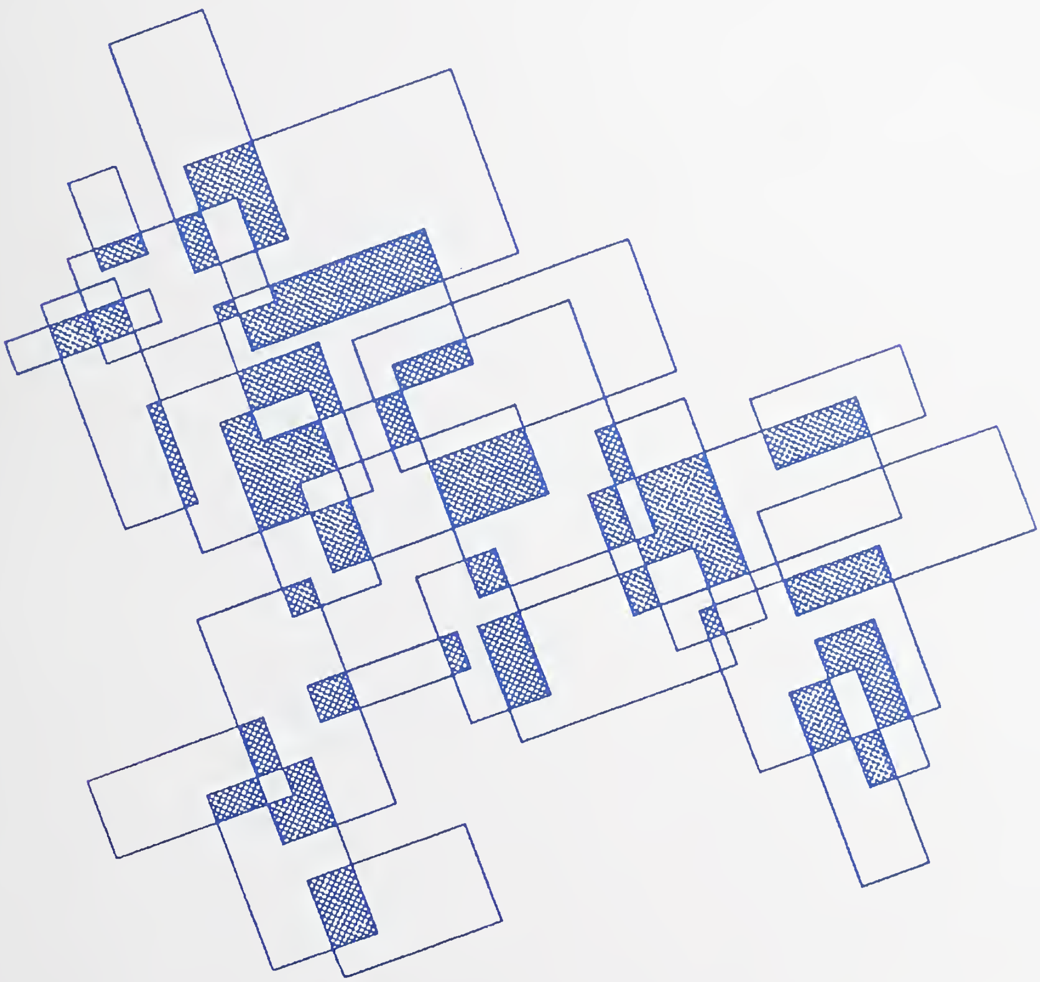
NIST Special Publication 500-244

Information Technology:

An Interpretation of the Guide to the Expression of Uncertainty in Measurement

Raghu Kacker



NIST
National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

The National Institute of Standards and Technology was established in 1988 by Congress to "assist industry in the development of technology . . . needed to improve product quality, to modernize manufacturing processes, to ensure product reliability . . . and to facilitate rapid commercialization . . . of products based on new scientific discoveries."

NIST, originally founded as the National Bureau of Standards in 1901, works to strengthen U.S. industry's competitiveness; advance science and engineering; and improve public health, safety, and the environment. One of the agency's basic functions is to develop, maintain, and retain custody of the national standards of measurement, and provide the means and methods for comparing standards used in science, engineering, manufacturing, commerce, industry, and education with the standards adopted or recognized by the Federal Government.

As an agency of the U.S. Commerce Department's Technology Administration, NIST conducts basic and applied research in the physical sciences and engineering, and develops measurement techniques, test methods, standards, and related services. The Institute does generic and precompetitive work on new and advanced technologies. NIST's research facilities are located at Gaithersburg, MD 20899, and at Boulder, CO 80303. Major technical operating units and their principal activities are listed below. For more information contact the Publications and Program Inquiries Desk, 301-975-3058.

Office of the Director
- National Quality Program
- International and Academic Affairs

Technology Services
- Standards Services
- Technology Partnerships
- Measurement Services
- Information Services

Advanced Technology Program
- Economic Assessment
- Information Technology and Applications
- Chemistry and Life Sciences
- Materials and Manufacturing Technology
- Electronics and Photonics Technology

Manufacturing Extension Partnership Program
- Regional Programs
- National Programs
- Program Development

Electronics and Electrical Engineering Laboratory
- Microelectronics
- Law Enforcement Standards
- Electricity
- Semiconductor Electronics
- Radio-Frequency Technology[1]
- Electromagnetic Technology[1]
- Optoelectronics[1]

Materials Science and Engineering Laboratory
- Intelligent Processing of Materials
- Ceramics
- Materials Reliability[1]
- Polymers
- Metallurgy
- NIST Center for Neutron Research

Chemical Science and Technology Laboratory
- Biotechnology
- Physical and Chemical Properties[2]
- Analytical Chemistry
- Process Measurements
- Surface and Microanalysis Science

Physics Laboratory
- Electron and Optical Physics
- Atomic Physics
- Optical Technology
- Ionizing Radiation
- Time and Frequency[1]
- Quantum Physics[1]

Manufacturing Engineering Laboratory
- Precision Engineering
- Automated Production Technology
- Intelligent Systems
- Fabrication Technology
- Manufacturing Systems Integration

Building and Fire Research Laboratory
- Applied Economics
- Structures
- Building Materials
- Building Environment
- Fire Safety Engineering
- Fire Science

Information Technology Laboratory
- Mathematical and Computational Sciences[2]
- Advanced Network Technologies
- Computer Security
- Information Access and User Interfaces
- High Performance Systems and Services
- Distributed Computing and Information Services
- Software Diagnostics and Conformance Testing
- Statistical Engineering

[1] At Boulder, CO 80303.
[2] Some elements at Boulder, CO.

