais existe alguma confusão na utilização ou na tradução.

In this article, it is presented the terminology and the general rules for the evaluating and expressing uncertainty in measurement, according to the recommendations of expressive international organizations (BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML). It seems important to divulge this terminology and the basic rules and to introduce them in experimental courses for exact sciences. Moreover, it is discussed the translation (for Brazil) or the utilization of some terms.

I Introdução

A terminologia sobre erros e incertezas de medição sempre foi um pouco confusa, a começar da confusão entre “erro” e “incerteza”. A forma para expressar a incerteza também nunca foi muito bem definida, desde o número de algarismos a serem escritos num resultado até o significado exato do que está sendo indicado como incerteza. Além disso, as regras para avaliação de incertezas também sempre foram um pouco controvertidas, no que se refere à avaliação de incertezas resultantes de efeitos sistemáticos e combinação destas com as resultantes de efeitos aleatórios. As diferentes concepções sobre estas regras é discutida em [1] e [2]. Além dessas questões, ainda pode ser mencionado o fato que as incertezas nos valores das grandezas físicas fundamentais têm sido historicamente subestimadas. Por exemplo [3], nos valores anteriores a 1963, as incertezas foram aproximadamente subestimadas por um fator 3, em média. É provável que o mesmo ou pior tenha ocorrido em outras medições físicas e que a subestimação de incertezas tenha persistido nas décadas seguintes.

Um esforço para resolver tais questões vem sendo realizado, há mais de 10 anos, por grupos de trabalho constituídos por especialistas indicados pelas organizações internacionais *Bureau International des Poids et Mesures (BIPM)*, *International Electrotechnical Commission (IEC)*, *International Federation of Clinical Chemistry (IFCC)*, *International Organization for Standardization (ISO)*, *International Union of*

Pure and Applied Chemistry (IUPAC), *International Union of Pure and Applied Physics (IUPAP)* e *International Organization of Legal Metrology (OIML)*. Um dos resultados dos trabalhos é o *Guide to the Expression of Uncertainty in Measurement*, conhecido pela sigla GUM (ou ISO-GUM) e publicado em 1993, em nome das organizações citadas [4]. Este documento de mais de 100 páginas e tiragem total de 7500 exemplares tem sido universalmente difundido e aceito, tendo sido já traduzido para mais de 10 idiomas. Estas informações e um histórico sobre a elaboração do GUM são dadas pelo Prof. Mathiesen [5]. Uma versão do GUM para o Brasil foi patrocinada pelo INMETRO [6], sendo que já existe uma tradução em Portugal. Documentos importantes tais como EAL-R2 [7] e NIST-TN 1297 [8] têm sido elaborados conforme terminologia e princípios básicos do GUM.

Quanto à nomenclatura, o *International Vocabulary of Basic and General Terms in Metrology (VIM)*, [9,10] resolve algumas questões. Entretanto, importantes expressões relativas a incertezas, que são propostas no GUM, ainda não foram incorporadas ao VIM. Como exemplos, podem ser mencionadas “incerteza padrão”, “incerteza tipo A”, “incerteza tipo B”, “incerteza combinada” e “incerteza expandida”.

A avaliação e expressão da incerteza de maneira uniformizada, se tornou muito importante em tecnologia, devido à globalização da economia. É evidente que um tratamento uniforme também deveria ser dedicado a medições físicas, em geral. A elaboração do GUM ocor-

reu por iniciativa do BIPM [11] e, posteriormente, contou com a participação das outras entidades, inclusive a IUPAP.

Neste artigo, procurou-se resumir as definições e regras gerais relativas a incertezas, conforme os princípios estabelecidos no GUM. Além de divulgar a terminologia correspondente (em inglês) proposta no GUM, procurou-se discutir a tradução de alguns termos controversos e também de outros, geralmente utilizados de maneira confusa. Parece importante conhecer esta nomenclatura e regras básicas, incorporando-as em disciplinas experimentais de cursos de graduação em ciências exatas.

II Discussão de alguns termos

Alguns termos importantes, sobre os quais existe alguma confusão, são discutidos nesta Seção. Um vocabulário relativo a incertezas, identificando as expressões que aparecem no VIM, no GUM ou não aparecem num ou noutro, é apresentado em [12].

Medição (Measurement)

A palavra “medição” é a recomendada na versão brasileira do VIM, aqui chamada de VIM-BR [10]. Esta é a palavra correta para se referir ao “ato de medir”, conforme dicionário.

A palavra “medida” é extensivamente usada com o sentido de “medição”. Entretanto, isto deveria ser evitado porque não é muito correto e também porque “medida” tem vários outros significados, tais como em “o desvio padrão é uma *medida* (measure) de dispersão” ou “a altura já foi *medida*” ou ainda “a última *medida* (resultado) é a melhor”.

Acurácia ou exatidão (accuracy)

A acurácia (ou exatidão) indica a qualidade do resultado da medição no que se refere à incerteza final. A tradução recomendada no VIM-BR é “exatidão”. Entretanto, os adjetivos correspondentes (exato, exata) são muito fortes e têm um significado bem definido, que é o indicado no dicionário e correspondente à palavra inglesa “exact”. Além disso, a palavra acurácia já tem sido usada em outros textos [12, 13]. Outra vantagem de “acurácia” é a similaridade com “accuracy”. Por estes motivos, a tradução recomendada no VIM-BR deveria ser revisada, pelo menos deixando “acurácia” como alternativa.

