Practical Methods for Analysis of Uncertainty Propagation[1]

Howard Castrup, Ph.D.
President
Integrated Sciences Group
Bakersfield, CA 93306

Key Words

Measurement
Uncertainty
Reliability
Calibration
Modeling
Product Quality
Error Modeling
Testing
Statistical Process Control
Measurement Decision Risk
Computer Workstations

Abstract

An uncertainty analysis methodology is described that is relevant to equipment tolerancing, analysis of experimental data, development of manufacturing templates and calibration of standards. By assembling the methodology from basic measurement principles, controversies regarding uncertainty combination are avoided. Whether applied to laboratory measurements or to product acceptance test data, the methodology leads to rigorous assessments of measurement accuracies and unambiguous evaluations of measurement decision risks.

Introduction

This paper reports an uncertainty analysis methodology that yields unambiguous results that can be applied directly to the assessment of measurement uncertainty. The methodology specifically addresses the final three stages of a four-part uncertainty analysis process:

- ❑ Measurement configuration description. Identification of measurement components and their interrelationships.
- ❑ Measurement error model development. Identification of error components and their interrelationships.
- ❑ Statistics development. Construction of statistical distributions for each measurement error component.
- ❑ Uncertainty analysis. Analysis and assessment of measurement uncertainty.

A methodology for describing measurement configurations is the topic of current research and will be reported in detail in the near future.

---

[1]Published in the *Proceedings of the 38th Annual ISA Instrumentation Symposium*, Las Vegas, April 1992.

Practical vs. Simple
The title of this paper claims that the methodology described is a practical one. This may imply that the methodology is simple or easy to use. If so, the implication is unintentional. Some of the mathematics tend to involve multiple terms, subscripts and superscripts and may appear a little daunting at times. In this paper the term "practical" is meant to mean usable or relevant to user objectives, such as equipment tolerancing or decision risk management. Simplicity and ease of use will follow once the methodology is embedded in user-interactive workstation applications, where the math can be largely hidden from view.

Departure from Tradition
Uncertainty analysis methodologies have traditionally been confined to techniques that are conceptually simple and straightforward. These methodologies have been developed in accordance with the available computational capabilities of the decades before desktop workstations became widespread. Unfortunately, while conventional methodologies are often easily understood, they are frequently ambiguous, restricted, and, sometimes useless or even dangerous. In contrast, the methods described in this paper are unambiguous, completely general and lead to a better understanding of the nature and extent of uncertainties surrounding a given measurement situation.

Accessibility to the Engineering Community
The complexity of the methodology of this paper can be made available to the general engineering community through dedicated software written for today's powerful desktop computers. What may have been considered to be hopelessly difficult in the past can now be made almost trivial from the standpoint of the analyst. Moreover, with the evolution of the desktop computer's graphical user interface (GUI), using a complex methodology, such as is described herein, can even be fun.

With these considerations in mind, it is the author's belief that the issue of uncertainty analysis needs to undergo a paradigm shift with a view toward achieving the following objectives:

- ❑ Develop uncertainty analysis methodologies that are relevant to scientific inquiry, standards calibration, parameter testing, production template development and other aspects of the marketplace.
- ❑ Implement these methodologies in menu-driven platforms with graphical user interfaces.

To explore in detail the issue of methodological relevance, it will be helpful to review some background on why measurements are made and how analyzing uncertainty leads to understanding, interpreting, and managing measurement results.

Why make measurements?

A variety of reasons for making measurements can be stated. We make measurements to discover new facts, verify hypotheses, transfer physical dimensions, make adjustments to physical attributes, or obtain information necessary to make decisions. The varied reasons for making physical measurements are found in the typical high-tech product development process. Each phase of this process involves the transfer of measurement information across an interface, as shown in Figure 1.

The process involves R&D, where new data are taken and hypotheses are tested; prototype development, where dimensions are transferred, attributes are adjusted or modified and decisions are made; design, where prototyping experience leads to decisions on optimal specs and allowable tolerances; production, where molds, jigs and templates transfer physical dimensions; testing, where decisions to accept or reject parts and assemblies are made; and usage, where customer response to product quality, reliability and performance is fed back in the form of sales activity, warranty claims, legal actions, publicity, etc.

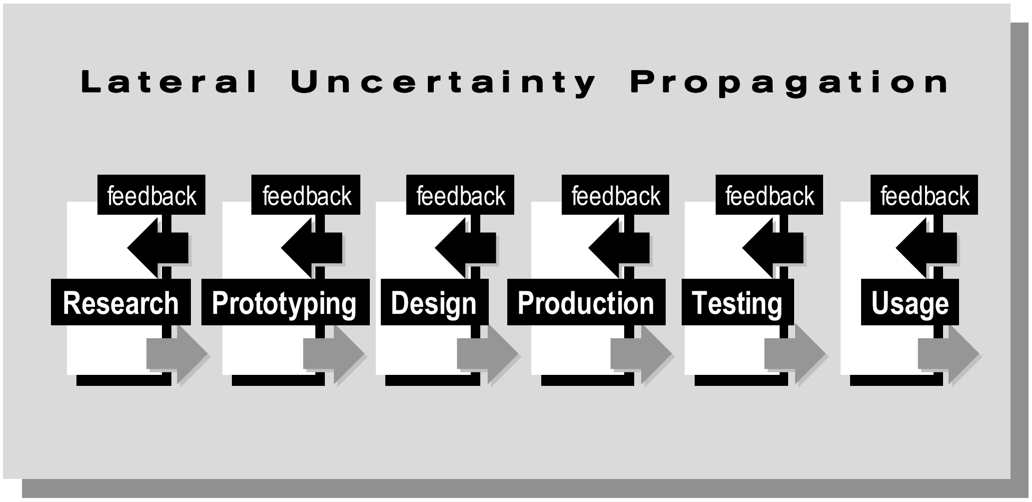

Figure 1. Lateral Uncertainty Propagation. Measurement results are transferred from stage to stage in the typical product development process. Measurement uncertainties accompany each measurement transferal. The appropriateness of measurement accuracies and other characteristics are fed back to modify and refine production process approaches and parameters.

Each product development interface shown in Figure 1 is supported by a measurement assurance infrastructure embodied in a test and calibration hierarchy. The basic hierarchy structure is shown in Figure 2.

In a typical hierarchy, testing of a given end item attribute by a test system yields a reported in-or out-of-tolerance indication, an adjustment if needed, and a beginning-of-period in-tolerance probability (measurement reliability). Similarly, the results of calibration of corresponding test system attributes include reported in- or out-of-tolerance indications, attribute adjustments and beginning-of-period measurement reliabilities. The same sort of data results from calibrating the supporting calibration systems and accompanies calibrations down through the hierarchy until a point is reached where the "unit under test" (UUT) of interest is a primary calibration standard.

Why estimate uncertainties?

All physical measurements are accompanied by measurement uncertainty. Since measurement results are transmitted laterally across development process interfaces and vertically across support hierarchy interfaces, uncertainties in these results also propagate both laterally and vertically.

Whether we use measurements to verify hypotheses, construct artifacts, or test products, we want to know how good our measurements are. Within the context of each application, this is synonymous with knowing the confidence with which our measurements allow us to make decisions, adjust parameters and so on.

A perhaps pessimistic, yet practical, way of looking at the situation is to say that we want to be able to assess the chances that negative consequences may result from applying knowledge obtained from measurements. It can be shown [1-11] that the probability for negative consequences increases with the uncertainty associated with a measurement result. Thus managing the risks involved in applying measurement results is intimately linked with managing measurement uncertainty.

Optimizing the management of measurement decision risks involves (1) linking specific values of a physical attribute with outcomes that may result from using the attribute and (2) estimating the probability of encountering these values in practice [4-9]. If high probabilities exist for unknowingly encountering attribute values associated with negative consequences, we say that our knowledge of the attribute's value is characterized by high levels of measurement uncertainty. If the reverse is the case, we say that measurement uncertainty is not significant.