NIST Special Publication 500-244

Information Technology:

An Interpretation of the Guide to the Expression of Uncertainty in Measurement

Raghu Kacker

Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20890-8980

May 2000

U.S. Department of Commerce
William M. Daley, Secretary

Technology Administration
Dr. Cheryl L. Shavers, Under Secretary of Commerce for Technology

National Institute of Standards and Technology
Raymond G. Kammer, Director

Reports on Information Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) stimulates U.S. economic growth and industrial competitiveness through technical leadership and collaborative research in critical infrastructure technology, including tests, test methods, reference data, and forward-looking standards, to advance the development and productive use of information technology. To overcome barriers to usability, scalability, interoperability, and security in information systems and networks, ITL programs focus on a broad range of networking, security, and advanced information technologies, as well as the mathematical, statistical, and computational sciences. This Special Publication 500-series reports on ITL's research in tests and test methods for information technology, and its collaborative activities with industry, government, and academic organizations.

National Institute of Standards and Technology
Special Publication 500-244
Natl. Inst. Stand. Technol.
Spec. Publ. 500-244
18 pages (May 2000)
CODEN: NSPUE2

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 2000

For sale by the Superintendent of Documents
U.S. Government Printing Office
Washington, DC 20402-9325

AN INTERPRETATION OF THE GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT

Raghu Kacker
National Institute of Standards and Technology
100 Bureau Drive, Stop 8980
Gaithersburg, MD 20899-8980
Raghu.kacker@nist.gov
Phone 301-975-2109

ABSTRACT

The Guide to the Expression of Uncertainty in Measurement (GUM) is intended for all scientific and technological measurements in science, engineering, commerce, industry, and regulation. So the GUM must have a clear interpretation. But it mixes up concepts from frequentist and Bayesian statistics in ambiguous ways. Therefore, as presented, the GUM is not clear and liable to be applied in more than one way, leading to more than one way of expressing uncertainty in measurement. This paper attempts to present a clear and coherent interpretation of the GUM and proposes a simple and widely applicable approach to construct expanded uncertainty intervals. Our hope is that the clarifications and the viewpoints presented here will promote a more consistent use of the GUM and facilitate its application to situations not explicitly covered in the original document.

Key Words: Bayesian Analysis, Expanded Uncertainty, Frequentist Statistics, Metrology, Statistics, Uncertainty

1. INTRODUCTION

The Guide to the Expression of Uncertainty in Measurement [1], commonly referred to as the GUM, is promulgating a standardized approach for evaluating and expressing uncertainty in measurement, and its impact is growing. In addition to providing a standardized approach for expressing uncertainty, the GUM has provided a practical approach for incorporating scientific judgment with the results of statistical analyses of measurement data. Both sources of knowledge are generally needed to evaluate uncertainty economically. Another advantage of the GUM is that the output from one stage of measurement may be used as an input to a subsequent stage. Thus the GUM has provided a practical way to partition a complex measurement problem into smaller, more manageable components and to inter-link a hierarchy of measurements. The latter benefit is useful in establishing the traceability of commercial and scientific measurements to the national and international standards.

The GUM is intended for all scientific and technological measurements in science, engineering, commerce, industry, and regulation. The GUM is now an "American National Standard for Expressing Uncertainty [2]." So the GUM must have an unambiguous interpretation. But even some of the most basic definitions of the GUM are not exactly clear. Consider the meaning of uncertainty. The GUM (Section 2.2.3) defines uncertainty of measurement as a "*parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.*" The GUM (Section 2.3.1) defines standard uncertainty as "*uncertainty of the result of a measurement expressed as a standard deviation.*" The GUM (Section 2.3.5) defines expanded uncertainty as a "*quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand.*"

The definition of uncertainty may be interpreted in the following two ways. Frequentist viewpoint: *uncertainty* is about the result of measurement assuming that the value of measurand is an unknown constant -- traditionally called the true value. Bayesian viewpoint: *uncertainty* is about the value of measurand, treated as a random variable, given that the result of measurement, the available measurement data, and scientific judgment are known quantities. The phrase "uncertainty of the result of measurement" in the definition of standard uncertainty supports the frequentist viewpoint. The frequentist viewpoint leads to the traditional concepts of true value and error. But the GUM (Annex D) discourages the use of these traditional concepts. In this sense, the GUM supports the Bayesian viewpoint. But then the GUM (Annex G) motivates the use of a Student's t-distribution from the viewpoint of frequentist sampling theory to assign the coverage probability to an interval about the result of measurement defined by expanded uncertainty. The frequentist viewpoint leads to the concept of confidence intervals. But the GUM (Section 6.2.2) states that the word "confidence" is not used to modify the word "interval" when referring to the interval defined by expanded uncertainty. This is what we mean when we say that the GUM mixes up concepts from frequentist and Bayesian statistics in ambiguous ways.

The consequences of this mix-up include the following. First, the GUM is liable to be applied in more than one way, leading to more than one way of expressing uncertainty in measurement. For example, many users believe that the frequentist confidence intervals, where the result of measurement and the standard uncertainty are treated as random variables, agree with the GUM as do the intervals defined by expanded uncertainty, where the value of measurand is a random variable. Second, the meaning of "being GUM compliant" is ambiguous. Third, the user would not be sure how to apply the GUM to situations not explicitly covered in the original document. In Section 2, we attempt to present a clear and coherent interpretation of the GUM and propose a simple and widely applicable approach for constructing expanded uncertainty intervals. A summary is given in Section 3, and a number of practical comments and recommendations are given in Section 4.

2. AN INTERPRETATION OF THE GUM

The *measurand* is a particular quantity subject to measurement. The object of measurement is to determine (assess) the *value of the measurand* (the GUM, Section 3.1.1). In some cases, the measurand is defined by a particular (standard) method of measurement. The GUM applies to measurands that are characterized by a scalar value.

Additional guidance is needed for measurands that are characterized by a vector or a function defined over some domain. Work is in progress to extend the GUM in this direction [3].

Since no measurement is perfect (except when counting the elements of a small set of discrete items), no measured quantity is known exactly (see, the GUM, Annex D). That is, the state of knowledge about the value of a measured quantity is uncertain. There are two well-established and distinct ways of defining and quantifying uncertainty: frequentist and Bayesian. The frequentist sampling theory assumes that the value of measurand is an unknown constant and the result of measurement is a random variable. A Bayesian approach treats the value of measurand as a random variable with a probability distribution representing the state of knowledge given that the result of measurement is a known quantity. The results of statistical analyses based on frequentist sampling theory are usually simpler and, for historical reasons, familiar to metrologists. The GUM was motivated in part to incorporate scientific judgment with the results of frequentist statistical analyses (see, the GUM, Section 0.7). So the GUM has mixed up frequentist and Bayesian concepts and introduced a new terminology. We will show that the GUM is clear and coherent if we adopt a Bayesian line of thinking. That is treat all quantities involved in measurement as random variables with probability distributions representing the states of knowledge, and treat the results of frequentist statistical analyses as approximations to the corresponding results of Bayesian analyses. Another advantage of the proposed interpretation of the GUM is that it affords a very simple approach for constructing expanded uncertainty intervals.