Precisão (precision)

A *precisão* é uma indicação parcial da qualidade da medição, que se refere apenas a flutuações aleatórias. Além de boa precisão, é necessário que os efeitos sistemáticos sejam pequenos para se ter boa acurácia. A

palavra “precisão” (precision) é universalmente aceita com este significado. Por isso, embora exista controvérsia entre os termos “acurácia” e “exatidão”, é inadmissível traduzir “accuracy” como “precisão” ou usar esta palavra para indicar a qualidade da incerteza final de um resultado, o que infelizmente tem ocorrido com frequência em manuais técnicos e até mesmo em textos científicos.

Mensurando (measurand)

Mensurando é definido no VIM como “a grandeza específica submetida a medição”. Entretanto, a definição do mensurando numa medição específica é uma questão um pouco delicada. Ocorre que, em geral, o nível de detalhamento da definição depende da acurácia permitida pelo próprio processo de medição.

Por exemplo, “índice de refração do ar” pode ser um mensurando bem definido para uma experiência simples. Numa medição mais elaborada devem ser especificados comprimento de onda, temperatura e pressão atmosférica na definição do mensurando. Melhorando mais ainda a acurácia, deve-se especificar também a composição da amostra de ar (inclusive impurezas) e assim por diante.

No formalismo para avaliação de incerteza, o valor (verdadeiro) do mensurando é uma quantidade *desconhecida* e *desconhecível* (que não pode ser conhecida). Deve ser observado que a palavra “verdadeiro” é redundante na expressão “valor verdadeiro do mensurando” e pode-se usar apenas “valor do mensurando”, como recomendado no GUM. Entretanto, em certas circunstâncias, especialmente para fins didáticos, pode ser útil ou importante enfatizar que se trata do valor verdadeiro.

Erro (error)

O erro η é a diferença entre o resultado y da medição e o valor do mensurando y_v :

$$\eta = y - y_v . \quad (1)$$

Uma vez que o valor do mensurando é uma quantidade *desconhecida* e *desconhecível*, resulta que o erro de medição também é uma quantidade *desconhecida* e *desconhecível*, no formalismo para avaliação de incertezas.

Em circunstâncias excepcionais, o mensurando é conhecido com acurácia muito melhor que a permitida pela medição. Por exemplo, isto pode ocorrer na aferição de um equipamento e também é comum em experiências didáticas. Em tais casos, o erro pode ser conhecido e isto gera certa confusão. Por exemplo, um estudante realiza a experiência de Millikan e a *incerteza* na carga do elétron deve ser avaliada, como parte do resultado final. Neste caso particular, poderá ser calculado também o “erro de medição”, porque a carga do elétron é conhecida com acurácia muitíssimo melhor

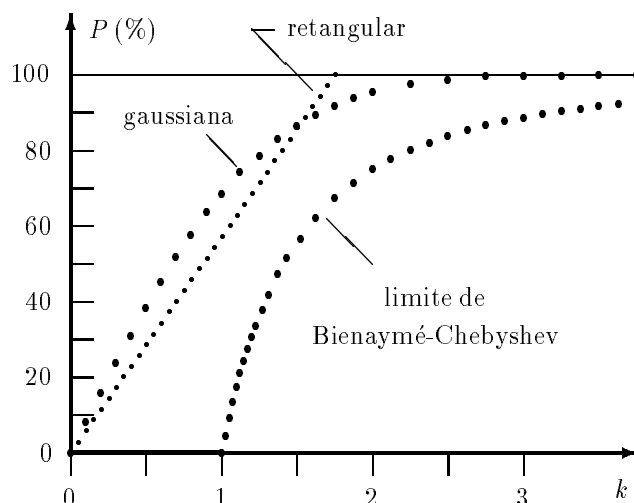


Figure 1.: Níveis de confiança para valores grandes de ν .

que a permitida por uma experiência didática simples. Entretanto, em nenhuma parte do formalismo para avaliação de incerteza, pode ser considerado que o mensurando ou o erro sejam conhecidos.

Incerteza (uncertainty)

Incerteza é um conceito qualitativo definido no VIM como “parâmetro associado ao resultado de uma medição que caracteriza a dispersão de valores que pode ser fundamentadamente atribuídos ao mensurando”. A incerteza, como conceito quantitativo, deve ser devidamente qualificada com adjetivo conveniente.

Como pode ser visto, “erro” e “incerteza” são conceitos bastante diferentes, que devem ser escrupulosamente distinguidos, especialmente para fins didáticos.

A *incerteza padrão* (standard uncertainty) é a incerteza dada na forma de *desvio padrão*.

A *incerteza tipo A* é a incerteza avaliada a partir da análise de uma série de observações, realizada conforme os métodos da *estatística clássica*.

