Analyzing Uncertainty for Risk Management[1]

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

Summary

A structured approach to uncertainty analysis is described that is applicable to product quality assessment and risk management. Expressions are derived that incorporate estimated uncertainties in risk analyses to determine whether product parameters will be acceptable for intended applications.

Key Words

consumer's risk
error model
measurement decision risk
producer's risk
uncertainty

Introduction

In recent years, ISO and NIST guidelines (ISO 1992, Taylor 1993) and recommendations have been developed that provide a framework for analyzing and communicating measurement uncertainties. These guidelines constitute a major step in building a common analytical language for both domestic and international trade. This language is essential for a rigorous analysis of decision risks associated with measurement and testing.

Measurement Decision Risk

In measuring or testing a product, there exists the possibility that out-of-tolerance parameters will be perceived to be in-tolerance. The probability that this will happen is called *false accept risk*. False accept risk constitutes a measure of the quality of a measurement process as viewed by individuals external to the measuring organization. The higher the false accept risk, the greater the chances for returned goods, loss of reputation, litigation and other undesirable outcomes. In a commercial context, individuals external to a measuring organization are often labeled "consumers." For this reason, false accept risk has traditionally been called *consumer's risk* (Eagle 1954).

A counterpart to false accept risk is false reject risk. False reject risk is the probability that in-tolerance parameters are perceived to be out-of-tolerance. False reject risk is a measure of the quality of a measuring process as viewed by individuals within the measuring organization. The higher the false reject risk, the greater the chances for unnecessary re-work and re-test. In a commercial context, a measuring organization is labeled the "producer," and false reject risk is called *producer's risk* (Eagle 1954).

False accept risk and false reject risk, taken together, are referred to as *measurement decision risk* (Castrup 1978). This paper examines the impact of measurement uncertainty on measurement decision risk. The discussion begins with a systematic prescription for analyzing uncertainty.

Error and Uncertainty

When we measure a physical parameter by any means (e.g., eyeballing, using off-the-shelf instruments, employing precise standards, etc.) we are making an estimate of the value of the quantity being measured. Two features of such estimates are *measurement error* and *measurement uncertainty*.

---

[1] Presented at the 49th ASQC Annual Quality Congress, Cincinnati, May 1995.

Measurement Error

The difference between the value of a measured quantity and a measurement estimate of its value is referred to as *measurement error*. Measurement error may be systematic or random.

Systematic errors are classified as those whose sign and magnitude remain fixed over a specified period of time or whose values change in a predictable way under specified conditions.

Random errors are those whose sign and/or magnitude may change randomly over a specified period of time or whose values are unpredictable, given randomly changing conditions.

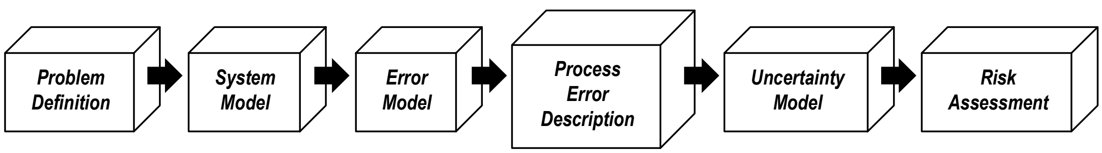
More will be said about random and systematic errors later.

Measurement Uncertainty

Measurement errors are never known exactly. In some instances they may be estimated and tolerated or corrected for. In others they may simply be acknowledged as being present. Whether an error is estimated or acknowledged, its existence introduces a certain amount of measurement uncertainty (Castrup 1992 and 1995).

Uncertainty Analysis Procedure

The analysis of measurement uncertainty and the assessment of risks associated with this uncertainty follows a procedure that is simple and straightforward. The basic framework of the procedure is illustrated below



The specific steps in the procedure are

1. Define the quantity of interest. Determine what variables need to be measured. Set quality objectives and define acceptable levels of risk.
2. Develop the system equation.
3. Develop an error model describing total error as a function of source errors.
4. Identify process error components for each source. Estimate measurement process uncertainties.
5. Estimate the total uncertainty.
6. Evaluate risks and take appropriate action.