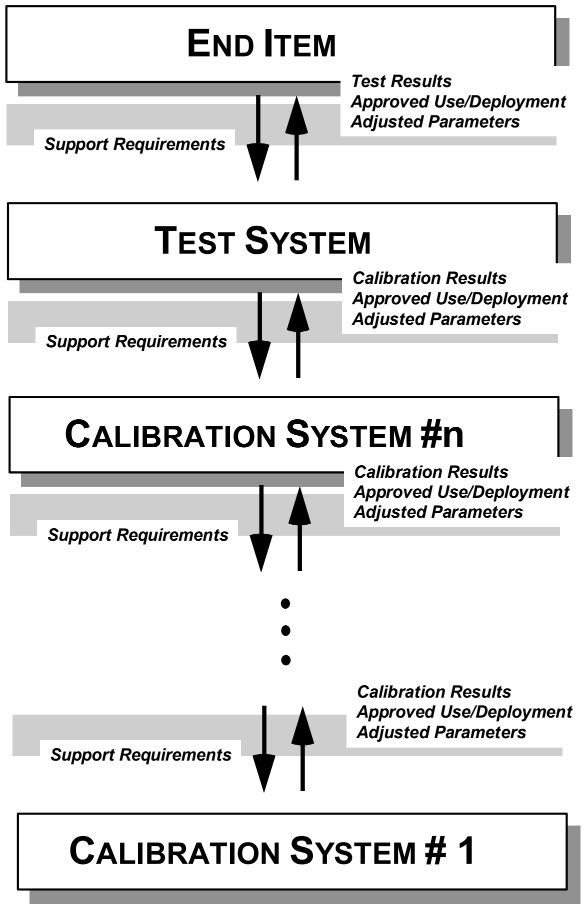

Figure 2. Vertical Uncertainty Propagation. Measurement accuracy requirements flow down from the end item or product through the measurement assurance support hierarchy. Calibrated and/or tested attributes, accompanied by measurementuncertainty, propagate upward.

If our approach to uncertainty analysis aids in estimating the probability of encountering attribute values associated with negative consequences then we have a workable, i.e., practical, measurement uncertainty analysis methodology.

Estimating Uncertainty

Conventional Methods

Conventional uncertainty analysis methodologies ordinarily employ the following steps:

1) Identify all components of error.
2) Estimate statistical or engineering variances for each component.[2]
3) Combine variances to achieve a total uncertainty estimate.
4) Estimate statistical confidence limits, based on the total estimate.

Statistical confidence limits are usually determined by assuming normally distributed error components and invoking Student's t-distribution [12-20].

Methodological Drawbacks

While step one is generally advisable, certain ambiguities and improprieties arise in the way that conventional methods address steps 2 through 4. This is due to three main drawbacks of conventional methods.

1. Arbitrary Uncertainty Models

The first drawback involves difficulties in gauging the relative impact of each component of error on total uncertainty. Some error components may contribute more significantly than others. Attempting to circumvent this problem can lead to extremely arbitrary and unwieldy weighting schemes whose applicability is often questionable.

Because of this, no universally applicable methodology exists for combining uncertainties. Most conventional approaches involve either a linear combination of component uncertainties (standard deviations) or confidence limits, or a linear combination of component variances. Linear combinations of uncertainties or confidence limits is ill-advised in virtually all cases.[3] Such combinations lead to what are often called "worst case" uncertainty estimates. They could also be called "worst guess" estimates.

---

[2]Such variances are referred to as Category A and Category B uncertainties, respectively [20].

[3]This is not the case for linear combinations of systematic measurement bias when signs are known and magnitudes can be estimated.

How uncertainties combine differs from situation to situation. Each situation requires the development of a valid model showing how error components contribute to the total error of a measurement result. In some cases, more than just a simple extrapolation from such a model is required. For example, if the appropriate error model is a linear combination of error components, it does not necessarily follow that total uncertainty can be determined from a linear combination of corresponding uncertainty component variances.

A large part of the problem stems from the fact that linear combinations of variances arising from various error components are not relevant except in cases where the error model is linear and all error components are statistically independent (s-independent). Moreover, even if s-independence pertains, linear combinations of variances are not generally useful unless all error components follow the same sort of statistical distribution and the distribution is symmetrical about the mean.

To get around these difficulties, the expedient of treating each error component as normally distributed is often employed. This is sometimes justified on the basis of the central limit theorem [20]. General methodologies for combining uncertainties with mixed statistical distributions have not been forthcoming.

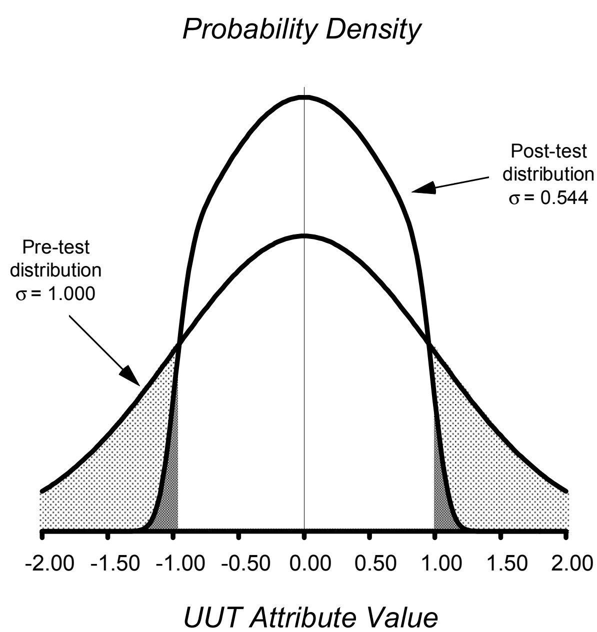

Figure 3. Pre-Test vs. Post-Test Attribute Populations. Typical statistical distributions for attribute values prior to and following test screening. The shaded areas represent probabilities for out-of-tolerance attributes. The pre-test in-tolerance percentage is approximately 68%.[4] The post-test curve corresponds to testing with a measuring system uncertainty (standard deviation) of approximately ten percent of the pre-test population uncertainty. As expected, the out-of-tolerance probability is lower after test screening than before test screening.

2. Misleading Variances - the Normality Assumption

The second drawback of the conventional approach is its reliance on statistical variance as the sole measure of uncertainty. Working with variances alone can produce misleading results. This is illustrated by considering the distributions shown in Figures 3 and 4. Figure 3 shows a population of product attribute values before and after test screening. Since testing has rejected most of the non conforming attributes, the post-test distribution's tails are pulled in toward the center [4-9].

---

[4]A not uncommon figure with products that are tested periodically as part of their in-use maintenance cycle.

From Figure 3, it is evident that, although the pre-test population is normally distributed, the post-test distribution of product attribute values is non-normal. Accordingly, treating post-test product attribute values as being normally distributed could lead to erroneous inferences about their uncertainty.[5]

This can be appreciated by considering the statistical standard deviation of post-test population values. Given the variance in the pre-test population and the accuracy of the test system, the standard deviation for the post-test distribution turns out to be approximately 0.544. If we were engaged in sampling post-test attribute values as part of a process control procedure, for example, we would likely obtain an estimate centered around this value.

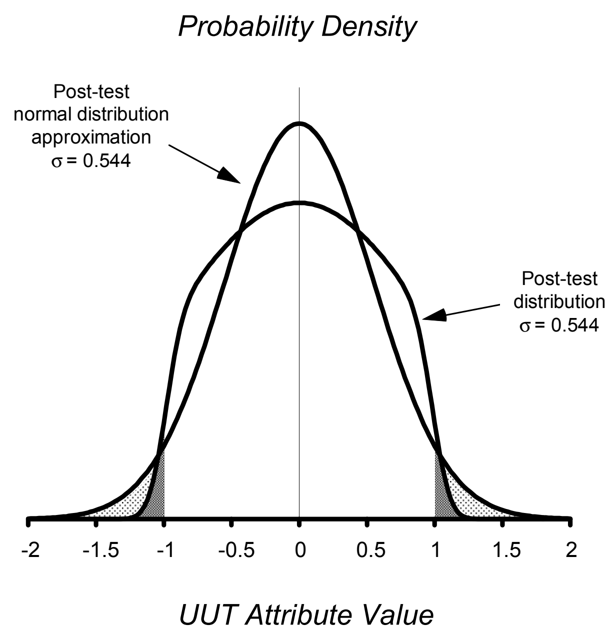

Figure 4. Post-Test Distribution Normal Approximation. The post-test distribution is contrasted with a normal distribution with equal variance. Not only are the out-of-tolerance probabilities (shaded areas) significantly different, the shapes of the distributions are dissimilar.

If we were to assume a normal distribution for the post-test population, a sampled standard deviation of 0.544 would correspond to an in-tolerance percentage of about 93% (see Figure 4). In contrast, the actual in-tolerance percentage is over 97%. When evaluating out-the-door quality levels, the difference between 93% and 97% in-tolerance can be astronomical. An erroneously low 93% level can result in unnecessary breaks in production, an unscheduled verification of parts and production machinery, and a reevaluation of the production process; all of which could be avoided by not assuming normality for the product attribute distribution.