A *incerteza tipo B* é a incerteza avaliada por quaisquer outros métodos, que não os métodos estatísticos clássicos. Em geral, para estimar a incerteza tipo B, os métodos empregados correspondem à *estatística bayesiana* [2, 14, 15].

Incerteza padrão tipo A e *incerteza padrão tipo B* são as incertezas tipo A e tipo B dadas na forma de desvio padrão, respectivamente.

Incerteza combinada é a que resulta da combinação de incertezas tipo A e tipo B, para se obter a incerteza final. *Incerteza padrão combinada* é a incerteza combinada dada na forma de desvio padrão.

Intervalo de confiança (interval of confidence)

A confiança P de uma afirmativa é a probabilidade de que esta afirmativa seja correta. Se Z é o valor verdadeiro de uma quantidade determinada por métodos estatísticos, pode-se considerar a afirmativa “ $z_1 \leq Z \leq z_2$ ” com probabilidade P de ser correta. O intervalo $[z_1, z_2]$ é um *intervalo de confiança* P para a quantidade Z e P é o *nível de confiança*.

No caso mais simples, z é a média de n medições de Z , sendo s o desvio padrão experimental na média. A precisão na determinação de Z pode ser avaliada pelo intervalo de confiança:

$$z - ks \leq Z \leq z + ks \quad (\text{com confiança } P). \quad (2)$$

A Figura 1 mostra $P(\%)$ em função de k para algumas distribuições de probabilidade, no limite de grandes valores de n (número de graus de liberdade ν muito grande). No caso em que o número de graus de liberdade é pequeno, o nível de confiança é menor. Por exemplo, no caso da distribuição gaussiana, o nível de confiança $P(k)$ para cada ν pode ser obtido a partir da *distribuição-t de Student* [4, 13, 14, por exemplo].

A expressão “confidence interval” é utilizada no GUM para se referir ao conceito desta expressão, tal como definido na estatística clássica. Para se referir a um “intervalo de confiança” para um mensurando determinado a partir de uma medição, envolvendo inclusive incertezas tipo B, a expressão utilizada no GUM é “interval of confidence”. Uma distinção deste tipo é impossível em português. Por isso, a expressão utilizada aqui é “intervalo de confiança” simplesmente, nos dois casos. As mesmas considerações se aplicam à expressão “nível de confiança”. Uma sugestão de tradução é “intervalo de credibilidade”. Na Referência [15], a expressão “credibility interval” é utilizada para se referir ao correspondente bayesiano do “intervalo de confiança clássico”.

Incerteza expandida (expanded uncertainty)

Incerteza expandida (expanded uncertainty) é a incerteza padrão multiplicada por um *fator de abrangência* k (coverage factor), de forma a definir um intervalo de confiança maior que o correspondente à incerteza padrão.

Repetitividade e reprodutibilidade

Repetitividade e reprodutibilidade (repeatability and reproducibility) são termos distinguidos no VIM. Repetitividade se refere ao grau de concordância entre resultados de medições repetidas exatamente nas mesmas condições. *Reprodutibilidade* se refere à medição do mesmo mensurando em condições modificadas ou diferentes.

Efeitos sistemáticos e aleatórios

Expressões tais como “erro aleatório”, “erro sistemático”, “incerteza aleatória” e “incerteza sistemática” são tradicionalmente usadas em física. Entretanto, esta nomenclatura não é usada no GUM que, além disso, recomenda que não seja utilizada. A justificativa para isto é o caráter relativo do que seja efeito sistemático ou efeito aleatório. Um exemplo simples é o erro de ajuste de “zero” de um instrumento, que pode ser sistemático para uma série de medições. Entretanto, se o “zero” é ajustado para cada medição, o erro se torna aleatório.

Numa medição específica, erros sistemáticos e erros aleatórios, bem como as respectivas incertezas, ficam bem caracterizados. Por isso, parece preferível manter esta nomenclatura tradicional, para fins didáticos. Entretanto, deve sempre ficar claro o caráter relativo da distinção entre “erro aleatório” e “erro sistemático”.

Independente destas questões, uma *correção* num determinado resultado nunca deveria ser identificada como “erro sistemático”, como muitas vezes ocorre. Em geral, uma “correção” acaba resultando em “erro sistemático” porque a correção nunca é perfeita.

As expressões “erro aleatório” (random error) e “erro estatístico” (statistical error) são utilizadas com o mesmo significado em muitos textos [12, 13, 14, por exemplo]. Entretanto, a expressão “erro aleatório” parece mais correta. A expressão “erro estatístico” é mais adequada para caracterizar o erro de “amostragem”, quando a própria grandeza sob medição é aleatória. Por exemplo, na medição de radioatividade, o sistema de detecção pode ser capaz de contar exatamente o número de partículas que atinge o detector. Entretanto, existe “erro estatístico” devido à flutuação intrínseca no número de partículas que atingem o detector, num determinado intervalo de tempo.

III Avaliação da incerteza tipo A

A incerteza padrão tipo A (u_A) pode ser identificada com o *desvio padrão experimental* que é uma estimativa não-tendenciosa (unbiased) para o desvio padrão [14, por exemplo].