The order in which these steps are taken is somewhat flexible. For instance, in some cases step 6 may be a prerequisite for step 1.

Problem Definition

Define the Quantity of Interest

We begin by focusing on a physical quantity y whose value is to be estimated through measurement. We identify the measurable variables $x_1, x_2, \ldots, x_n$ that are needed to estimate y.

Establish Quality Objectives

In estimating the value of y, we will need to express the quality or "accuracy" of our estimate. We do this in terms of limiting values that can be said to contain the "true" value of the quantity with some specified probability. Indeed, we *equate* the accuracy of the estimate with the limiting values and the associated containment probability. The smaller the separation of the limits and the higher the containment probability, the better the quality of the estimate.

It is important to note that, stating *either* the limiting values *or* the containment probability is *not* sufficient to quantify accuracy. Both parameters are needed.[2]

The probability of containment is called the *confidence level* and the limiting values are referred to as *confidence limits*. We say that we have a "good" estimate if it falls within appropriate confidence limits with an acceptable confidence level.

It is easy to see that a set of confidence limits and an associated confidence level can be viewed as parameters for controlling risk. Saying, for example, that there is a 98% chance that the limits $Y \pm b$ contain the quantity y, is equivalent to saying that there is an estimated 2% risk that y will be found outside $Y \pm b$.

System Modeling

We express the quantity y as a function of the n measurable variables x_i, $i = 1,2, \cdots , n$

$$y = y(x_1, x_2, \cdots, x_n) \ . \qquad (1)$$

Equation (1) is called the *system equation* for the measurement.

Error Modeling

We recognize that each variable in the system equation is a potential source of error. Accordingly, we develop the error model by expanding Eq. (1) in a Taylor series

$$\varepsilon(y) = \sum_{i=1}^{n} \left(\frac{\partial y}{\partial x_i} \right)_0 \varepsilon(x_i) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial^2 y}{\partial x_i \partial x_j} \right)_0 \varepsilon(x_i)\varepsilon(x_j) + \cdots , \qquad (2)$$

where the notation $\varepsilon(\cdot)$ represents the error in the bracketed variable, and the zero subscripts indicate that the variables (error sources) x_i and x_j are to be taken at their nominal (or "errorless") values for the bracketed partial derivatives. In cases where the errors in measurement are small, the second and higher order terms may be dropped, and Eq. (2) becomes

$$\varepsilon(y) \cong \sum_{i=1}^{n} \left(\frac{\partial y}{\partial x_i} \right)_0 \varepsilon(x_i) \ . \qquad (3)$$

Eq. (3) is the error model for the determination of the quantity y. The partial derivatives serve as weighting coefficients for the error sources

$$c_i = \left(\frac{\partial y}{\partial x_i} \right)_0 , \ \ i = 1,2, \cdots ,n \ , \qquad (4)$$

so that

$$\varepsilon(y) \cong \sum_{i=1}^{n} c_i \varepsilon(x_i) \ . \qquad (5)$$

Process Error Description

The perceived values of the variables x_i in Eq. (1) are obtained in measurement processes. If each process involves a definable set of process error components $\varepsilon_r(x_i)$, $r = 1,2, \cdots , n_i$, then the errors $\varepsilon(x_i)$ can be expressed as

$$\varepsilon(x_i) = \sum_{r=1}^{n_i} \varepsilon_r(x_i) = \sum_{r=1}^{n_i} \varepsilon_{ir} \ , \qquad (6)$$

where we use the notation $\varepsilon_{ir} = \varepsilon_r(x_i)$. The variable ε_{ir} is the *rth* process error component of the *ith* error source.

---

[2]This has implications for equipment tolerances. Tolerances expressed only in terms of limiting values, without associated containment probabilities, are essentially meaningless numbers.