The GUM is mainly concerned with the expected values and the standard deviations of the random variables involved in measurement rather than with the fully characterized probability distributions. The reason, we believe, is that it is easier to estimate or assess the expected value and the standard deviation of a random variable than judge the complete probability distribution. The expected value and the standard deviation of a random variable are said to characterize its probability distribution. Since Bayesian methods work with the probability distributions of the involved variables, the GUM is not intended to be a completely Bayesian approach in our view. Another researcher has shown that the recommendations of the GUM can be regarded as approximate solutions to certain frequentist and Bayesian inference problems [4].

The GUM is based on the concept of measurement equation. A *measurement equation* is a functional relationship that expresses the value of measurand as a function of all those variables that affect its assessment. The expected value and the standard deviation of an input variable to the measurement equation are evaluated from statistical analysis of measurement data and/or by scientific judgment. The method of evaluation is referred to as Type A evaluation when measurement data are used, and Type B evaluation when scientific judgment is used. These two modes of evaluation are not necessarily mutually exclusive [3]. The evaluated expected values and the standard deviations of the input variables are then combined through the measurement equation to obtain the expected value and the standard deviation of the value of measurand. The expected value is taken as the estimated value of measurand and the standard deviation as the combined standard uncertainty concerning the value of measurand.

2.1 Type A and Type B Evaluations of Standard Uncertainty

The statistical methods employed for Type A evaluation may be either Bayesian or frequentist. But for simplicity, Type A evaluations are usually frequentist estimates. We will briefly describe the two approaches. Type A evaluations from Bayesian analyses and Type B evaluations from scientific judgment are mathematically compatible inputs to the measurement equation because both treat the input quantities as random variables. But Type A evaluations from frequentist analyses are not mathematically compatible with Type B evaluations, because the frequentist methods treat the input quantities as unknown constants. We will illustrate that, in the practical cases of interest, the frequentist estimates may be regarded as approximations to the corresponding results from Bayesian analyses based on non-informative prior distributions. Therefore, it is legitimate to treat frequentist estimates and Type B assessments as mathematically compatible inputs to the measurement equation.

Suppose the value of the quantity of interest is X. A Bayesian analysis starts with a prior probability distribution representing the state of knowledge about X before measurement. The expected value, the variance, and the standard deviation (square root of variance) of the prior distribution are called prior expected value, prior variance, and prior standard deviation, and denoted by E(X), V(X), and SD(X) respectively. The relationship between the value of X and the statistical measurement data is expressed by a "likelihood function." Generally, both Bayesians and frequentists agree on the likelihood function. The prior distribution and the likelihood function are then combined by Bayes theorem [5] to obtain a posterior distribution representing the state of knowledge about X after measurement. The expected value, the variance, and the standard deviation of the posterior distribution are called posterior expected value, posterior variance, and posterior standard deviation, and denoted by E(X | data), V(X | data), and SD(X | data) respectively. This notation indicates that the posterior distribution is conditional on the data. The posterior distribution can be used as a prior distribution in a subsequent measurement of the same quantity, and the process can be repeated any number of times. The posterior expected value E(X | data) is taken as an estimate of X, and SD(X | data) is taken as a measure of the uncertainty concerning X after measurement.

In a frequentist analysis, the value of the quantity of interest X is treated as an unknown constant -- traditionally called the true value. The output of a frequentist statistical analysis is an estimate of X and an estimated standard deviation of the estimate. Consider the simple case where X is estimated from a sample (set) of n measurements that are assumed to be independent and identically normally distributed random variables with expected value X and some variance σ^2. Let x and s^2 denote the sample mean and the sample variance of the n measurements. Then x, s^2, and s are the estimates of X, σ^2, and σ respectively. The probability distribution of x, called a sampling distribution, is also normal but with expected value X and variance σ^2/n. The ratio $s/\sqrt{n}$ is an estimate $\sigma/\sqrt{n}$. The standard deviation $\sigma/\sqrt{n}$, called population standard deviation of the mean, characterizes the tightness of the sampling distribution of x about E(x) = X. So $s/\sqrt{n}$ is an estimate of the tightness of the sampling distribution of x about X. Thus $s/\sqrt{n}$, called sample standard deviation of the mean, is a measure of the doubt about x as an estimate of X.

The frequentist estimates x and $s/\sqrt{n}$ may be viewed as approximations to the Bayesian posterior expected value E(X | data) and the standard deviation SD(X | data) respectively based on a class on prior distributions called non-informative prior distributions [5]. A non-informative prior distribution represents the situation that relatively little is known a priori about the value X of the quantity of interest in advance of measurement. It can be shown that the Bayesian posterior expected value and variance based on non-informative prior distributions are approximately equal to the corresponding estimates from frequentist sampling theory, provided the number of independent measurements on which the estimates are based is not too small [5]. This assertion is illustrated in the Appendix.

Note: In the case of n independent and identically normally distributed measurements with mean x and standard deviation s, the Bayesian posterior distribution of $(X - x)/(s/\sqrt{n})$, based on a pair of common non-informative prior distributions, is the t-distribution with (n -1) degrees of freedom [5]. Thus $SD(X \mid data) = \sqrt{[(n-1)/(n-3)]} \times (s/\sqrt{n})$, which is defined only when n is four or more. Therefore, at least four independent measurements are required to claim that the frequentist estimate $s/\sqrt{n}$ approximates the Bayesian posterior standard deviation SD(X | data).

Frequently, the data structures and the statistical models underlying the frequentist analyses are more complicated than the simple example of a series of independent and identically normally distributed measurements discussed above. The outputs of the data analysis are, nonetheless, an estimate of a parameter and an estimated standard deviation of the estimate. Even with more complicated analyses, in the practical cases of interest, the frequentist estimates may be regarded as approximations of the Bayesian posterior expected value and standard deviation corresponding to some (proper or improper) non-informative prior distributions [5]. This relationship between the frequentist and the Bayesian results enables us to interpret the GUM from a Bayesian line of thinking and still employ frequentist statistics for Type A evaluations.