No caso mais simples, a medição é repetida n vezes exatamente nas mesmas condições obtendo-se os resultados y_1, y_2, \dots, y_n . A melhor estimativa para o valor do mensurando é a média \bar{y} e a estimativa não-tendenciosa para a incerteza tipo A é

$$u_A = \frac{u}{\sqrt{n}}, \quad \text{onde} \quad u^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2. \quad (3)$$

Neste caso, o *número de graus de liberdade* é $\nu = (n - 1)$.

Casos mais complicados, tais como grandezas determinadas a partir de ajuste de funções, são usualmente discutidos em livros de tratamento estatístico de dados experimentais [12 a 15, por exemplo]. Além da incerteza padrão tipo A em cada grandeza, devem ser indicadas também as *covariâncias* e o correspondente *número de graus de liberdade*.

IV Avaliação da incerteza tipo B

A incerteza padrão tipo B também deve ser dada na forma de desvio padrão. Entretanto, não existe a estatística convencional para fazer isto, simplesmente porque não existem várias observações. No que segue, são transcritos em itálico, alguns trechos do GUM.

Se uma quantidade de entrada X não é determinada por meio de observações repetidas, a incerteza tipo B é *avaliada pelo julgamento científico baseado em toda informação disponível sobre a variabilidade da quantidade de entrada. O conjunto (pool) de informações pode incluir dados de experiências prévias, experiência ou conhecimento geral do comportamento e*

propriedades dos materiais e instrumentos relevantes, especificações de fabricantes, dados fornecidos em certificados de calibração e outros certificados e incertezas atribuídas a dados de referência obtidos em manuais (handbooks).

Um dos problemas é que a avaliação da incerteza tipo B é bastante subjetiva pois reflete, em grande parte, o grau de conhecimento do avaliador sobre o mensurando e a medição. No que segue, é transcrito um trecho do GUM [4] que, embora não seja referente especificamente a incerteza tipo B, parece se aplicar particularmente bem a este tipo de avaliação.

“3.4.8. Although this guide provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and the utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value.”

O procedimento para determinação da incerteza tipo B consiste em admitir, para os valores possíveis de X , uma distribuição de probabilidades que esteja de acordo com todo conhecimento e informação disponíveis sobre a “variabilidade” desta quantidade. O termo “variabilidade” se refere a valores possíveis de X , que tem valor único. O procedimento estabelecido no GUM para avaliação de incerteza tipo B corresponde ao princípio inicial da *estatística bayesiana*, que consiste em admitir uma distribuição de probabilidades a priori para a variável aleatória. Discussões mais detalhadas são apresentadas em [2], [14] e [15], por exemplo. A seguir, são discutidos alguns exemplos.

a. Distribuição retangular

Como exemplo, uma quantidade de entrada X está num intervalo entre a_- e a_+ , sendo que isto é tudo o que se sabe sobre a “variabilidade” de X . A única alternativa aceitável é admitir que X pode estar em qualquer ponto do intervalo com igual probabilidade. Isto é, pode-se admitir, para os valores possíveis de X , uma distribuição retangular de probabilidades de largura $2a$. Assim, resulta que a melhor estimativa para X é

$$x = \frac{a_+ + a_-}{2} \quad (4)$$

e a incerteza padrão é o desvio padrão da distribuição retangular:

$$u = \frac{a}{\sqrt{3}}, \quad \text{onde } a = \frac{a_+ - a_-}{2}. \quad (5)$$

Este modelo para a distribuição de probabilidades se aplica quando não existe absolutamente nenhuma

informação sobre X , exceto que esta quantidade está entre a_- e a_+ . Qualquer informação adicional implica na necessidade de modificar o modelo. Se, por exemplo, qualquer informação adicional indicar que é mais provável que X esteja no centro do intervalo, pode-se adotar uma distribuição “trapezoidal” (triângulo sobre o retângulo) ou uma distribuição triangular. Estas distribuições e exemplos são discutidos no GUM [4,6].

b. Distribuição de Laplace-Gauss

A distribuição de Laplace-Gauss, também chamada de gaussiana ou normal, é bastante usada para representar a dispersão de valores possíveis de uma quantidade. Uma justificativa para isto é o *Teorema Central do Limite*, que em sua versão mais geral também é denominado *Teorema de Lindeberg-Feller*, [16, por exemplo].

Numa linguagem bastante simplificada e adaptada para o problema em questão, o *Teorema Central do Limite* estabelece que se um erro η é a soma de um número muito grande de pequenos erros que têm distribuições aleatórias quaisquer com variâncias finitas, então a distribuição resultante para η é uma distribuição gaussiana. Além disso, a variância da distribuição resultante é a soma das variâncias das distribuições envolvidas na convolução. Uma discussão mais detalhada é apresentada em [16].

Na prática, a convolução de poucas distribuições de erros comparáveis converge muito rapidamente para uma distribuição gaussiana. Por exemplo, a combinação de 3 erros comparáveis, seguindo distribuições retangulares, resulta numa distribuição muito próxima da gaussiana [12, 17, por exemplo].