Process Error Components

It has been found useful to break process error components down as follows:

Error Component	Description
Subject Parameter Bias	Systematic discrepancy between the "true" value and the nominal or reading value of a parameter being measured.
Measuring Parameter Bias	Systematic discrepancy between the "true" value and the nominal or reading value of a parameter performing a measurement.
Subject Parameter Random	Random fluctuations in the value of a parameter being measured.
Measuring Parameter Random	Random fluctuations in the value of a parameter performing a measurement.
Subject Parameter Resolution	Error due to the finite precision with which values of a parameter being measured can be perceived.
Measuring Parameter Resolution	Error due to the precision with which values of a parameter performing a measurement can be perceived.
Data Acquisition	Error due to acquiring data from measurements. Includes data sampling error, computation or "round off" error and operator bias.
Stress Response	Error due to stresses of shipping and handling of an item following measurement. Stress response error is important in cases where a measured parameter's value is reported externally and the measured item is physically moved from the measurement environment to another location.
Environment/Ancillary Equipment	Error due to environmental factors or to ancillary equipment, such as temperature monitoring devices.
Other	Error due to sources peculiar to a given measurement scenario.

Measurement Process Uncertainty Analysis

Statistical (Category A) Estimates

A statistical estimate of random uncertainty may be made on the basis of a sample of measurement data. Random uncertainties are due to random fluctuations in measurements made with a measuring parameter or to random fluctuations in the value of the subject parameter. If a sample of n measurements yields the values $X_1, X_2, \cdots, X_n$, then the random uncertainty in a measurement X is estimated by the sample standard deviation

$$u_{ran} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2} \ , \tag{7}$$

where

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \ . \tag{8}$$

In cases where the estimate u_{ran} is said to represent the uncertainty in the mean value $\bar{X}$ rather than the uncertainty in a single measurement X, the applicable expression is

$$u_{ran} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(X_i - \bar{X})^2} \ . \tag{9}$$

An uncertainty estimate obtained by statistical sampling is called a **Category A estimate**.

Category A Degrees of Freedom

The degrees of freedom ν for an uncertainty estimate based on a sample of size n is given by

$$\nu = n - 1 .$$

Heuristic (Category B) Estimates

Estimating an uncertainty for errors from a given source where data samples are not available is very simple and straightforward. Such an estimate can be obtained from heuristic limits that serve as bounds for errors from the source together with an estimate for the probability that these limits are expected to contain the errors.

For this reason, it is tempting to refer to the heuristically estimated limits as "confidence limits." However, confidence limits are quantities that are usually determined statistically. To avoid confusing heuristic limits with statistical confidence limits, we sometimes call them "error limits" or "containment limits." The probability that they contain errors from a source of interest is labeled the "containment probability."

For normally distributed errors, with an estimated containment probability P and symmetric two-sided error limits $\pm L$, the uncertainty in the errors is obtained from

$$u = \frac{L}{\Phi^{-1}\left(\frac{1+P}{2}\right)} , \tag{10}$$

where the function $\Phi^{-1}(\cdot)$ is the inverse normal distribution function.[3] Likewise, the uncertainty in normally distributed errors bounded by a lower or upper single-sided limit, L_1 or L_2, may be estimated from

$$u = \frac{L_1}{\Phi^{-1}(1-P)} \quad \text{or} \qquad u = \frac{L_2}{\Phi^{-1}(P)} \ . \tag{11}$$

Not all errors are normally distributed. For instance, if errors are uniformly distributed with 100% containment limits $\pm L$, the uncertainty estimate is

$$u = \frac{L}{\sqrt{3}} . \tag{12}$$

Other distributions are also possible (Castrup 1992, ISG 1992-1994).

There are several ways to determine heuristic uncertainty estimates. In many cases, the above prescription is applicable. In other cases, uncertainties may be estimated directly without recourse to containment limits or containment probabilities. However it is arrived at, a heuristic uncertainty estimate is termed a **Category B estimate**. It has been argued that Category B estimates are sometimes just as rigorous as statistical estimates (Castrup 1995).

Category B Degrees of Freedom

A method is available for computing degrees of freedom for Category B estimates (ISO 1992). The use of this method requires estimating what are essentially confidence limits for the estimated error limits. This is a refinement that may be beyond the scope of what can be done in practical situations. Moreover, since Category B estimates are heuristic in character and, therefore, of a somewhat subjective nature anyway, embellishing a Category B estimate with such a refinement is rarely beneficial in the first place. Consequently, the degrees of freedom for Category B estimates may be regarded as infinite in most cases without incurring any loss of credibility.