In a Type B evaluation, scientific judgment is expressed in terms of a fully characterized probability distribution for X. Thus the expected value and the variance of X are specified values. The GUM treats Type B evaluations of the expected value and the variance in exactly the same way as it treats Type A evaluations. One should not belabor the distinction between the two modes of evaluation [3]. We need a general notation for the expected value, the variance, and the standard deviation of an input variable regardless of the mode of evaluation. We will denote the current state of knowledge about the expected value, the variance, and the standard deviation of an input variable X based on all available information as E(X | .), V(X | .), and SD(X | .) respectively. The expected value E(X | .), denoted by x, is taken as the estimated value of X and the standard deviation SD(X | .), denoted by u(x), is referred to as the standard uncertainty concerning X. The variance V(X | .) is equal to $u^2(x)$.

2.2 Measurement Equation

Let Y denote the value of measurand, treated as a random variable with a probability distribution representing the state of knowledge. In the GUM paradigm, the primary object of measurement is to evaluate the expected value and the standard deviation of the value of measurand Y from all available measurement data and scientific judgment. Following the notation of Subsection 2.1, we will denote the expected value, the variance, and the standard deviation of Y as E(Y | .), V(Y | .), and SD(Y | .) respectively. The GUM is concerned with applications where E(Y | .) and SD(Y | .) are determined from the expected values and the standard deviations of some number N of input variables $X_1, X_2, \ldots, X_N$ through a functional relationship, denoted by f, and called the *measurement equation*:

$$Y = f(X_1, X_2, \ldots, X_N). \tag{1}$$

In a broad sense, the measurement equation represents the procedure used to determine the value of measurand. Some of the input variables X_i may themselves be viewed as measurands and functions of additional input variables. Therefore, the measurement equation provides a practical way to partition a complex measurement problem into smaller more manageable components and to inter-link a hierarchy of measurements. In some cases, the function f is expressed as a system of equations. In some other cases, the function f may be the identity function Y = X or may be expressed as $Y = X + C_1 + C_2 + \ldots + C_M$, where $C_1, C_2, \ldots, C_M$, are correction for systematic (non-random) effects. The function f may be determined experimentally or may exist only as an algorithm that is evaluated numerically.

The expected value $E(X_i \mid .)$, the variance $V(X_i \mid .)$, and the standard deviation $SD(X_i \mid .)$ of an input variable X_i for i = 1, …, N may be estimated from measurement data (Type A) and/or assessed by scientific judgment (Type B). Therefore, the measurement equation provides a practical way to combine scientific judgment and the results of statistical analyses of measurement data. The expected value E(Y | .) is obtained by substituting the expected values $E(X_i \mid .)$ for the input variables X_i for i = 1, …, N in the measurement equation:

$$E(Y \mid .) = f(E(X_1 \mid .), E(X_2 \mid .), \ldots, E(X_N \mid .)). \tag{2}$$

In order to determine V(Y | .), the measurement equation is approximated by a first-order Taylor series. This provides the following equation called the law of propagation of uncertainty:

$$V(Y \mid .) = \sum_{i=1}^{N} c_i^2 V(X_i \mid .) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} c_i c_j SD(X_i \mid .) SD(X_j \mid .) r(X_i, X_j), \tag{3}$$

where c_i represents the partial derivative of the function f with respect to X_i evaluated at $E(X_i)$, and $r(X_i, X_j)$ denotes the correlation coefficient between X_i and X_j for i, j = 1, 2, …, N. The GUM (Section F.1.2) describes a number of approaches to quantify correlation coefficient. As discussed in the GUM (Section 5.1.2, Note), equation (3) may be expanded to include higher order terms from the Taylor series. Then, SD(Y | .) is $\sqrt{(V(Y \mid .))}$. The method of evaluating the expected value E(Y | .) and the standard deviation SD(Y | .) from equations (2) and (3) respectively is referred to as the method of propagating uncertainties. The effectiveness of this method depends on the thoroughness of the measurement equation and the adequacy of the expected values and the standard deviations of the input variables. An alternative to propagating uncertainties is indicted later in this section. The *estimated value* of measurand, denoted by y, is the expected value $E(Y \mid .)$, and the *standard uncertainty* concerning the value of measurand, denoted by u(y), is the standard deviation $SD(Y \mid .)$. The estimated value $E(Y \mid .) = y$ is also referred to as the *result of measurement*. The variance $V(Y \mid .)$ is equal to $u^2(y)$. The quantities y and u(y) represent the current state of knowledge about the expected value and the standard deviation of Y based on all available information. According to this interpretation, any probability distribution that has the expected value y and the standard deviation u(y) qualifies as a state-of-knowledge distribution of Y. Thus *standard uncertainty* is the standard deviation of a state-of-knowledge distribution of the value of measurand. A probability distribution characterized by y and u(y) is not necessarily the same as a mathematically derived probability distribution of Y. Note that equation (3) propagates uncertainties rather than distributions. When it is useful to indicate that u(y) has been obtained by combining a number of uncertainty components, the standard uncertainty is termed *combined standard uncertainty*, and denoted by $u_c(y)$.

An alternative to the method of propagating uncertainties is numerical simulation. Numerical simulation avoids approximating the function f, of equation (1), by a Taylor series. Simulation is possible whenever the measurement equation (1) can be numerically evaluated. Using assumed or derived forms for the probability distributions characterized by the expected values and the standard deviations of the input variables, a sufficient number of the values of Y may be simulated numerically. The simulated values of Y then provide $E(Y \mid .) = y$ and $SD(Y \mid .) = u_c(y)$. Numerical simulation is a legitimate approach because the probability distributions of all input variables are fully characterized. This approach may be referred to as a propagation of distributions by numerical simulation rather than a propagation of uncertainties. Work is progressing in this direction [3].

Note: Suppose extensive experimental, scientific, and theoretical knowledge exists to afford a fully Bayesian approach to determine the (posterior) probability distribution of the value of measurand. In that case one may use a fully Bayesian approach. The results would be "GUM compliant" with the identity function Y = X as measurement equation.