Como pode ser visto do Teorema Central do Limite, a distribuição normal deve descrever bem a “variabilidade” dos valores possíveis de X , quando ficar claro que esta resulta de várias contribuições comparáveis e os limites a_- e a_+ não sejam totalmente confiáveis. A melhor estimativa de X é dada pela Eq. (4). Entretanto, sempre existe a dificuldade de atribuir um nível de confiança P ao intervalo $[a_-, a_+]$.

Uma vez que o valor de P seja estabelecido conforme a informação disponível, a incerteza padrão u pode ser estimada:

$$u = \frac{a_+ - a_-}{2k} = \frac{a}{k}, \quad (6)$$

onde k é obtido da função de distribuição gaussiana $F(ku)$: $100 \times F(ku = a) = P(\%)$. Deve ser observado que $F(ku)$ é a integral da função de densidade de probabilidade entre os limites $(x - ku)$ e $(x + ku)$. Valores de P e k são dados na Fig. 1.

c. Distribuição qualquer

Uma relação que pode ser útil na avaliação de incerteza tipo B é a *inequação de Bienaymé-Chebyshev* [18].

Para uma função de densidade de probabilidade *qualquer* com média x e variância finita u^2 , a probabilidade de ocorrer $|X - x| \leq ku$ é

$$P \geq \left(1 - \frac{1}{k^2}\right) \quad (|X - x| \leq ku). \quad (7)$$

Assim, para um dado valor de k , a inequação de Bienaymé-Chebyshev garante um nível de confiança mínimo, *qualquer que seja a distribuição de probabilidades*. (Fig. 1).

Inversamente, dado um nível de confiança P , a inequação indica o valor máximo para k (Fig. 1). Isto pode auxiliar na escolha de k para estimar a incerteza tipo B ($u = a/k$).

Exemplo simples

Como exemplo simples, pode-se considerar um instrumento para o qual o fabricante estabelece um “limite de erro” a para o valor indicado. Deve ser escolhida uma distribuição para os valores possíveis da quantidade, a partir de dados fornecidos pelo fabricante e a partir de toda informação e conhecimento a respeito do instrumento e da calibração do mesmo.

Para instrumentos analógicos, existia uma regra antiga, segundo a qual “erro” (de calibração) nunca deveria ser maior que a menor divisão. Assim, o “limite de erro” a é igual a menor divisão. Admitindo distribuição gaussiana e 95% de confiança para o limite de erro ($k = 2$), obtém-se a famosa regra da “metade da menor divisão ($u = a/2$). Se fosse admitida uma distribuição retangular, o resultado não seria muito diferente ($u = a/1,73$).

Entretanto, deve ser observado que esta regra, atualmente não é mais válida, em geral. Por exemplo, para uma régua metálica de boa qualidade, o limite de erro é seguramente menor que a menor divisão. Por outro lado, existem paquímetros (de boa qualidade) com nônio de 50 divisões, para os quais o limite de erro de calibração é maior que a menor divisão ($20 \mu m$). Além disso, instrumentos digitais, têm erro de calibração bem maior que a menor leitura, em geral.

V Incerteza combinada

Em geral, o mensurando Y é admitido como sendo dado por

$$Y = f(X_1, X_2, \dots, X_N), \quad (8)$$

onde X_i são valores verdadeiros das quantidades de entrada. Se esses valores são estimados por valores x_i , obtém-se $y = f(x_1, x_2, \dots, x_N)$ como estimativa para o mensurando.

A combinação de incertezas deve ser feita pela fórmula usual de propagação de incertezas [4, 12, 13,

por exemplo]. Na ausência de correlação entre as quantidades de entrada, a incerteza combinada é dada por

$$u_c^2(y) = \sum_{i=1}^N u_i^2(y), \quad (9)$$

onde

$$u_i(y) = |c_i| u_i(x_i), \quad \text{sendo } c_i = \left[\frac{\partial f}{\partial X_i} \right]_{x_1, x_2, \dots, x_N} \quad (10)$$

O coeficientes c_i são chamados *coeficientes de sensibilidade* para a variável X_i .

Se existe correlação entre as quantidades de entrada, deve-se usar a fórmula de propagação de incertezas completa, incluindo covariâncias [4, 12, 13, por exemplo].

A fórmula de propagação de incertezas pode ser usada para incerteza padrão ou para incerteza expandida. Se as incertezas padrões são multiplicadas pelo mesmo fator de abrangência k , resulta a incerteza padrão combinada multiplicada pelo mesmo fator de abrangência. Isto é, obtém-se a *incerteza combinada expandida*.