3. Ambiguity of Application

The third drawback with conventional methods is that they produce results that are not readily applicable. The use of conventional methods typically yields an estimate of the total variance of measurement values. What then to do with this variance? True it can be used to calculate confidence limits (again, assuming normal distributions of measurements), but confidence limits are not always useful. In general, by themselves they constitute weak decision variables.

---

[5]In this context, attribute uncertainty may be equated with the probability that a product item drawn at random from the post-test population will be in-tolerance.

The relationship of statistical variances or confidence limits to probabilities associated with negative consequences, referred to earlier, is often ambiguous. Unless a statistical variance enables us to infer the statistical distribution that it characterizes, its function is primarily ornamental. Without knowledge of this distribution, we are at a loss to determine the probability that parts manufactured by one source will mate with parts manufactured by another, or the probability that calibrated test sytems will incorrectly accept out-of-tolerance products.

Methodology Requirements

Given these observations on conventional methods, it appears that what is needed is an uncertainty analysis methodology that directly generates probability estimates for attribute values. The methodology should not be restricted with regard to statistical distributions of error components, nor to assumptions of s-independence. Moreover, it should yield results that can be used in managing measurement decision risk.

Such a methodology is referred to below as the "new method."

The New Method

The new method employs an analysis procedure that differs from that followed by conventional approaches. The procedure it follows is

1) Define the measurement mathematically.
2) Identify all components of error for a given quantity of interest.
3) Construct an appropriate total error model.
4) Determine statistical distributions for each error component.
    - ❏ Identify all error sources for each error component.
    - ❏ Obtain technical information from which to identify the statistical distribution appropriate for each error source.
    - ❏ Construct a composite statistical distribution for each error component based on its source distributions.
5) Develop a total error statistical distribution from the distributions for each error component.
6) Compute confidence limits, expectation values, measurement decision risks, etc. using the total error statistical distribution.

The Error Model

The error model should describe how error components combine to produce the total error of a measurement result. As an example, consider the determination of particle velocity (v) through measurements of time (t) and distance (d). We first define the measurement with the familiar relation $v = d/t$. If errors are represented by the symbol ε then, if errors in time are small compared to the magnitude of the time measurement itself, the appropriate error model is

$$\begin{aligned} v + \varepsilon_v &= \frac{d + \varepsilon_d}{t + \varepsilon_t} \\ &\cong \frac{d}{t}\left(1 + \varepsilon_d / d\right)\left(1 - \varepsilon_t / t\right) \\ &\cong v\left(1 + \varepsilon_d / d - \varepsilon_t / t\right), \end{aligned}$$

and

$$\varepsilon_v \cong v\left(\frac{\varepsilon_d}{d} - \frac{\varepsilon_t}{t}\right)$$
$$= \varepsilon_1 + \varepsilon_2 ,$$

where

$$\varepsilon_1 = v\varepsilon_d / d$$

and

$$\varepsilon_2 = -v\varepsilon_t / t .$$

Note that the same expressions result from using the conventional Taylor series expansion for small measurement errors [19]:

$$\varepsilon_v = \left(\frac{\partial v}{\partial d}\right)\varepsilon_d + \left(\frac{\partial v}{\partial t}\right)\varepsilon_t .$$

In general, if the determination of a given quantity is based on a set of *n* measured attributes, the total error of the quantity can be expressed in the functional relationship

$$\begin{aligned}\varepsilon_{total} &= \varepsilon_{total}(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \\ &= \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_n .\end{aligned} \qquad (1)$$

As with all measurement errors, each of the variables ε_i is composed of both process errors e_p (physical discrepancies between measurement results and true measurand values) and errors of perception e_o (discrepancies between measurement results and the perception of these results):

$$\varepsilon_i = \varepsilon_i(e_p, e_o), \quad i = 1, 2, \cdots, n . \qquad (2)$$

Steps three and four of the methodology described in this paper involve determining the statistical distributions for each error component and using these component distributions to form a statistical distribution for the total error. Returning to the particle velocity example, the statistical distribution for ε_v can be obtained from a joint distribution for ε_1 and ε_2. Representing this joint distribution by the probability density function (pdf) $f(\varepsilon_1, \varepsilon_2)$, the pdf for ε_v can be found using

$$f(\varepsilon_v) = \int_{-\infty}^{\infty} d\varepsilon_1 f(\varepsilon_1, \varepsilon_v - \varepsilon_1) . \qquad (3)$$

In cases where the error components are s-independent, as is commonly the case, this expression becomes

$$f(\varepsilon_v) = \int_{-\infty}^{\infty} d\varepsilon_1 f_1(\varepsilon_1) f_2(\varepsilon_v - \varepsilon_1) , \qquad (4)$$

where $f_1(\cdot)$ and $f_2(\cdot)$ are the pdfs for the individual error components ε_1 and ε_2. In this example, these pdfs are related to the pdfs for distance and time according to

$$f_1(\varepsilon_1) = \frac{d}{v} f_{\varepsilon_d}(\varepsilon_1 d / v) , \qquad (5)$$

and

$$f_2(\varepsilon_2) = \frac{t}{v} f_{e_t}(-\varepsilon_2 t / v) , \qquad (6)$$

The remainder of this paper focuses on the construction of pdfs for individual error components. As Eqs (1) through (6) indicate, once these pdfs are obtained, a pdf for total measurement error can be developed. Using the total error pdf, a description of total measurement uncertainty becomes possible.

To illustrate, suppose that errors in distance are normally distributed around the distance measurement with standard deviation σ_d, while time measurements are uniformly distributed within $\pm\tau$ of the time measurement. Then

$$f_{\varepsilon_d}(\varepsilon_d) = \frac{1}{\sqrt{2\pi}\sigma_d} e^{-(\varepsilon_d - d)^2/2\sigma_d^2}$$

and

$$f_{\varepsilon_t}(\varepsilon_t) = \begin{cases} 1/2\tau, & t-\tau \le \varepsilon_t \le t+\tau \\ 0, & \text{otherwise}. \end{cases}$$

Eqs (5) and (6) yield

$$f_1(\varepsilon_1) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(\varepsilon_1 - v)^2/2\sigma_1^2}$$

where $\sigma_1^2 = (v/d)\sigma_d$, and

$$f_2(\varepsilon_2) = \begin{cases} t/2v\tau, & -(v/t)(t+\tau) \le \varepsilon_2 \le -(v/t)(t-\tau) \\ 0, & \text{otherwise}. \end{cases}$$

Substituting these pdfs in Eq. (4) gives

$$\begin{aligned} f(\varepsilon_v) &= \frac{1}{\sqrt{2\pi}\sigma_1}\frac{t}{2v\tau}\int_{\varepsilon_v+\frac{v}{t}(t-\tau)}^{\varepsilon_v+\frac{v}{t}(t+\tau)} e^{-(\varepsilon_1 - v)^2/2\sigma_1^2}\, d\varepsilon_1 \\ &= \frac{t}{2v\tau}\left[\Phi\left(\frac{\varepsilon_v + v\tau/t}{\sigma_1}\right) - \Phi\left(\frac{\varepsilon_v - v\tau/t}{\sigma_1}\right)\right], \end{aligned}$$

where the function Φ is the cumulative normal distribution function defined by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\varsigma^2/2}\, d\varsigma .$$

Accounting for Process Error

Process error e_p arises from errors in the measurement system (e_{ms}), from the measuring environment (e_e), and from the set-up and configuration of the measurement system (e_s):

$$\begin{aligned} e_p &= e_p(e_{ms}, e_e, e_s) \\ &= e_{ms} + e_e + e_s . \end{aligned} \tag{7}$$

In Eq (7), the subscripts *ms*, *e* and *s* refer to "measuring system," "environment," and "set up," respectively. Measurement system and environmental process errors are broken down into a bias (b) and a random error (ε). Set-up error is conceived as constituting a bias only:

$$\begin{aligned} e_{ms} &= b_{ms} + \varepsilon_{ms} \\ e_e &= b_e + \varepsilon_e \\ e_s &= b_s . \end{aligned} \tag{8}$$

In discussing given measurement situations, the "true" value of the measurand (attribute being measured) will be denoted x and the measured value (measurement result) will be labeled y. Thus the system measures the value x and returns the result y. A measurement result returned by the measuring system can be described by a statistical distribution which is conditional on both the measurand's value and on the measurement process errors. Such a statistical distribution is described by the "conditional" pdf $f(y\,|\,x, e_p)$. This function is read "f of y given x and e_p." It represents the probability of obtaining a measurement result y, given a true value x and a process error e_p.