---

[3]The normal distribution function is given by

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

Uncertainty Modeling

Statistical Variance

A useful variable in analyzing measurement uncertainty is statistical variance. The variance in a measurement may be thought of as the mean square error in the measurement. It is defined in terms of an *expectation value*. That is, if the mean or expected value of a variable x is μ_x, then the variance in x is given by

$$\mathrm{var}(x) = E[(x-\mu_x)^2], \tag{13}$$

where the function $E(\cdot)$ represents a statistical average over all possible values of x. If x represents a measured value, then the error in x is written

$$\varepsilon(x) = x - \mu_x, \tag{14}$$

and

$$\begin{aligned} \mathrm{var}[\varepsilon(x)] &= E[\varepsilon(x)^2] \\ &= E[(x-\mu_x)^2] \\ &= \mathrm{var}(x). \end{aligned} \tag{15}$$

Eq. (15) is an important axiom in analyzing uncertainties due to measurement...

> **Axiom 1:** The variance in the measured value of a quantity is equal to the variance in the measurement error for the quantity.

Combining Variances

There is a simple rule that governs the variance of the sum of two quantities x_1 and x_2. This rule states that, if a and b are constants (or "coefficients"), then

$$\mathrm{var}(ax_1 + bx_2) = a^2\,\mathrm{var}(x_1) + b^2\,\mathrm{var}(x_2) + 2ab\,\mathrm{cov}(x_1, x_2) . \tag{16}$$

The term $\mathrm{cov}(x_1,x_2)$ is the "covariance" of x_1 and x_2 defined by

$$\mathrm{cov}(x_1, x_2) = E[(x_1-\mu_1)(x_2-\mu_2)] . \tag{17}$$

Extending Eqs. (15) and (16) to Eq. (5) gives

$$\mathrm{var}[\varepsilon(y)] \cong \sum_{i=1}^{n} c_i^2\,\mathrm{var}[\varepsilon(x_i)] + 2\sum_{i=1}^{n}\sum_{j>i} c_i c_j\,\mathrm{cov}[\varepsilon(x_i), \varepsilon(x_j)] . \tag{18}$$

We now introduce the notation $\sigma_y = \mathrm{var}(y)$ and $\sigma_i = \mathrm{var}(x_i)$. With this notation, substitution of Eq. (15) in Eq. (18) gives

$$\sigma_y^2 \cong \sum_{i=1}^{n} c_i^2\sigma_i^2 + 2\sum_{i=1}^{n}\sum_{j>i} c_i c_j\,\mathrm{cov}(x_i, x_j) . \tag{19}$$

Statistical Independence

If variables exhibit no tendency to vary together, they are said to be *statistically independent*. If two variables are statistically independent , their covariance is zero.

Correlated Coefficients

If two error sources are *not* statistically independent, then we say that they are correlated. This correlation is quantified by a *correlation coefficient*. A correlation coefficient of +1 means that two quantities vary in perfect step with one another. If one goes up or down by a certain amount, the other goes up or down by a proportional amount. A correlation coefficient of zero means that two quantities are statistically independent. A correlation coefficient of

-1 means that two quantities vary in perfect step with one another but in opposite directions. If one goes up or down by a certain amount, the other goes down or up by a proportional amount. The correlation coefficient can be either estimated heuristically or computed from statistical samples (Castrup 1995, ISG 1994).