Exemplo simples

Um exemplo simples e comum é a medição com um instrumento na qual, em primeira aproximação, podem ser considerados apenas erros aleatórios de medição e o erro de calibração do instrumento. Isto é, o mensurando Y é dado por

$$Y = X + C \quad (11)$$

onde X é a quantidade de entrada, obtida diretamente na medição e C é uma correção desconhecida, devida a erro de calibração. As quantidades X e C são estimadas como x e c , respectivamente, e $y = (x + c)$. A incerteza padrão combinada é obtida a partir de (11) e (9):

$$u = \sqrt{u_x^2 + u_c^2}. \quad (12)$$

Por hipótese, x é obtido a partir de n medições e só tem erros aleatórios. Assim, a incerteza em x é de tipo A e obtida pela Eq. (3) ($u_x = u_A$). Por hipótese, a incerteza em c é apenas devida a erro de calibração e é de tipo B ($u_c = u_B$). Assim, resulta a incerteza padrão combinada para y :

$$u = \sqrt{u_A^2 + u_B^2}. \quad (13)$$

Em geral, quando não se tem nenhuma estimativa para a correção de calibração do instrumento, $c = 0$ e $y = x$. A incerteza $u_B = u_c$ deve ser estimada, usando todo conhecimento e informação disponível sobre o instrumento. Deve-se avaliar o limite de erro de calibração (a), respectivo nível de confiança e adotar uma distribuição para valores possíveis do mensurando C . Por exemplo, se a distribuição é retangular e o nível de confiança para o intervalo $2a$ é 100%, resulta $u_B = a/\sqrt{3}$.

Em certos casos, a correção C é desconhecida, mas c pode ser diferente de zero e deve ser estimada. Um exemplo é o chamado “erro cosseno”. Por exemplo, isto ocorre se

$$C = (1 - \cos \alpha) \quad (14)$$

e $\alpha = 0$, mas tem um erro de calibração. Mesmo que a distribuição de valores possíveis de α seja centrada em zero, os valores possíveis de C não se distribuem em torno de zero e o valor médio correspondente é positivo. Erros do tipo “cosseno” são discutidos em [4] e [17], por exemplo.

VI Número de graus de liberdade

Na avaliação da incerteza padrão tipo A, o número de graus de liberdade é $\nu_A = n - m$, onde n é o número de observações independentes e m é o número de quantidades determinadas. No caso mais simples, a quantidade é a média e o número de graus de liberdade é $\nu_A = n - 1$.

A incerteza padrão estimada pela Eq. (3) tem uma incerteza padrão dada por [4, 13, por exemplo]

$$\Delta u_A = \frac{u_A}{\sqrt{2\nu_A}}. \quad (15)$$

Esta relação mostra que a “incerteza na incerteza” está diretamente relacionada com o número de graus de liberdade ν . A Eq. (15) permite associar um número de graus de liberdade ν_B à incerteza padrão tipo B (u_B). Se Δu_B é a incerteza padrão em u_B :

$$\nu_B \approx \frac{1}{2} \left[\frac{u_B}{\Delta u_B} \right]^2. \quad (16)$$

Assim, se a “incerteza” na incerteza u_B é estimada, obtém-se uma estimativa para ν_B . Por exemplo, se a incerteza em u_B é estimada em 25%, resulta $\nu_B = 8$.

O número de graus de liberdade efetivo ν_{ef} para a incerteza combinada também pode ser estimado. Se cada incerteza $u_i(y)$ na Eq. (9) tem uma incerteza Δu_i , a “incerteza” na incerteza combinada é obtida pela própria fórmula de propagação de incertezas. Substituindo as incertezas Δu_i em termos dos respectivos graus de liberdade, obtém-se a fórmula de Welch-Satterthwaite [4,6]:

$$\frac{u_c^4}{\nu_{ef}} = \sum_{i=1}^N \frac{u_i^4}{\nu_i}. \quad (17)$$

Para número de graus de liberdade da ordem de 10, a Eq. (15) mostra que a “incerteza” na incerteza é da ordem de 22%. Isto significa que a incerteza de tipo A é determinada com “incerteza” relativamente grande. Assim, a determinação de incerteza tipo B pode ser tão boa ou até melhor que a de tipo A, quando o número de

graus de liberdade estatístico não é muito grande. Esta observação é importante porque métodos estatísticos convencionais (clássicos) são inquestionados e aceitos como de alta qualidade, enquanto que procedimentos tais como os indicados na Seção IV para determinar a incerteza de tipo B são criticados.

VII Incerteza expandida

Para distribuição gaussiana de erros, a incerteza padrão u define um intervalo de confiança de aproximadamente 68% para os valores possíveis do mensurando, no caso de número de graus de liberdade razoavelmente grande. Neste caso, as incertezas expandidas ku , com valores de k de 2 a 3, definem intervalos com confiança de 95 a 99%, aproximadamente (Eq. (2) e Fig. 1).

É importante observar que para $k = 1$, o nível de confiança varia de 62,8% a 68,3% conforme ν varia de 4 até ∞ . Entretanto, no caso $k = 3$, para a mesma variação de ν , o nível de confiança varia de 95,9% a 99,7%. Em geral, esta diferença tem implicações práticas muito mais graves que a diferença para $k = 1$.

Além disso, os níveis de confiança para incerteza expandida ($k \geq 2$) são muito dependentes da forma da distribuição. A Fig. 1 mostra que para $k = 1,5$ os níveis de confiança para as distribuições gaussiana e retangular são próximos. Para uma distribuição triangular, o valor também é próximo. Até mesmo o valor mínimo dado pela inequação de Bienaymé-Chebyshev para uma distribuição qualquer, não é muito diferente. Entretanto, para $k = 2$ os valores são completamente discrepantes. Estas considerações mostram as dificuldades e inconveniências de se indicar incerteza expandida para grandes valores do fator de abrangência k . Como regra geral, deve-se evitar indicar a incerteza expandida e, quando isto for necessário, o fator de abrangência nunca deve ser maior que 2 ou 3. A atribuição de um nível de confiança útil exige que a distribuição de probabilidades seja razoavelmente bem determinada e o número de graus de liberdade razoavelmente grande.