In a typical measuring situation, the process error e_p is not known (nor is the true value x), and the measuring individual or other "operator" (such as an automated control system) will not be able to obtain the function $f(y\,|\,x, e_p)$ explicitly. Instead, what could be attempted is an estimate of a corresponding function $f(y\,|\,x)$ that is an "average" or "expectation value" for $f(y\,|\,x, e_p)$. The probability density function $f(y\,|\,x)$ is obtained by averaging over ranges of values accessible to e_{ms}, e_e and e_s (the sources of e_p). The averaging process is described in Appendix A.

Obtaining information about e_{ms}, e_e and e_s and constructing the functional form of $f(y\,|\,x)$ is accomplished through a structured question and answer process to be reported in detail in a future paper. Briefly, the process consists of extracting all known engineering and other technical knowledge about the attribute under consideration and the measuring system and environment. In some cases, access to test and calibration history data bases is also involved. Recent experience with a prototype test and calibration management decision support system [8,9] suggests that the process of constructing $f(y\,|\,x)$ can be implemented in a user-interactive computer workstation environment.

Accounting for Perception Error

Once a measuring system returns a result, the result is perceived by the operator. This perception is usually subject to error. Perception errors arise in a number of ways. For example, in reading an analog meter, errors due to discrepancies between the operator's vantage point and the nominal meter reading position may arise (parallax errors). In reading a ruler, weighing device or digital voltmeter, errors due to discrepancies between the measurand's value and the measuring system's nominal scale or readout points often occur (resolution errors). The reader can readily imagine other examples.

Thus, the perceived or "reported" result may differ from the result y returned by the measurement system. These differences are assumed to be distributed around the value of y and are said to be conditional on this value. Thus, denoting the perceived result by the variable z, this distribution is given by the function $f(z\,|\,y)$. If the pdfs $f(y\,|\,x)$ and $f(z\,|\,y)$ can be determined, then the distribution of perceived results around the true value of the measurand can be constructed. As one might suspect, this pdf is denoted $f(z\,|\,x)$. The construction of $f(z\,|\,x)$ is described in Appendix A.

Measurement Uncertainty Estimation

The pdf $f(z\,|\,x)$ provides a description of the probabilities associated with obtaining perceived or reported values z, given that the true value being measured is x. Both measurement process errors and perception errors influence the characteristics of $f(z\,|\,x)$.

Determination of Confidence Limits for z

Estimating statistical confidence limits in the measurement of a quantity is a major facet of conventional uncertainty analysis methods. As discussed earlier, most conventional methods (which assume normal error distributions) conclude by forming normal or Student's t confidence limit estimates based on measurement variance.

The method described in this paper takes a more general tack by employing the pdf $f(z\,|\,x)$ directly rather than by merely focusing on one of its parameters (i.e., the variance). This permits uncertainty estimation in cases afflicted with non-normally distributed errors, as is shown in Appendices A and B. Unlike conventional methods, statistical confidence limits for z are obtained through integration of $f(z\,|\,x)$ directly. This does not involve the usual process of attempting to base confidence limits on some multiple of the standard deviation in z.

Estimation of the Measurand Value *x*

Appendix A shows how the method reported in this paper can also be used to estimate values for the measurand *x* based on the measurement *z*, the process error e_p and the perception error ε_0. This feature is unavailable with conventional methods.

Determination of Confidence Limits for the Measurand

In addition to estimates of measurand value, Appendix A provides a prescription for obtaining upper and lower bounds that can be said to contain the measurand value with a given level of statistical significance. This is another feature of the new method that has been previously unavailable.

Management of Measurement Decision Risks

As stated earlier, if we can estimate the probability of encountering attribute values associated with negative consequences, then we have a practical uncertainty analysis methodology. One application of such estimates is the determination of consumer and producer risk [1-7]. Consumer and producer risk can be determined through the use of $f(z\,|\,x)$ and the *a priori* distribution for *x*, $f(x)$. The procedure is outlined in Appendix A.

Conclusion

Because of its ability to unambiguously determine measurement uncertainty and to enable the effective management of this uncertainty in practical situations, the new method is decidedly superior to conventional methods.

Conventional methods require less mathematical effort, but do not yield results that are generally valid. Moreover, the new method, by working directly with error source distributions, does not require the development of techniques for combining uncertainties *per se*. Consequently, it avoids philosophical difficulties that have chronically plagued conventional uncertainty analysis methodologies and have constituted a stumbling block to progress in this area.

The proliferation of desktop computing capability throughout industry has removed the primary obstacle to implementing complex mathematical methods in the work environment. Hence, there are no overriding practical reasons why the methodology described in this paper cannot be put to use by scientific and engineering personnel. Some additional work is required, however, to bring this to fruition. Future efforts are principally needed in the areas of error model development and construction of error source distributions.

Constructing Error Models

The development of applicable error models requires engineering knowledge of how measurements are made and knowledge of the sensitivity of measurement parameters to sources of error. Constructing error models based on this knowledge would involve supplying information to a user-interactive desktop application. The desktop application would then develop an appropriate configuration analysis model describing the measurement process and set-up. Once a measurement configuration model is constructed, the appropriate error model follows directly.

Development of generally applicable user-interactive desktop applications for modeling measurement configurations is currently in the progress.

Constructing Source Distributions

Once error sources are identified, their respective statistical distributions need to be determined. For some error sources, such as measuring system error, these distributions can be developed from engineering knowledge of ranges of values accessible to measurement attributes and from the results of audits or tests or from calibration history [9]. The construction of other distributions requires the application of knowledge gained from experience (e.g., testing or calibration) with attributes of interest.

The prognosis for developing tools for constructing error source distributions is good. Promising results have already been obtained during beta testing of a measurement decision risk management prototype developed by the U.S. Navy [8,9].

Generalization of the Mathematical Methods

The methodology, as described Appendices A and B, illustrates many of its concepts by obtaining results in closed form or in the form of integral equations. Implementation of the methodology would not require that this be done. Interfacing the basic methodological approach with off-the-shelf mathematical analysis software would be sufficient to employ the methodology in a completely general way, without restrictions concerning error models employed or corresponding source distributions.

Acknowledgment

The author is indebted to Dr. Klaus Jaeger of Lockheed Missiles & Space Company and to Mr. Dave Abell of Hewlett Packard for provocative discussions which helped focus the direction of research toward the methodology described in this paper. Acknowledgment is also due Professor John Ferling of Claremont McKenna College, Mr. Jerry Hayes of Hayes Technology, Mr. Del Caldwell of the Navy Warfare Assessment Center, and Dr. Ken Kuskey of Comarco for many collaborations on related measurement science and technology management subjects. I would also like to thank my wife, Ms. Suzanne Griston, for patiently wading through the first draft and providing many cogent and worthwhile suggestions.

References

[1] Hayes, J., "Factors Affecting Measuring Reliability," U.S. Naval Ordnance Laboratory, TM No. 63-106, 24 October 1955.

[2] Castrup, H., *Evaluation of Customer and Manufacturer Risk vs. Acceptance Test In-Tolerance Level*, TRW Technical Report No. 99900-7871-RU-00, April 1978.

[3] Ferling, J., *Calibration Interval Analysis Model*, SAI Comsystems Technical Report, Prepared for the U.S. Navy Metrology Engineering Center, Contract N00123-87-C-0589, April 1979.

[4] Kuskey, K., *Cost-Benefit Model for Analysis of Calibration-System Designs in the Case of Random-Walk Equipment Behavior*, SAI Comsystems Technical Report, Prepared for the U.S. Navy Metrology Engineering Center, Contract N00123-76-C-0589, January 1979.

[5] Kuskey, K. and Kraft, R., *An Analysis and Evaluation of Management and Information Systems for the Calibration and Repair of the Navy's Electronic Test Equipment*, DDI Technical Report TR 81-9-177, December 1981.

[6] Ferling, J., "The Role of Accuracy Ratios in Test and Measurement Processes," *Proceedings of the 1984 Measurement Science Conference*, Long Beach, January 19-20.

[7] Weber, S. and Hillstrom, A., *Economic Model of Calibration Improvements for Automatic Test Equipment*, NBS Special Publication 673, April 1984.