The correlation coefficient ρ_{ij} for two error sources or error components x_i and x_j is defined by the relation

$$\rho_{ij} = \frac{\operatorname{cov}(x_i, x_j)}{\sigma_i \sigma_j}, \tag{20}$$

where σ_i and σ_j are the standard deviations in x_i and x_j. Using this definition, we can rewrite Eq. (19) as

$$\sigma_y^2 \cong \sum_{i=1}^{n} c_i^2 \sigma_i^2 + 2\sum_{i=1}^{n} \sum_{j>i} c_i c_j \rho_{ij} \sigma_i \sigma_j . \tag{21}$$

Statistical Uncertainty - Standard Deviations

When we sample a value of a random variable, we obtain a number that may take on a range of values. In many cases, the range of values accessible to a variable is infinite. This does not mean, however, that all values are equally likely. For the most part, sampled values of a random variable tend to be distributed about some mean or mode value. The statistic that quantifies this spread is the **standard deviation**. The standard deviation is just the square root of the variance

$$\sigma_y \cong \sqrt{\sum_{i=1}^{n} c_i^2 \sigma_i^2 + 2\sum_{i=1}^{n} \sum_{j>i} c_i c_j \rho_{ij} \sigma_i \sigma_j} . \tag{22}$$

In general, the greater the spread, the larger the standard deviation. This means that, with large standard deviations, errors tend not to be "localized," i.e., the confidence with which they are known tends to be low. Equating the word "confidence" with the less precise but more comfortable word "certainty," we argue that the standard deviation in the measurement of a quantity is synonymous with the quantity's uncertainty. To make the equivalence clearer, we will henceforth denote standard deviations and other uncertainties by the letter u, rather than the symbol σ. Making this substitution in Eq. (22) gives

$$u \cong \sqrt{\sum_{i=1}^{n} c_i^2 u_i^2 + 2\sum_{i=1}^{n} \sum_{j>i} c_i c_j \rho_{ij} u_i u_j} , \tag{23}$$

Where the "y" subscript has been dropped for simplicity. Note that, since measurement uncertainty is equated with the square root of measurement variance, Axiom 1 can be restated as

> **Axiom 2:** The uncertainty in the measured value of a quantity is equal to the uncertainty in the measurement error for the quantity.

What this axiom tells us is that, since measurement error is the discrepancy between the actual value of a parameter and a perceived or measured value, we can think of measurement uncertainty as either a lack of knowledge concerning the value of a measured parameter or as a lack of knowledge concerning the error in this value.

The latter view is useful for two reasons. First, it sets the focus of uncertainty analysis directly on the distributions of measurement errors, rather than on the distributions of variables participating in measurement. Second, it forces us to a realization that, even though the error from a given source may be fixed over the course of a measurement, the *uncertainty* in this error is a statistical quantity that we estimate from *a priori* or other knowledge. This allows us to avoid being sidetracked by semantics over whether errors are random or systematic. Rather than bogging down in idle philosophical speculation, we attempt immediately to determine the vital statistics of each error distribution of interest and to estimate the uncertainties due to the relevant error sources.

Combining Uncertainties

From Eq (23) we write

$$u^2 = \sum_{i=1}^{n} c_i^2 u_i^2 + 2\sum_{i=1}^{n}\sum_{j>i} c_i c_j \rho_{ij} u_i u_j \,. \tag{24}$$

If the *ith* error source is subject to n_i process errors, then Eqs. (16) and (6) give

$$u_i^2 = \sum_{r=1}^{n_i} u_{ir}^2 + 2\sum_{r=1}^{n_i}\sum_{q>r} \rho_{irq} u_{ir} u_{iq} \,, \tag{25}$$

where the coefficient ρ_{irq} is the correlation coefficient between the *rth* and *qth* process error components of the *ith* error source.

Cross Correlations

In some cases, the process error of one error source is correlated with the process error of another. The correlation coefficients ρ_{ij} are thus influenced by these correlations. We can derive an expression for ρ_{ij} from the definition of covariance. From Eqs. (60) and (17), the covariance in two errors ε_i and ε_j is given by

$$\begin{aligned} \mathrm{cov}(\varepsilon_i, \varepsilon_j) &= E(\varepsilon_i \varepsilon_j) \\ &= \rho_{ij} u_i u_j \,, \end{aligned} \tag{26}$$