2.3 Expanded Uncertainty, Coverage Factor, and Coverage Probability

In certain applications, it is necessary to express the uncertainty as an interval about the estimated value of measurand. The GUM concepts of expanded uncertainty, coverage factor, and coverage probability relate to this need. We will interpret these concepts from the viewpoint of treating the value of measurand as a random variable. The *expanded uncertainty*, denoted by U, is obtained by multiplying the standard uncertainty $SD(Y \mid .) = u_c(y)$ by a factor denoted by k. Thus $U = k \times SD(Y \mid .)$. Expanded uncertainty defines the interval $[E(Y \mid .) - k \times SD(Y \mid .), E(Y \mid .) + k \times SD(Y \mid .)]$ about $E(Y \mid .) = y$. The GUM has not assigned a name to this interval. We will call this interval an *expanded uncertainty interval* and write it as $[E(Y \mid .) \pm k \times SD(Y \mid .)] \equiv [y \pm k \times u_c(y)]$. This interval may alternatively be referred to as a *k-standard uncertainty interval*. The *coverage probability* associated with the expanded uncertainty interval is the probability $\Pr[E(Y \mid .) - k \times SD(Y \mid .) \leq Y \leq E(Y \mid .) + k \times SD(Y \mid .)]$, where Y is a random variable, and $E(Y \mid .) = y$, $SD(Y \mid .) = u_c(y)$, and k are treated as constants. The coverage probability concerns a state-of-knowledge distribution of Y, and it is a conditional statement given that the evaluated expected value y and the evaluated standard uncertainty $u_c(y)$ are known quantities. The multiple k determines the width of the interval and thus the coverage probability. Hence k is called a *coverage factor*. In order to establish a relationship between the coverage factor and the coverage probability, some assumption about the form of the state-of-knowledge distribution of Y is required. The relationship between the coverage probability and the coverage factor is indicated in the GUM by writing the latter as k_p where p is coverage probability.

Note: The GUM (Section 6.2.2) uses the words "level of confidence" as a synonym for "coverage probability." Since the term level of confidence is usually associated with frequentist confidence intervals, we do not recommend its use in connection with expanded uncertainty intervals.

2.4 Doubt About Evaluated Combined Standard Uncertainty

The evaluated combined standard uncertainty $u_c(y)$ could be doubtful for a number of reasons. In order to assure that $u_c(y)$ is adequate for the needs, all of the following sources of doubt must be considered. Only a small number of independent measurements were used in a Type A evaluation. The GUM (Section E.4.3, Table E.1) shows that the doubt about a Type A standard uncertainty arising from purely statistical reason of limited sampling can be surprisingly large when the number of independent measurements is small. Likewise, $u_c(y)$ could be doubtful because a Type B assessment is not very reliable. Frequently, the main source of doubt is the inadequate effort made to identify significant influence quantities and the failure to include in $u_c(y)$ the corresponding components of uncertainty. Some influence quantities may be deemed to be significant, but the corresponding components of uncertainty cannot be assessed for lack of sufficient experimental or scientific knowledge. The law of propagation of uncertainty could itself be an important source of doubt about $u_c(y)$. Use of second order terms as discussed in the GUM (Section 5.1.2, Note) is a helpful step in the right direction. But how does one know the importance of second order terms in advance of actually computing them? Also, $u_c(y)$ may be doubtful because the measurements may not be independent and representative for the intended scope of the measurement environment (see, Subsection 2.7). The quantity actually measured may be an approximation of the quantity whose value is desired. In such cases, the discrepancy between the intended measurand and the quantity realized for measurement could be an important source of doubt about $u_c(y)$. Inadequate specification of the measurand could be an important source of doubt (see, the GUM, Section D.6.2). In addition, the doubt about $u_c(y)$ due to unrecognized effects could be important. Presence of such effects is suggested by significant differences in the estimated values of a common measurand by two or more methods (or laboratories). In general, the doubt about evaluated combined standard uncertainty $u_c(y)$ cannot be quantified.

2.5 Use of a Student's t-Distribution

Often, metrologists associate coverage factors of 2 and 3 with approximate 95 % and 99 % coverage probabilities respectively. This relationship between the coverage factor and the coverage probability presumes an approximate normal distribution for the value of measurand. The GUM prescribes an alternative to the normal distribution that accounts for the doubt about standard uncertainty $u_c(y)$ due to the small number of independent measurements used in Type A evaluations and/or the poor reliability of Type B assessments. The GUM prescription involves the use of a Student's t-distribution with effective degrees of freedom as determined by the Welch-Satterthwaite approximation. We will discuss the advantage of a t-distribution and the applicability of the GUM prescription. The t-distribution is named after its developer W. S. Gosset, who wrote under the pen name Student.

First, consider the special case where the value of measurand Y, treated as an unknown constant, is estimated from a frequentist analysis of a series of n measurements that are assumed to be independent and identically normally distributed with expected value Y and some standard deviation σ. Suppose the sample mean and sample standard deviation are y and s respectively. Then y is an estimate of Y and $s/\sqrt{n}$ is an estimate of the population standard deviation of the mean $\sigma/\sqrt{n}$. It can be shown that the ratio $(y - Y)/(s/\sqrt{n})$ has the Student's t-distribution with $\nu = (n - 1)$ degrees of freedom (d.f.) [6]. Consequently, $\Pr[y - t_p(\nu) \times s/\sqrt{n} \leq Y \leq y + t_p(\nu) \times s/\sqrt{n}] = p$, where $t_p(\nu)$ denotes a value of the t-distribution with d.f. $\nu = (n - 1)$ such that $\Pr[-t_p(\nu) \leq t \leq t_p(\nu)] = p$. The interval $[y \pm t_p(\nu) \times s/\sqrt{n}]$ is called a confidence interval with confidence level p. In this confidence interval, y and $s/\sqrt{n}$ are random variables and Y is an unknown constant. A confidence interval is not an expanded uncertainty interval because in the latter case Y is a random variable, and y and $s/\sqrt{n}$ are constants. It turns out that in this particular case, a Bayesian interval exists that is numerically identical to the corresponding confidence interval. This result comes from the following Theorem [5].

Theorem 1: Let the sample quantities y and s^2 be independently distributed as normal $N(Y, \sigma^2/n)$ with expected value Y and variance σ^2/n, and (σ^2/ν) times chi-square $\chi^2(\nu)$ distribution with ν degrees of freedom respectively. Suppose a priori that Y and $\log \sigma$ are approximately independent and locally uniform. Then, given y and s^2, (a) σ is distributed as $(\sqrt{\nu} \times s)$ times $\chi^{-1}(\nu)$ distribution, (b) conditional on σ, Y is distributed as $N(y, \sigma^2/n)$, and (c) unconditionally, $(Y - y)/(s/\sqrt{n})$ has the Student's t-distribution with $\nu = (n - 1)$ degrees of freedom.

The prior distributions stipulated in this theorem are non-informative. From this theorem, it follows that $\Pr[y - t_p(\nu) \times s/\sqrt{n} \leq Y \leq y + t_p(\nu) \times s/\sqrt{n}] = p$, where Y is a random variable, and y and $s/\sqrt{n}$ are constants. Thus the interval $[y \pm k_p \times u(y)]$, where $u(y) = s/\sqrt{n}$ and $k_p = t_p(\nu)$ qualifies as an expanded uncertainty interval with coverage probability p. Thus, in the special case of independent and identically normally distributed measurements, a frequentist confidence interval is numerically identical to the corresponding expanded uncertainty interval.