[8] Castrup, H., "Navy Analytical Metrology RD&E," *Navy Metrology Research & Development Program Conference Report*, Dept. of the Navy, Metrology Engineering Center, NWS, Seal Beach, Corona Annex, March 1988.

[9] Castrup, H., "Calibration Requirements Analysis System," *Proceedings of the 1989 NCSL Workshop and Symposium*, Denver, July 1989.

[10] Castrup, H., *NASA Handbook NHB* 5330.9(1A), *Metrology and Calibration Provisions Guidelines*, Chapters 3.3, 3.4, 4.6, 4.7, 5.6 - 5.8, 7 and Appendices B, C and D, Jet Propulsion Laboratory, August 1991.

[11] Ferling, J. and Caldwell, D., "Analytic Modeling for Electronic Test Equipment Adjustment Policies," Presentation, *1989 NCSL Workshop and Symposium*, Denver, July 1989.

[12] 'Student,' (William S. Gosset), *Biometrica*, **6**, 1908.

[13] Proschan, F., "Confidence and Tolerance Intervals for the Normal Distribution," *Proc. Annual American Statistical Association*, Boston, December 1951.

[14] Youden, W., "Uncertainties in Calibration," *Proc. 1962 International Conference on Precision Electromagnetic Measurements*, August 1962.

[15] Eisenhart, C., "Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems," J. Res. NBS, Vol. 67C, No. 2, April-June 1963.

[16] Ku, H., "Statistical Concepts in Metrology," *Handbook of Industrial Metrology*, American Society of Tool and Manufacturing Engineers, Prentice-Hall, Inc., New York, 1967.

[17] Ku, H., "Expressions of Imprecision, Systematic Error, and Uncertainty Associated with a Reported Value," *Precision Measurements and Calibration, Statistical Concepts and Procedures*, NBS Special Pub. 300, Vol. 1, Ed. by H. Ku, February 1969.

[18] Abernathy, R.B., et al., *Measurement Uncertainty Handbook*, Research Triangle Park, Instrument Society of America, 1980.

[19] Taylor, J., *Fundamentals of Measurement Error*, Neff Instrument Corporation, 1988.

[20] ISO/TAG4/WG3, *Guide to the Expression of Uncertainty in Measurement*, 1st Ed., June 1992.

[21] Taylor, B. and Kuyatt, C., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, January 1993.

Appendix A - Construction of Component Probability Density Functions

This appendix addresses the construction of pdfs for the components of error that combine to make the total error of Eq (1) in the text. If the joint pdf for component errors is $f(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$, then the pdf for the total error is given by

$$f(\varepsilon_{total}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\varepsilon_2 d\varepsilon_3 \cdots d\varepsilon_n \, f(\varepsilon_{total} - \varepsilon_2 - \cdots - \varepsilon_n, \varepsilon_2, \cdots, \varepsilon_n) \, . \quad \text{(A-1)}$$

Each of the error components is a function of both process errors, arising from various facets of the measurement process, and errors of perception, arising from the perception of measurement results. Both process errors and errors of perception are discussed in this appendix in some detail.

Given a functional form for the joint distribution, it can be constructed from knowledge of the individual pdfs of the error components. The construction of each component pdf involves several steps:

Process Error:

- ❑ Development of a process error model for each error component.
- ❑ Development of a pdf describing the distribution of measurement results, given specific process error component values.
- ❑ Determination of the expectation value for the measurement results pdf.

Perception Error:

- ❑ Development of a perception error model.
- ❑ Development of a pdf describing the distribution of perceived measurement values, given a specific measurement result.
- ❑ Determination of the expectation value for the perceived measurement values distribution.

Appendix B shows how pdfs constructed using this procedure are employed to estimate measurement uncertainty limits, measurand expectation values and measurement decision risks.

Process Error

The Process Error Model

From observed measurement results, we make inferences about the value of a given measurand and about the uncertainty in our knowledge of this value. To develop a methodological framework for making such inferences, it is helpful to view the measurand as representing some deviation from a nominal or target value.[6] In the present discussion, deviations from nominal are treated as measurement biases or errors whose description can be accomplished by constructing pdfs that represent their statistical distributions.

Knowledge of these distributions is acquired through measurement, tempered by certain *a priori* knowledge of their makeup and of the uncertainties surrounding the measurement process.

Whether the measurand is an element of a derived quantity (such as distance is an element of velocity) or stands alone as the quantity of interest, deviations of its true value from nominal are referred to herein as "error components." Errors inherent in measurements of these components are labeled process errors.

From Eqs (7) and (8), process error is given by:

$$e_p = b_{ms} + b_e + b_s + \varepsilon_{ms} + \varepsilon_e \, . \qquad \text{(A-2)}$$

Development of the Measurement Results pdf

Let the variable x represent the deviation from nominal of a measured quantity (i.e., the error component of the quantity). Development of the pdf $f(y\,|\,x)$ for results produced by the measuring system begins by viewing the measurement result within the context of a given set of process errors. The pdf is written

$$f(y\,|\,x,e_p) = f(y\,|\,x,b_{ms},b_e,b_s,\varepsilon_{ms},\varepsilon_e) \, . \qquad \text{(A-3)}$$

Determining the Expectation Value for the Measurement Results pdf

The pdf $f(y\,|\,x)$ is found by averaging the error sources in Eq (A-3) over their respective distributions.

1. General Case

The general expression for performing this average is

[6]Examples of such nominal values are the length of a yardstick, the volume of a quart of milk, and the weight of a four-ounce sinker.

$$f(y\,|\,x) = \int_{process\ errors} f(e_p)f(y\,|\,x,e_p)de_p$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} db_{ms}db_e db_s d\varepsilon_{ms}d\varepsilon_e\ f(e_p)f(y\,|\,x,b_{ms},b_e,b_s,\varepsilon_{ms},\varepsilon_e)\,. \tag{A-4}$$

2. s-Independent Sources

If the error sources are s-independent, then the joint pdf $f(y\,|\,x,\ e_p)$ is the product of the pdfs of the source distributions:

$$f(e_p) = f(b_{ms})f(b_e)f(b_s)f(\varepsilon_{ms})f(\varepsilon_e)\,. \tag{A-5}$$

With s-independent error sources, Eq (A-4) can then be solved in a straightforward manner. The order of integration is usually unimportant. For example, we might first consider measurement uncertainty due to random fluctuations in the measuring environment. These fluctuations are accounted for by averaging Eq (A-4) over the variable ε_e:

$$f(y\,|\,x,b_{ms},b_e,b_s,\varepsilon_{ms}) = \int_{-\infty}^{\infty} d\varepsilon_e f(\varepsilon_e)f(y\,|\,x,b_{ms},b_e,b_s,\varepsilon_{ms},\varepsilon_e)\,.$$

The other error sources are averaged in the same way.

The Perception Error Model

Once the measurement result y is obtained, it is perceived by the operator to have the value z. The distribution of z around y, described by the conditional pdf $f(z\,|\,y)$ can usually be determined by engineering analysis.

Determination of the pdf for Perceived Measurement Values

Using Eq (A-4), the pdfs $f(z\,|\,y)$ and $f(y\,|\,x)$ can be used to determine the pdf for observed measurements of the value of the measurand:

$$f(z\,|\,x) = \int_{-\infty}^{\infty} f(z\,|\,y)f(y\,|\,x)dy$$
$$= \int_{-\infty}^{\infty} dy \int_{process\ error} de_p f(z\,|\,y)f(y\,|\,x,e_p)\,. \tag{A-6}$$

Eq (A-6) describes a pdf for observed measurements taken on a given measurand value x. Prior to measurement, the available information on this value consists of knowing that the measurand attribute was drawn from a population of like attributes whose values are distributed according to some pdf $f(x)$. In many instances, sufficient *a priori* knowledge is available on this population to enable an approximate specification of the population's distribution prior to measurement. To illustrate, suppose the measuring situation is product acceptance testing. In this case, *a priori* knowledge of $f(x)$ can be obtained from design and manufacturing considerations and from product testing history data.

Armed with an *a priori* pdf $f(x)$, the expected distribution of observed measurements is given by

$$f(z) = \int_{-\infty}^{\infty} f(z\,|\,x)f(x)dx\,, \tag{A-7}$$

where $f(z\,|\,x)$ is given in Eq (A-6).