O Teorema Central do Limite estabelece um atenuante para as dificuldades de interpretar a incerteza combinada e atribuir intervalos de confiança, quando incertezas tipo B são envolvidas. Quando a incerteza combinada resulta de várias contribuições comparáveis, mesmo com diferentes distribuições de erros, o Teorema Central do Limite indica que a convolução dos diferentes erros deve se aproximar de uma distribuição gaussiana. Uma discussão mais detalhada é apresentada na Referência 16.

VIII Formas de apresentar a incerteza

Em princípio, a incerteza final no resultado da medição deve ser apresentada de forma que permita ao leitor refazer completamente os cálculos, quando necessário. Por exemplo, valores mais acurados para constantes usadas nos cálculos, podem se tornar acessíveis. Um exemplo clássico desta situação ocorreu na experiência de Millikan. O valor da carga do elétron foi corrigido 16 anos mais tarde, usando um valor mais acurado para a viscosidade do ar.

Num resultado final, a incerteza deve ser qualificada, indicando explicitamente se é a incerteza padrão ou uma incerteza expandida com um dado fator de abrangência. Além disso, devem ser apresentados o número de graus de liberdade, covariâncias quando for o caso, descrição detalhada do método de cálculo e listagem completa de todas as quantidades de entrada, importadas ou determinadas experimentalmente, juntamente com as respectivas incertezas tipo A e tipo B. Também devem ser indicadas explicitamente as incertezas tipo A e tipo B do resultado final e respectivos graus de liberdade.

Algarismos significativos

Não existem regras bem estabelecidas para o número de algarismos a ser indicado na incerteza. Entretanto, é consenso que, *não mais que 2 algarismos significativos* sejam indicados na incerteza padrão, exceto em casos muito excepcionais. A justificativa para isto é que a “incerteza” na incerteza nunca é muito pequena, exceto em casos excepcionais em que o número de graus de liberdade seja excepcionalmente grande.

Textos bastante expressivos tais como o GUM ou a *listagem de valores do CODATA para as constantes físicas fundamentais* [19] usam sempre 2 algarismos significativos para a incerteza. Parece que o mais razoável é adotar esta regra geral, embora não seja muito consistente. Por exemplo, se a incerteza num diâmetro é conhecida apenas como sendo $u_d = 0,18 m$, não é muito consistente dizer que a incerteza no raio é $u_r = 0,090 m$. Seria mais consistente escrever $u_d = 0,09 m$. Por outro lado, é inaceitável arredondar $u = 0,14 m$ para $u = 0,1 m$ ou arredondar $u = 0,16 m$ para $u = 0,2 m$. Em [12], é sugerida a regra de usar *obrigatoriamente* 2 algarismos significativos quando o primeiro algarismo na incerteza padrão é 1, 2 ou 3.

Independente de se usar um ou dois algarismos para a incerteza num resultado final, os cálculos intermediários, devem sempre ser feitos com mais algarismos, preferivelmente 3, quando disponíveis, para evitar erros de arredondamento.

Num resultado final, a quantidade deve ser sempre indicada com os algarismos consistentes com a incerteza padrão. Em cálculos intermediários, devem ser usados

mais algarismos, quando disponíveis.

Formas compactas

É evidente a necessidade de formas compactas para indicar a incerteza e não há muito consenso sobre isto. A Tabela 1 mostra 3 opções para representar a incerteza padrão.

Tabela 1. *Formas compactas para indicar a incerteza padrão.*

1	$(12,435 \pm 0,067) mm$
2	$12,435 (67) mm$
3	$12,435 (0,067) mm$

A forma 1 (\pm) é a tradicionalmente usada em física experimental e estatística para indicar desvio padrão ou incerteza dada nesta forma (u). Isto é, define um intervalo de largura $2u$ com confiança de 68%, aproximadamente, no caso de distribuição gaussiana.

O problema é que a mesma forma também é usada em outras áreas, por exemplo, em especificações técnicas, para intervalo $2u$ com confiança bem maior, próxima ou igual a 100%. De qualquer modo, quando esta forma é utilizada, deve ser explicitamente mencionado em alguma parte do texto se indica incerteza padrão ou incerteza expandida.

A forma 2 é inconveniente no manuseio de dados, porque é sempre a incerteza que define os algarismos que devem ser considerados em cada etapa. Além disso tem a desvantagem de não apresentar nenhuma redundância. Do ponto de vista didático, isto é inconveniente, porque o aluno iniciante não sabe ainda manusear com segurança os algarismos significativos.

A forma 3 elimina, em grande parte, todos os problemas anteriores. Por isso, parece ser a melhor. Entretanto, esta forma tem sido muito pouco usada, infelizmente.