The use of a t-distribution in place of the normal distribution accounts for the doubt about s^2/n as represented by the degrees of freedom $\nu = (n - 1)$ by increasing the coverage factor k_p and hence the width of the interval $[y \pm k_p \times s/\sqrt{n}]$ for a fixed coverage probability. For example, suppose $n = 4$ and d.f. $\nu = (n - 1) = 3$. Now suppose the coverage probability is fixed at $p = 95$ %, then the coverage factors for the t-distribution and the normal distribution are $t_{0.95}(3) = 3.18$ and $k_{0.95} = 1.96$ respectively, a difference of 62 %. For ν of 15 or more the difference between the coverage factors for the t-distribution and the normal distribution is less than 9 %, when the intended coverage probability is 95 %. This illustrates that a t-distribution is useful when Y is estimated from a small number of measurements that are believed, based on experimental and theoretical knowledge, to be approximately independent and identically normally distributed. But the benefit of using a t-distribution rather than the normal distribution is insignificant when the number of independent measurements is more than 15.

The GUM (Section G.6.4) prescription for using a Student's t-distribution is as follows. Evaluate the expected value $E(Y \mid .) = y$, and the combined standard uncertainty $SD(Y \mid .) = u_c(y)$ from equations (2) and (3) respectively. Estimate the effective degrees of freedom ν_{eff} of $u_c(y)$ from the Welch-Satterthwaite approximation as discussed in the GUM (Section G.4). Obtain $t_p(\nu_{eff})$ for the required coverage probability p from a table of Student's t-distribution. Take $k_p = t_p(\nu_{eff})$ and calculate the expanded uncertainty $U = k_p \times u_c(y)$.

The concept of degrees of freedom as used by the GUM (Section E.4.3) for independent and identically normally distributed measurements has been extended by the GUM (Section G.4.2) for the "reliability" of Type B evaluations. This extension has been developed to enable the use of Welch-Satterthwaite approximation for both Type A and Type B evaluations of the standard uncertainties. The Welch-Satterthwaite approximation applies to those input variables $X_1, X_2, \ldots, X_N$ that are not mutually correlated.

The GUM prescription may be argued as an approximation when the measurement equation is a linear function of N independent variables $X_1, X_2, \ldots, X_N$, and X_i is estimated from a series of n_i measurements that are assumed to be independent and identically normally distributed for every $i = 1, 2, \ldots, N$. Research is needed to understand the reasonableness of this approximation. The GUM prescription may not be a reasonable approximation when not all of $X_1, X_2, \ldots, X_N$ are estimated from a series of independent and identically normally distributed measurements or some of the $X_1, X_2, \ldots, X_N$ are correlated or the measurement equation is a highly non-linear function of $X_1, X_2, \ldots, X_N$. Conclusion: the GUM prescription may not be a reasonable approximation in many applications.

2.6 Expanded Uncertainty Intervals Based on the Chebyshev and the Gauss Inequalities

Some assumption about the form of the distribution of Y, as characterized by the result of measurement y and the standard uncertainty $u_c(y)$, is required to relate the coverage probability and the coverage factor used to define an expanded uncertainty interval. The coverage probability associated with an expanded uncertainty interval is doubtful to the extent that the assumed form of the distribution of Y is doubtful. Therefore, we propose that the metrologist report the minimum coverage probability for a class of probability distributions rather than a specified coverage probability for a particular assumed distribution. We will describe a simple and widely applicable approach to set the coverage factor that defines an expanded uncertainty interval with a desired minimum coverage probability for two common classes of distributions.

When nothing can be assumed about the distribution of Y except that $E(Y \mid .) = y$ and $SD(Y \mid .) = u_c(y)$, then the Chebyshev inequality [6] applies. Accordingly, $\Pr[y - k \times u_c(y) \leq Y \leq y + k \times u_c(y)] \geq (1 - 1/k^2)$. In particular, the coverage probability associated with the expanded uncertainty interval $[y \pm k \times u_c(y)]$ is at least 75 % for $k = 2$, and is at least 89 % for $k = 3$. Suppose the desired minimum coverage probability is 85 %, then by setting $1 - 1/k^2 = 0.85$, we get the coverage factor k as 2.58.

In many applications, it is reasonable to assume that the distribution of Y, as characterized by $E(Y \mid .) = y$ and $SD(Y \mid .) = u_c(y)$, is symmetric and unimodal about y. With this assumption, we can invoke the Gauss inequality [6], and claim that $\Pr[y - k \times u_c(y) \leq Y \leq y + k \times u_c(y)] \geq [1 - 4/(9k^2)]$. In particular, the coverage probability associated with the expanded uncertainty interval $[y \pm k \times u_c(y)]$ is at least 89 % for $k = 2$, and is at least 95 % for $k = 3$, when the distribution of Y is symmetric and unimodal about y. Suppose the desired minimum coverage probability is 90 %, then by setting $1 - 4/(9k^2) = 0.90$, we get the coverage factor k as 2.11.

Note: A t-distribution is symmetric and unimodal. But the coverage probability associated with the interval $[y \pm k \times s/\sqrt{n}]$ for $k = 2$, based on the t-distribution with $\nu = (n - 1) = 3$ degrees of freedom, is 86 % rather than 89 % or more. This is because the standard deviation of the t-distribution with ν degrees of freedom is $\sqrt{[\nu/(\nu - 2)]}$. When the degrees of freedom $\nu = (n - 1) = 3$, $SD(Y \mid .) = \sqrt{3} \times s/\sqrt{n} = 1.732 \times s/\sqrt{n}$. Hence $s/\sqrt{n}$ is less than $SD(Y \mid .)$. The minimum coverage probability of 89 %, associated with the interval $[y \pm k \times s/\sqrt{n}]$ for $k = 2$, applies to symmetric and unimodal distributions that are characterized by the expected value $E(Y \mid .) = y$ and the standard deviation $SD(Y \mid .) = s/\sqrt{n}$.

2.7 Additional Comments

The choice of coverage factor k involves a trade-off between the width of the expanded uncertainty interval $[y \pm k \times u_c(y)]$ and the corresponding coverage probability for the assumed probability distribution of Y. In the practical cases of interest, narrower intervals corresponding to smaller values of k are more interesting. But they have lower coverage probabilities. The choice of coverage factor k = 2 provides a reasonable balance between the width of the interval and the coverage probability for the commonly assumed forms of distributions. The coverage probability associated with a 2-standard uncertainty interval is at least 75 % regardless of the form of the distribution of Y characterized by the result of measurement and the standard uncertainty. The coverage probability jumps to at least 89 % when the distribution can be assumed to be symmetric and unimodal.