Inferences Concerning Measurand Values

From a measurement or a set of measurements, we can infer what the most likely distribution of values for the measurand x might be. This is the distribution that could lead to obtaining the perceived values z from measurements of x. Of course, to be precise, the measurand's value is usually a fixed, true quantity, not a distribution of values. However, this quantity is unknown. In forming an estimate of its distribution, what we are really trying to do is determine probabilities for incremental ranges or neighborhoods of values that contain the true value.

The pdf $f(x|z)$ for the distribution of values of x, given the observed measured values z, is obtained from the expression

$$f(x|z) = \frac{f(z|x)f(x)}{f(z)} . \tag{A-8}$$

The pdf $f(z|x)$ is given in Eq (A-6) and the pdf $f(z)$ is computed using Eq (A-7). The *a priori* pdf $f(x)$ is determined as described in the previous section. Eq (A-8) will be used in Appendix B to determine confidence limits for x and to estimate the most probable value for x, given a perceived measurement z.

Example: Normally Distributed s-Independent Sources

For s-independent error sources, Eq (A-5) is substituted into Eq (A-4). If all error sources are normally distributed, performing the integration yields the result

$$f(y|x) = \frac{1}{\sqrt{2\pi}\sigma_p} e^{-(y-x)^2/2\sigma_p^2} , \tag{A-9}$$

where

$$\sigma_p^2 = \sigma_{b_{ms}}^2 + \sigma_{b_e}^2 + \sigma_{b_s}^2 + \sigma_{\varepsilon_{ms}}^2 + \sigma_{e_e}^2 . \tag{A-10}$$

If errors of perception are normally distributed, as is the case with those that stem from random cognitive processes (such as parallax errors), the pdf $f(z|y)$ can be written

$$f(z|y) = \frac{1}{\sqrt{2\pi}\sigma_{\varepsilon_0}} e^{-(z-y)^2/2\sigma_{\varepsilon_0}^2} , \tag{A-11}$$

where the variable ε_0 is the (random) perception or "observation" error. Substitution of Eqs (A-11) and (A-9) in Eq (A-6) yields

$$f(z|x) = \frac{1}{\sqrt{2\pi}\sigma_m} e^{-(z-x)^2/2\sigma_m^2} , \tag{A-12}$$

where

$$\sigma_m^2 = \sigma_p^2 + \sigma_{\varepsilon_0}^2 . \tag{A-13}$$

For normally distributed measurand values, the *a priori* pdf $f(x)$ is (assuming zero population bias)

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-x^2/2\sigma_x^2} . \tag{A-14}$$

Using this expression with Eq (A-13) in Eq (A-7) gives the expected distribution of measured values:

$$f(z) = \frac{1}{\sqrt{2\pi}\sigma_z} e^{-z^2/2\sigma_z^2} , \tag{A-15}$$

where

$$\sigma_z^2 = \sigma_m^2 + \sigma_x^2 \,. \tag{A-16}$$

Combining Eqs (A-15), (A-14) and (A-12) in Eq (A-8) gives

$$\begin{aligned} f(x \mid z) &= \frac{\sigma_z}{\sqrt{2\pi}\sigma_m \sigma_x} e^{-(z-x)^2/2\sigma_m^2} e^{-x^2/2\sigma_x^2} e^{-z^2/2\sigma_z^2} \\ &= \frac{1}{\sqrt{2\pi}\sigma_{x|z}} e^{-(x-\beta z)^2/2\sigma_{x|z}^2} \,, \end{aligned} \tag{A-17}$$

where

$$\beta = \frac{1}{1 + (\sigma_m / \sigma_x)^2} \tag{A-18}$$

and

$$\sigma_{x|z} = \sqrt{\beta}\sigma_m \,. \tag{A-19}$$

From Eqs (A-9) through (A-13) it is obvious that the component pdfs obtained using the foregoing procedure could be calculated by recognizing that, it the error sources are normally distributed, the component distributions are also normal with variances equal to the sums of the variances of the error sources. This is the familiar RSS result found in many treatments on uncertainty analysis [13-17, 20]. Note that the conditions for its validity are that error sources be both s-independent and normally distributed.

For such situations, the statistical distribution construction procedure described above is pure overkill. The procedure becomes more relevant (practical) in cases where one or more error sources are not normally distributed.

Example: Mixed Error Source Distributions

Consider, for purposes of illustration, a case where all error sources are normally distributed except for perception error. An example of such a case is one where perception uncertainty is due to random fluctuations in the least significant digit of a digital device readout. In using the device, the operator obtains a perceived value *z*. If there are *k* significant digits following the decimal, then the limits of uncertainty due to the least significant digit can be expressed according to

$$y = z \pm \rho_k \,,$$

where $\rho_k = 5 \times 10^{-(k+1)}$.

The measuring system readout informs the operator that the measurement result is somewhere between $z - \rho_k$ and $z + \rho_k$ with uniform probability. The conditional distribution that applies to this uniformly distributed perception error is

$$f(z \mid y) = \begin{cases} \dfrac{1}{2\rho_k}, & y - \rho_k \le z \le y + \rho_k \\ 0, & \text{otherwise} \,. \end{cases} \tag{A-20}$$

Substitution of this pdf in Eq (A-12) yields

$$f(z\,|\,x) = \frac{1}{2\rho_k\sqrt{2\pi}\sigma_p}\int_{z-\rho_k}^{z+\rho_k} e^{-(y-x)^2/2\sigma_p^2}dy$$
$$= \frac{1}{2\rho_k}\left[\Phi\left(\frac{z-x+\rho_k}{\sigma_p}\right) - \Phi\left(\frac{z-x-\rho_k}{\sigma_p}\right)\right], \tag{A-21}$$

where the variable σ_p is defined in Eq (A-10). The function Φ is the Gaussian cumulative distribution function.

To obtain the pdf $f(z)$, rather than plugging Eq (A-21) in Eq (A-7), it is more convenient to substitute Eq (A-6) in Eq (A-7) and perform the integration over first x then y:

$$f(z) = \int_{-\infty}^{\infty} f(z\,|\,x)f(x)dx$$
$$= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, f(z\,|\,y)f(y\,|\,x)f(x)$$
$$= \frac{1}{2\rho_k}\int_{z-\rho_k}^{z+\rho_k} dy \int_{-\infty}^{\infty} dx\, f(y\,|\,x)f(x)$$
$$= \frac{1}{2\rho_k}\left[\Phi\left(\frac{z+\rho_k}{\sigma_z}\right) - \Phi\left(\frac{z-\rho_k}{\sigma_z}\right)\right], \tag{A-22}$$

where σ_z is now given by

$$\sigma_z^2 = \sigma_p^2 + \sigma_x^2\,. \tag{A-23}$$

The construction of the pdf $f(x\,|\,z)$ follows the same procedure as with normally distributed components. Using Eqs (A-14), (A-21) and (A-22) in Eq (A-8), the pdf $f(x\,|\,z)$ can be written

$$f(x\,|\,z) = \frac{f(z\,|\,x)f(x)}{f(z)}$$
$$= \left[\Phi\left(\frac{z-x+\rho_k}{\sigma_p}\right) - \Phi\left(\frac{z-x-\rho_k}{\sigma_p}\right)\right]\frac{1}{\varphi(z,\rho_k,\sigma_z)\sqrt{2\pi}\sigma_x}e^{-x^2/2\sigma_x^2}\,, \tag{A-24}$$

where σ_z is defined in Eq (A-23) and

$$\varphi(z,\rho_k,\sigma_z) = \Phi\left(\frac{z+\rho_k}{\sigma_z}\right) - \Phi\left(\frac{z-\rho_k}{\sigma_z}\right). \tag{A-25}$$

Comparing Eq (A-24) with Eq (A-17) shows that, if even a single error source is non normal, the resultant pdf may be substantially different in character than if all sources are normally distributed. This point will be returned to in Appendix B.

Appendix B - Applications

Estimating Measurement Confidence Limits

Conventional methodologies calculate statistical confidence limits for measurements by inferring these limits from computed measurement variances. Alternatively, using the methodology in this paper, statistical confidence limits for observed measurements can be estimated directly using the pdf $f(z|x)$. For a (1 - α) X 100% confidence level, the appropriate expressions are

$$\frac{\alpha}{2} = \int_{-\infty}^{L_1} f(z \mid x) dz \,, \qquad \text{(lower limit)} \qquad \text{(B-1)}$$

and

$$\frac{\alpha}{2} = \int_{L_2}^{\infty} f(z \mid x) dz \,. \qquad \text{(upper limit)} \qquad \text{(B-2)}$$

Estimating Measurand Values

In making measurements, we are often primarily interested in ascertaining an estimate of the value of the measurand and in obtaining some confidence that this estimate is sufficiently accurate to suit our purposes. Extension of the foregoing methodology enables meeting this objective.