IX Comentários e conclusões

Não é objetivo deste artigo detalhar procedimentos para avaliação de incertezas, mas apenas resumir os princípios gerais contidos no GUM [4,6]. Os procedimentos para avaliação de incerteza tipo A são métodos estatísticos bem conhecidos. Quanto à avaliação da incerteza tipo B, não existe nenhum procedimento detalhado, exceto que o avaliador deve escolher uma distribuição de probabilidades para valores possíveis da grandeza, sendo que esta escolha deve incorporar toda informação disponível sobre o mensurando e a medição. Isto corresponde ao procedimento inicial da *estatística bayesiana*.

Outro objetivo deste artigo é divulgar a nomenclatura estabelecida no GUM (em inglês). Entretanto, deve ser observado que o GUM é um documento que tem tido ampla aceitação por metrologistas e instituições nacionais ligadas à Metrologia. Embora seja endossado pela IUPAP, ainda existem controvérsias e questões a serem resolvidas. Além disso, o documento é relativamente recente e, portanto deve ser submetido a revisões e incorporar sugestões. De qualquer modo, é pouco provável que a nomenclatura e os princípios básicos estabelecidos no GUM, com exceção de detalhes, não venham a ser assimilados por físicos experimentais e outros profissionais de ciências exatas e tecnologia. Por isso, parece importante que esta nomenclatura seja introduzida em disciplinas experimentais de cursos de graduação nestas áreas.

Agradecimentos ao Prof. Giorgio Moscati, pelo apoio e incentivo, manifestados em diversas ocasiões, com relação ao estudo deste tema.

References

- [1] Colclough, A.R., 1987, *Two Theories of Experimental Error*, J. Research National Bureau of Standards, Vol. 92/3, pp.167.
- [2] Weise, K. and Woger, W., 1992, *A Bayesian theory of measurement uncertainty*, Meas. Sci. Technol. **3**, pp.1.
- [3] Petley, B.W., 1985, *The Fundamental Physical Constants and the Frontier of Measurement*, Adam Hilger Ltd, London.
- [4] BIPM/IEC/IFCC/ISO/IUPAC/IUPAP/OIML, 1993, *Guide to the Expression of Uncertainty in Measurement*, (corrected and reprinted, 1995) International Organization for Standardization (ISO), Geneva.
- [5] Mathiesen, O., 1997, *Evaluation of Uncertainty in Measurement*, Lecture in the Advanced School of Metrology (INMETRO), Angra dos Reis, March 1997.
- [6] BIPM/IEC/IFCC/ISO/IUPAC/IUPAP/OIML, 1998, *Guia para Expressão da Incerteza de Medição, 2ª Versão Brasileira*, INMETRO, Rio de Janeiro.
- [7] EAL-R2, *Expression of the Uncertainty of Measurement in Calibration*, European Cooperation for Accreditation Laboratories, Utrecht, The Netherlands (1997).
- [8] NIST Technical Note 1297, 1994, *Guidelines for Evaluation and Expressing Uncertainty of NIST Measurement Results*, U.S. Government Printing Office (Washington, USA).
- [9] BIPM/IEC/IFCC/ISO/IUPAC/IUPAP/OIML, 1993, *International Vocabulary of Basic and General Terms in Metrology, 2nd Edition*, International Organization for Standardization, Geneva.
- [10] INMETRO, 1995, *Vocabulário Internacional de Termos Fundamentais e Gerais em Metrologia*, INMETRO, Duque de Caxias (RJ), (Versão traduzida da Referência anterior).
- [11] Collé, R. and Karp, P., 1987, *Measurement Uncertainties: Report of an International Working Group Meeting*, J. Research National Bureau of Standards, Vol. 92/3, pp.243.
- [12] Vuolo, J. H., 1996, *Fundamentos da Teoria dos Erros, 2ª Edição*, Editora Edgard Blucher, São Paulo.
- [13] Helene, O.A.M. e Vanin, V.R., 1981, *Tratamento Estatístico de Dados em Física Experimental*, Ed. Edgard Blücher Ltda, São Paulo.
- [14] Vanin, V.R. e Gouffon, P., 1996, *Tópicos Avançados em Tratamento Estatístico de Dados Experimentais*, Edição Preliminar (livro), Instituto de Física da USP, São Paulo.
- [15] Harry, F. M. and Waller, R. A., 1994, *Bayesian Methods*, in *Statistical Methods for Physical Science*, J. L. Stanford and S. B. Vardeman, Eds., Academic Press, New York.
- [16] Woodroffe, M., 1975, *Probability with Applications*, McGraw-Hill Kogakusha, Tokyo.
- [17] Cohen, E. R., 1992, *Uncertainty and Error in Physical Measurements* in Proceedings of the International School of Physics "Enrico Fermi", Course CX, North-Holland, Amsterdam.
- [18] Estler, W. T., 1997, *A Distribution-Independent Bound on the Level of Confidence in the Result of a Measurement*, J. Res. Natl. Inst. Stand. Technol. Vol. 102/5, pp. 587.
- [19] Cohen, E. R., and Taylor, B. N., 1987, *The 1986 CODATA Recommended Values of the Fundamental Physical Constants*, J. Res. National Bureau of Standards, Vol. 92/2, pp.85.