One of the most critical assumptions in statistical analyses is the independence of measurements. Suppose, for example, the intended scope of the measurement environment is long-term involving a number of influence quantities that may not change appreciably over short periods of time. Now suppose the available data are short-term measurements during which a number of important influence quantities remained constant. Then the short-term measurements could be positively correlated resulting in under-evaluation of long-term uncertainty. So it is important to clarify the intended scope of the measurement environment. Then the measurement protocol should be designed to assure that the measurement data represent variation in all relevant significant influence quantities, and that the data conform to the assumption of independence built in the statistical model used for data analysis.

3. SUMMARY

We have shown that the GUM is clear and coherent when interpreted with the following precepts. First, all quantities involved in measurement are random variables with probability distributions that represent the state of knowledge about them, a la Bayesian statistics. Second, the GUM is mainly concerned with the expected values and the standard deviations of the random variables involved in measurement rather than with the fully characterized probability distributions. Third, Type A estimates obtained from frequentist analyses of measurement data are regarded as approximations to the corresponding results from Bayesian analyses based on non-informative prior distributions.

The GUM is based on the concept of measurement equation. A measurement equation expresses the value of measurand as a function of all those variables that affect its assessment. The expected value and the standard deviation of an input variable to the measurement equation are evaluated from statistical analysis of measurement data (Type A) and/or by scientific judgment (Type B). The statistical methods employed for Type A evaluation may be either Bayesian or frequentist. Type A evaluations from Bayesian analyses of measurement data and Type B evaluations from scientific judgment are mathematically compatible inputs to the measurement equation because both treat the input quantities as random variables. But Type A evaluations are usually frequentist estimates. They are not mathematically compatible with Type B evaluations because the frequentist methods treat the input quantities as unknown constants. However, in the practical cases of interest, the frequentist estimates may be regarded as approximations to the corresponding results from Bayesian analyses based on non-informative prior distributions. Therefore, it is legitimate to treat frequentist estimates and Type B evaluations as mathematically compatible inputs to the measurement equation. The evaluated expected values and the standard deviations of the input variables are then combined by the method of propagating uncertainties, as discussed in Subsection 2.2, to obtain the expected value and the standard deviation of the value of measurand. The expected value is taken as the estimated value of measurand and the standard deviation is taken as the combined standard uncertainty concerning the value of measurand. The estimated value of measurand is also referred to as the result of measurement. Any probability distribution whose parameters match the expected value and the standard deviation of the value of measurand qualifies as a state-of-knowledge probability distribution of Y. Numerical simulation is an alternative to the method of propagating uncertainties. Indeed, simulation may be a preferred approach when the measurement equation can be numerically evaluated. The expanded uncertainty is a multiple of the standard uncertainty that defines an interval about the estimated value of measurand that is presumed to cover a large fraction of the distribution of Y. The multiple is called coverage factor and the fraction of distribution covered is called coverage probability. The coverage probability associated with an expanded uncertainty interval is a conditional statement given that the evaluated expected value and the evaluated standard deviation of the value of measurand are known quantities. The GUM prescription to construct expanded uncertainty intervals, involving the use of a Student's t-distribution with effective degrees of freedom as determined by the Welch-Satterthwaite approximation, may not be a reasonable approximation in many applications.

Some assumption about the form of the distribution of Y, as characterized by the result of measurement y and the standard uncertainty $u_c(y)$, is required to relate the coverage probability and the coverage factor used to define an expanded uncertainty interval. The coverage probability associated with an expanded uncertainty interval is doubtful to the extent that the assumed form of the distribution of Y is doubtful. Therefore, we have proposed the use of the Chebyshev and the Gauss inequalities to construct expanded uncertainty intervals with a minimum coverage probability for a class of probability distributions.

4. COMMENTS AND RECOMMENDATIONS

An effective approach to quantify uncertainty is to make an "uncertainty budget" that includes the important components of uncertainty and identifies their interrelationships. Then, have the uncertainty budget reviewed by peer subject matter experts to assure that no potentially significant sources of uncertainty have been ignored, within the limits of available knowledge, and that the estimates of the components of uncertainty seem reasonable. Usually, the combined standard uncertainty is reported to at most two significant digits. The components of uncertainty that contribute only a small fraction to the combined standard uncertainty are often identified in the budget as insignificant and neglected.

Clarify the intended scope of the measurement environment for the specified measurand. Is it short-term or long-term? Then make sure that the measurement data represent variation in all significant influence quantities for the intended scope of measurements.

When using a series of measurements to estimate the value of measurand, demonstrate that the measurement process is in a state of statistical control. The number of independent measurements used for each component of uncertainty should be as-large-as-practical but not less than four. (Estimates based on fewer than four measurements may be used when they are believed to be reliable based on scientific judgment and prior experience.)

For archives of measurement data, tabulate standard uncertainties with comments rather than expanded uncertainty intervals because it is standard uncertainties rather than expanded uncertainty intervals that get propagated through a hierarchy of measurements.

A frequentist confidence interval is not an expanded uncertainty interval because in the latter case the value of measurand is a random variable, and the result of measurement and the standard uncertainty are known quantities.

The degrees of freedom, as evaluated by the Welch-Satterthwaite approximation, may not be an adequate measure of the doubt about evaluated combined standard uncertainty. Reason: the important sources of doubt may not be limited to the small number of independent measurements used in Type A evaluations and/or the poor reliability of Type B evaluations.

As a general rule, use the coverage factor two to construct expanded uncertainty intervals. The choice of coverage factor requires some assumption about the form of the distribution of the value of measurement, and involves a trade-off between the width of the expanded uncertainty interval and the coverage probability. The coverage factor two provides a reasonable balance between the width of the interval and the coverage probability for the commonly assumed forms of distributions.

Use the Gauss inequality to set the coverage factor for a desired minimum coverage probability when the distribution of the value of measurand, as characterized by the result of measurement and the standard uncertainty, can be assumed to be symmetric and unimodal.

When a particular probability distribution, such as the normal or a t-distribution, is used to set the coverage factor for a desired coverage probability, provide some justification that the assumed form of distribution is reasonable. The justification may have experimental and/or theoretical basis.

Quantification of uncertainty requires expenditure of cost and time. The effort expended must be proportional to the quality of uncertainty statement that is need by the potential users of the result of measurement.

We are interested in receiving feedback from the users of the GUM about the viewpoints expressed in this paper.

ACKNOWLEDGMENT

The suggestions received from Norm Berk, Raju Datla, Tyler Estler, Keith Eberhardt, Charles Hagwood, Gail Hixenbaugh, Al Jones, Steve Phillips, Andrew Rukhin, and Barry Taylor are gratefully acknowledged.