In making this extension, we employ the pdf $f(x \mid z)$ to obtain a statistical expectation value for *x*, given a perceived measurement result *z*. The relevant expression is

$$\langle x \mid z \rangle = \int_{-\infty}^{\infty} x f(x \mid z) \, dx \,. \qquad \text{(B-3)}$$

Estimating Confidence Limits for *x*

The conditional pdf $f(x \mid z)$ can be used to find upper and lower bounds for a neighborhood of measurand values that contains the true value of the measurand with a specified level of confidence. If this level of confidence is (1- α) × 100%, then the confidence limits L_1 and L_2 for *x* are found by solving

$$\begin{aligned} \frac{\alpha}{2} &= \int_{L_2}^{\infty} f(x \mid z) \, dx \\ &= \int_{-\infty}^{L_1} f(x \mid z) \, dx \,. \end{aligned} \qquad \text{(B-4)}$$

Estimating Measurement Decision Risk

The analysis of risks accompanying measurement decisions is a subject of current research [8-11]. In the course of this research, two of the most powerful indicators measurement decision risk have been found to be producer and consumer risk.

Consumer risk is defined at the probability that measurements of out-of-tolerance attributes will be perceived as being in-tolerance. Producer risk is defined as the probability that measurements of in-tolerance attributes will be perceived as being out-of-tolerance. Both variables are useful indicators of the quality or accuracy of a measuring process.

If the variable *A* denotes the acceptable (in-tolerance) range of attribute values and its complement $\overline{A}$ denotes the corresponding range of out-of-tolerance values, then consumer risk (CR) and producer risk (PR) are calculated according to

$$\begin{aligned} CR &= P(z \in A, x \in \overline{A}) \\ &= P(z \in A) - P(z \in A, x \in A) \\ &= \int_A dz \, f(z) - \int_A dx \int_A dz \, f(z \mid x) f(x) \,, \end{aligned} \qquad \text{(B-5)}$$

and

$$\begin{aligned} PR &= P(z \in \overline{A}, x \in A) \\ &= P(x \in A) - P(z \in A, x \in A) \\ &= \int_A dx \, f(x) - \int_A dx \int_A dz \, f(z \mid x) f(x) \,. \end{aligned} \qquad \text{(B-6)}$$

Example: Normally Distributed s-Independent Sources

The pdfs for normally distributed s-independent sources will be employed in Eqs (B-1) through (B-6) to estimate measurement confidence limits, measurand bias, confidence limits for this bias, and consumer and producer risks accompanying measurements.

1. Measurement Confidence Limits

Substitution of Eq (A-12) in Eqs (B-1) and (B-2) gives the (1 - α) × 100% confidence limits for observed measurement z :

$$L_1 = x - \sigma_m \Phi^{-1}(1-\alpha/2) ,$$

and

$$L_2 = x + \sigma_m \Phi^{-1}(1-\alpha/2) ,$$

or, alternatively,

$$x - \sigma_m \Phi^{-1}(1-\alpha/2) \le z \le x + \sigma_m \Phi^{-1}(1-\alpha/2) . \qquad \text{(B-7)}$$

The operator $\Phi^{-1}(\cdot)$ is the inverse cumulative normal function, and the measurement standard deviation σ_m is defined in Eq (A-13).

2. Measurand Bias

By substituting Eq (A-17) into Eq (B-3), the most likely value for the measurand, given the perceived measurement result z, turns out to be

$$\langle x \mid z \rangle = \beta z . \qquad \text{(B-8)}$$

Note that, since $\beta > 1$ (unless $\sigma_m = 0$), the magnitude of the maximum likelihood estimate of x is larger than the magnitude of z. This can be understood by recalling that the variable x is being treated as a deviation from nominal, and noting that normally distributed measurements tend to regress toward nominal. With these considerations in mind, it can be anticipated that the maximum likelihood estimate of the true deviation from nominal would be larger than the perceived or measured deviation from nominal.

It should be pointed out that the process of estimating a maximum likelihood value for an attribute involves both measuring the attribute and making *a priori* statements about its distribution. If, in the development of Eq (A-17), a nonzero mean value had been specified in the *a prior* distribution of x, then the resultant maximum likelihood value would have been centered around the nonzero mean value (i.e., away from nominal).

3. Measurand Confidence Limits

Upper and lower confidence limits for the measurand are obtained by substituting $f(x|z)$ from Eq (A-17) in Eq (B-4). The result is

$$\beta z - \sigma_{x|z} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \le x \le \beta z + \sigma_{x|z} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right). \qquad \text{(B-9)}$$

4. Consumer/Producer Risk

To simplify the discussion, assume that the acceptance region for attribute deviations from nominal, represented by the variable x, is symmetrical about zero, i.e., that A=[-L, L]. From Eqs (B-5) and (B-6), consumer risk and producer risk are given by

$$CR = P(z \in A) - P(z \in A, x \in A) , \qquad \text{(B-10)}$$

and

$$PR = P(x \in A) - P(z \in A, x \in A) . \qquad \text{(B-11)}$$

The component parts of these relations are easily calculated. From Eq (A-15),

$$P(z \in A) = 2\Phi\left(\frac{L}{\sigma_z}\right) - 1\,, \tag{B-12}$$

where σ_z is defined in Eq (A-16). From Eq (A-12), the joint probability for both *z* and *x* lying within *A* is given by

$$\begin{aligned} P(z \in A, x \in A) &= \int_{-L}^{L} dz \int_{-L}^{L} dx\ f(z\,|\,x) f(x) \\ &= \frac{1}{\sqrt{2\pi}\sigma_x}\left[\Phi\left(\frac{L+x}{\sigma_m}\right) + \Phi\left(\frac{L-x}{\sigma_m}\right) - 1\right] e^{-x^2/2\sigma_x^2}\,, \end{aligned} \tag{B-13}$$

where σ_m is given in Eq (A-13). Finally, using Eq (A-14) yields

$$P(x \in A) = 2\Phi\left(\frac{L}{\sigma_x}\right) - 1\,. \tag{B-14}$$

Equations (B-12) and (B-13) are substituted into Equation (B-10) to get an estimate of consumer risk. Equations (B-13) and (B-14) are substituted into Equation (B-11) to get the corresponding producer risk.

Example: s-independent Error Sources with Mixed Distributions

As in Appendix A, the example for cases involving mixed distributions considered here is one in which perception errors are uniformly distributed, and errors from all other sources are normally distributed.

1. Measurement Confidence Limits

The same procedure is used to estimate confidence limits for mixed distribution error sources as for normally distributed error sources. For uniformly distributed errors of perception, the lower and upper confidence limits can be obtained from

$$\begin{aligned} \frac{\alpha}{2} &= \int_{-\infty}^{L_1} f(z\,|\,x) dz \\ &= \int_{-\infty}^{L_1} dz \int_{-\infty}^{\infty} dy\ f(z\,|\,y) f(y\,|\,x) \\ &= \int_{-\infty}^{L_1} dy\ f(y\,|\,x) \int_{y-\rho_k}^{y+\rho_k} dz\ f(z\,|\,y) + \int_{L_1-\rho_k}^{L_1+\rho_k} dy\ f(y\,|\,x) \int_{y-\rho_k}^{L_1} dz\ f(z\,|\,y) \\ &= \frac{1}{2\rho_k}\left\{ (L_1+\rho_k-x)\Phi\left(\frac{L_1+\rho_k-x}{\sigma_p}\right) - (L_1-\rho_k-x)\Phi\left(\frac{L_1-\rho_k-x}{\sigma_p}\right) \right. \\ &\qquad \left. + \frac{1}{\sqrt{2\pi}}\left[e^{-(L_1+\rho_k-x)^2/2\sigma_p^2} - e^{-(L_1-\rho_k-x)^2/2\sigma_p^2} \right]\right\}. \end{aligned} \tag{B-15}$$

and

$$\begin{aligned}
\frac{\alpha}{2} &= \int_{L_2}^{\infty} f(z\,|\,x)dz \\
&= \int_{L_2}^{\infty} dz \int_{-\infty}^{\infty} dy\; f(z\,|\,y) f(y\,|\,x) \\
&= \int_{L_2-\rho_k}^{L_2+\rho_k} dy\; f(y\,|\,x) \int_{L_2}^{y+\rho_k} dz\; f(z\,|\,y) + \int_{L_2+\rho_k}^{\infty} dy\; f(y\,|\,x) \int_{y-\rho_k}^{y+\rho_k} dz\; f(z\,|\,y) \qquad \text{(B-16)} \\
&= \frac{1}{2\rho_k} \left\{ (L_2 + \rho_k - x)\Phi\left(\frac{L_2 + \rho_k - x}{\sigma_p}\right) - (L_2 - \rho_k - x)\Phi\left(\frac{L_2 - \rho_k - x}{\sigma_p}\right) \right. \\
&\qquad \left. + \frac{1}{\sqrt{2\pi}} \left[e^{-(L_2+\rho_k-x)^2/2\sigma_p^2} - e^{-(L_2-\rho_k-x)^2/2\sigma_p^2} \right] \right\}.
\end{aligned}$$

Solving for L_1 and L_2 from Eqs (B-15) and (B-16) requires the use of numerical or graphical methods.

2. Measurand Bias Estimate

For the present example, the expectation value for the measurand is obtained from

$$\begin{aligned}
\langle x\,|\,z \rangle &= \int_{-\infty}^{\infty} x\; f(x\,|\,z)dx \\
&= \int_{-\infty}^{\infty} x\, \frac{f(z\,|\,x) f(x)}{f(z)} dx \\
&= \frac{1}{f(z)} \int_{-\infty}^{\infty} x\; f(x)dx \int_{-\infty}^{\infty} f(z\,|\,y) f(y\,|\,z)dy \\
&= \frac{1}{f(z)} \int_{-\infty}^{\infty} dy\; f(z\,|\,y) \int_{-\infty}^{\infty} x\; f(y\,|\,x) f(x)dx\,.
\end{aligned}$$

Using Eqs (A-9), (A-14), (A-20), (A-22) and (A-23) and integrating gives

$$\langle x\,|\,z \rangle = \frac{2\rho_k \gamma}{\sqrt{2\pi}\sigma_z \varphi(z, \rho_k, \sigma_z)}, \qquad \text{(B-17)}$$

where

$$\gamma = \frac{1}{1 + (\sigma_p / \sigma_x)^2}. \qquad \text{(B-18)}$$

3. Measurand Confidence Limits

Upper and lower confidence limits are calculated for this example by numerically or graphically solving the following expressions for L_1 and L_2

$$\frac{\alpha}{2} = \frac{1}{\varphi(z, \rho_k \sigma_z)\sqrt{2\pi}\sigma_x} \int_{-\infty}^{L_1} \left[\Phi\left(\frac{z - x + \rho_k}{\sigma_p}\right) - \Phi\left(\frac{z - x - \rho_k}{\sigma_p}\right) \right] e^{-x^2/2\sigma_x^2}\, dx \qquad \text{(B-19)}$$

and

$$\frac{\alpha}{2} = \frac{1}{\varphi(z, \rho_k \sigma_z)\sqrt{2\pi}\sigma_x} \int_{L_2}^{\infty} \left[\Phi\left(\frac{z - x + \rho_k}{\sigma_p}\right) - \Phi\left(\frac{z - x - \rho_k}{\sigma_p}\right) \right] e^{-x^2/2\sigma_x^2}\, dx\,. \qquad \text{(B-20)}$$

4. Consumer/Producer Risk

As with the example of normally distributed error sources, assume that the acceptance region A in Eqs (B-5) and (B-6) is symmetrical about zero, i.e., $A = [-L, L]$. Using Eqs (A-14), (A-21) and (A-22) yields the expressions

$$P(x \in A) = 2\Phi\left(\frac{L}{\sigma_x}\right) - 1, \quad \text{(B-21)}$$

$$P(z \in A) = \frac{1}{2\rho_k}\int_{-L}^{L}\left[\Phi\left(\frac{z+\rho_k}{\sigma_z}\right) - \Phi\left(\frac{z-\rho_k}{\sigma_z}\right)\right]dz \quad \text{(B-22)}$$

and

$$P(z \in A, x \in A) = \frac{1/2\rho_k}{\sqrt{2\pi}\sigma_x}\int_{-L}^{L} dz \int_{-L}^{L} dx\left[\Phi\left(\frac{z-x+\rho_k}{\sigma_p}\right) - \Phi\left(\frac{z-x-\rho_k}{\sigma_p}\right)\right]e^{-x^2/2\sigma_x^2}. \quad \text{(B-23)}$$

Contrasting Eqs (B-22) and (B-23) with Eqs (B-12) and (B-13), respectively, shows that applying the assumption of normality to cases with mixed error component distributions may compromise the validity of measurement decision risk management.

Appendix C - Nomenclature

The following are terms and variables used in this paper. The definitions pertain to the usage of these terms and variables in this paper and do not necessarily reflect their general usage within given fields of study.

attribute - A measurable parameter or function.

measurement reliability - The probability that an attribute is in conformance with stated accuracy specifications.

total error - The total deviation from nominal of the value of an attribute.

error component - If an attribute is a function of one or more variables, the deviation from nominal of a each variable is an error component .

s-independent - Statistical independence. Two variables are said to be s-independent if the values adopted by one have no influence on the values adopted by the other.

statistical variance - The expectation value of the square of the deviation of a quantity from its mean value. A measure of the magnitude of the spread of values adopted by a variable.

population - All items exhibiting a given measurable property.

distribution - A mathematical expression describing the probabilities associated with obtaining specific values for a given attribute.

error model - A mathematical expression describing the relationship of an error to its error components.

error source - A variable that influences the value of an error component.

confidence limits - Limits which are estimated to contain a given variable with a specified probability.

expectation value - The most probable value of an attribute or variable.

measurement decision risk - The probability of an undesirable outcome resulting from a decision based on measurements.

probability density function (pdf) - A mathematical expression describing the functional relationship between a specific value of an attribute or variable and the probability of obtaining that value.

ε_{total} - Total error.

ε_i - The *ith* error component of the total error.

e_p - Measurement process error. Error due to the measuring system, environment and set-up.

e_{ms} - Error due to the measuring system.

e_e - Error due to the measuring environment.

e_s - Error due to the set-up and configuration of the measuring system.

b_{ms} - The part of measuring system error that remains fixed during a given measurement or set of measurements.

ε_{ms} - The part of measuring system error that varies randomly during a given measurement or set of measurements.

b_e - The part of measuring environment error that remains fixed during a given measurement or set of measurements.

ε_e - The part of measuring environment error that varies randomly during a given measurement or set of measurements.

b_s - Synonymous with e_s.

x - The true value of the deviation from nominal of an attribute being measured.

y - The value returned by the measuring system for a measurement of x.

z - The value of a measurement perceived or observed by the operator of the measuring system.

$f(y|x)$ - The pdf for obtaining a measured value y from a measurement of x.

$f(z|y)$ - The dpf for perceiving a measurement result z from a measured value y.

$f(z|x)$ - The pdf for a measurement result z being perceived from a measurement of x.

$f(x|z)$ - The pdf for an attribute having a value x given that its measurement is perceived to be z.

$f(x)$ - The *a priori* pdf for attribute values prior to measurement.

$f(z)$ - The pdf for perceived measurements taken on an attribute population.

L_1 - Lower confidence limit.

L_2 - Upper confidence limit.

$\langle x|z \rangle$ - The most probable value for an attribute being measured, given that its perceived measurement value is z.

CR - Consumer risk.

PR - Producer risk.

$P(z \in A)$ - The probability that measurements of an attribute will be perceived to be in conformance with stated specifications.

$P(x \in A)$ - The probability that an attribute is in conformance with specifications prior to measurement.

$P(z \in A, x \in A)$ - The probability that an attribute is in conformance with specifications and is perceived to be in conformance with specifications.

$\Phi(\cdot)$ - The cumulative normal distribution function.

$\Phi^{-1}(\cdot)$ - The inverse of $\Phi(\cdot)$.

σ_p - The standard deviation for measurement process errors.

σ_{ε_0} - The standard deviation for errors of perception.

σ_m - The standard deviation for perceived measurement results.

σ_z - The standard deviation for perceived measurement results for measurements taken on an attribute population.

$\sigma_{x|z}$ - The standard deviation for the estimated distribution of true attribute values that is most likely to produce a perceived measurement result z.

ρ_k - One half the magnitude of the maximum range of perceived values that can contain a measurement result.