REFERENCES

[1] *Guide to the Expression of Uncertainty in Measurement* (ISBN 92-67-10188-9), International Organization for Standardization, Case postale 56, CH-1211, Geneva 20, Switzerland, 1995.

[2] *American National Standard for Expressing Uncertainty - U.S. Guide to the Expression of Uncertainty in Measurement* (ANSI/NCSL Z540-2-1997), National Conference of Standards Laboratories, 1800 30th Street, Suite 305B, Boulder, CO 80301, 1997.

[3] Barry N. Taylor: Personal Communication, March 2000.

[4] Gleser, Leon. J., Assessing Uncertainty in Measurement, *Statistical Science*, 1998, 13, pp. 277-290.

[5] Box, George E. P., and Tiao, George C., *Bayesian Inference in Statistical Analysis*, Addison -Wesley Publishing Company, Reading, MA, 1973. (Theorem 2.4.1, and other parts of the book.)

[6] Stuart, Alan, and Ord, J. Keith, Kendall's *Advanced Theory of Statistics, Distribution Theory*, Vol. 1, Oxford University Press, NY, 1987.

APPENDIX

The following example illustrates the concept of non-informative prior distributions, and shows that the expected value of a Bayesian posterior distribution based on a non-informative prior distribution is approximately equal to the corresponding estimate from frequentist analysis. Suppose the measurand is the mean breaking strength X of a large batch of certain parts. Suppose a random sample (set) of $n = 12$ parts is selected from the batch, and their mean breaking strength is determined to be x based on destructive testing. Suppose the standard deviation of each measurement, including the test and the part variation, is known with high reliability to be $\sigma = 17.3$ units. For simplicity, we are assuming that the standard deviation σ is known. We will assume that the sampling distribution of x can be taken as normal with expected value X and standard deviation $\sigma/\sqrt{n} = 17.3/\sqrt{12} = 5.0$ units. Now suppose the value of sample mean x is 70.0 units. Then the frequentist estimate of X is 70.0 units with a standard deviation 5.0 units.

A Bayesian analysis starts with a prior distribution, representing prior state of knowledge, about X, then updates the state of knowledge based on the results of measurement. Suppose that, based on prior knowledge of the manufacturing process, the prior distribution of X can be assumed to be normal with some expected value μ_0 and some standard deviation σ_0. Since the probability distribution of the mean x is assumed to be normal with expected value X and standard deviation $\sigma/\sqrt{n} = 5.0$ units, the probability density function $p(x \mid X)$ of x given X and $\sigma/\sqrt{n}$ is proportional to $\exp[-(n/2)((x - X)/\sigma)^2]$, where $n = 12$, and $\sigma = 17.3$ units. Now given $x = 70.0$ units, the probability function $p(x \mid X)$ may be regarded as a function not of x but of X. When so regarded the function $p(x \mid X)$ is called a likelihood function of X given x and denoted by $l(X \mid x)$. Thus

$$l(X \mid x) \propto \exp[-(n/2)((x - X)/\sigma)^2], \tag{4}$$

where $x = 70.0$ units, $n = 12$, and $\sigma = 17.3$ units. Then by Bayes theorem [5], the posterior distribution of X given x is also normal with expected value $E(X \mid x)$ and standard deviation $SD(X \mid x)$ where

$$E(X \mid x) = [1/(1+r)]x + [1 - 1/(1+r)]\mu_0, \tag{5}$$

$$SD(X \mid x) = (\sigma/\sqrt{n}) \times (1/\sqrt{1+r}), \tag{6}$$

and

$$r = (\sigma^2/n)/\sigma_0^2, \tag{7}$$

is the ratio of the variance of the sampling distribution of x to the prior variance of X. The ratio r represents the importance of the prior distribution relative to the current measurement data. Clearly as r tends to 0, $E(X \mid x)$ tends to x, where x is the frequentist estimate of X, and $SD(X \mid x)$ tends to $\sigma/\sqrt{n}$, the standard deviation of the sampling distribution of x. The ratio r is close to zero when the prior variance σ_0^2 is very large relative to σ^2/n (or the sample size n is extremely large). Such values of r represent the situation that the prior state of knowledge is meager in relation to the information in the current measurement data. Prior distributions for which r is close to zero are appropriately called non-informative prior distributions.

Consider two different prior distributions. Prior distribution 1 is normal $N(\mu_0, \sigma_0^2)$ with $\mu_0 = 60.0$ and $\sigma_0 = 10.0$. Prior distribution 2 is normal $N(\mu_0, \sigma_0^2)$ with $\mu_0 = 60.0$ and $\sigma_0 = 1000.0$. For prior distribution 1, $r = 0.25$. Thus $E(X \mid x) = 68.0$ and $SD(X \mid x) = 4.47$. The posterior expected value $E(X \mid x) = 68.0$ is closer to the sample mean $x = 70.0$ than the prior expected value $\mu_0 = 60.0$. Such is often the case, because in many scientific applications the ratio r is small. For prior distribution 2, $r = 0.000025$ indicating that the prior distribution 2 is non-informative relative to the information in the current measurement data. In this case $E(X \mid x) = 70.0$, the same result as obtained from frequentist analysis.

Dated: Thursday, May 25, 2000
Place: NIST, Gaithersburg, MD 20899-8980

NIST Technical Publications

Periodical

Journal of Research of the National Institute of Standards and Technology—Reports NIST research and development in those disciplines of the physical and engineering sciences in which the Institute is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Institute's technical and scientific programs. Issued six times a year.

Nonperiodicals

Monographs—Major contributions to the technical literature on various subjects related to the Institute's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NIST, NIST annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NIST under the authority of the National Standard Data Act (Public Law 90-396). NOTE: The Journal of Physical and Chemical Reference Data (JPCRD) is published bimonthly for NIST by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements are available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

Building Science Series—Disseminates technical information developed at the Institute on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NIST under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NIST administers this program in support of the efforts of private-sector standardizing organizations.

Order the **following** *NIST publications—FIPS and NISTIRs—from the National Technical Information Service, Springfield, VA 22161.*

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NIST pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NIST Interagency or Internal Reports (NISTIR)—The series includes interim or final reports on work performed by NIST for outside sponsors (both government and nongovernment). In general, initial distribution is handled by the sponsor; public distribution is handled by sales through the National Technical Information Service, Springfield, VA 22161, in hard copy, electronic media, or microfiche form. NISTIR's may also report results of NIST projects of transitory or limited interest, including those that will be published subsequently in more comprehensive form.

U.S. Department of Commerce
National Institute of Standards
and Technology
Gaithersburg, MD 20899–0001

Official Business
Penalty for Private Use $300