so that

$$\rho_{ij} = \frac{E(\varepsilon_i \varepsilon_j)}{u_i u_j} \,. \tag{27}$$

Substituting from Eq. (6) in Eq. (27) yields

$$\begin{aligned} \rho_{ij} &= \frac{1}{u_i u_j} E\left(\sum_{r=1}^{n_i} \varepsilon_{ir} \sum_{q=1}^{n_j} \varepsilon_{jq} \right) \\ &= \frac{1}{u_i u_j} \sum_{r=1}^{n_i}\sum_{q=1}^{n_j} E(\varepsilon_{ir} \varepsilon_{jq}) \\ &= \frac{1}{u_i u_j} \sum_{r=1}^{n_i}\sum_{q=1}^{n_j} \rho_{ijrq} u_{ir} u_{jq} \,, \end{aligned} \tag{28}$$

where ρ_{ijrq} is the correlation coefficient between the *rth* process error component of the *ith* error source and the *qth* process error component of the *jth* error source. Substituting Eq. (28) in Eq. (24) gives the total uncertainty as

$$u^2 = \sum_{i=1}^{n} c_i^2 u_i^2 + 2\sum_{i=1}^{n_i}\sum_{j>i} c_i c_j \sum_{r=1}^{n_i}\sum_{q=1}^{n_j} \rho_{ijrq} u_{ir} u_{jq} \,, \tag{29}$$

where u_i^2 is given in Eq. (25).

Risk Evaluation

After computing the uncertainty in a quantity of interest, we can establish total error confidence limits or containment limits. Such limits may serve as factors for setting design goals for product parameters to ensure that acceptable levels of false accept or false reject risk are maintained.

Category A Confidence Limits

Determining confidence limits for a Category A uncertainty estimate is a straightforward and simple exercise. For normally distributed errors, confidence limits are determined under the assumption that the uncertainty estimate is *t*-distributed with ν degrees of freedom.

When an uncertainty estimate is a combination of uncertainty estimates for *n* sources of error, each with n_i normally distributed process error components, $i = 1,2, \cdots, n$, the degrees of freedom is determined from the Welch-Satterthwaite formula

$$\nu_i = \frac{u_i^4}{\sum_{r=1}^{n_i} \frac{u_{ir}^4}{\nu_{ir}}} , \tag{30}$$

and

$$\nu = \frac{u^4}{\sum_{i=1}^{n} \frac{u_i^4}{\nu_i}} , \tag{31}$$

where u is the combined uncertainty estimate and u_i and ν_i, $i = 1,2, \cdots, n$, are the uncertainty estimates and degrees of freedom, respectively, for the *ith* error source.

Once the degrees of freedom has been established for an uncertainty estimate u, two-sided confidence limits $\pm L$, corresponding to an error containment probability P, are determined according to

$$L = t_{\alpha,\nu} u , \tag{32}$$

where the variable α is given by

$$\alpha = (1 + P)/2 . \tag{33}$$

Similarly, lower and upper one-sided confidence limits L_1 and L_2 can be obtained from

$$L_1 = t_{1-P,\nu} u \text{ and } \quad L_2 = t_{P,\nu} u . \tag{34}$$

Note:
The assumption of a *t*-distributed total uncertainty breaks down if any component of the total process uncertainty is not normally distributed or approximately normally distributed. In this event, the validity of a confidence limit for the total uncertainty is compromised (although the estimate of the uncertainty itself remains valid). The extent of this compromise depends on the number of non-normal components and the departure from normality in each component.

It should be mentioned that the assumption of a *t*-distributed total uncertainty is ordinarily justifiable. Nevertheless, general computing methods are currently being developed to enable convolving the distributions of process uncertainty components into a total uncertainty distribution from which generally valid confidence limits can be determined (ISG 95).

Category B Confidence Limits

As argued earlier, the effective degrees of freedom for Category B can usually be set to infinity. For normally distributed errors and an uncertainty estimate *u* with infinite degrees of freedom, two-sided error limits $\pm L$ are determined from

$$L = \Phi^{-1}\left(\frac{1+P}{2}\right) u , \tag{35}$$

Single-sided lower and upper error limits L_1 and L_2 are obtained using

$$L_1 = \Phi^{-1}(1-P) u \text{ and } \quad L_{2,P} = \Phi^{-1}(P) u . \tag{36}$$

For uniformly distributed errors that are symmetrically distributed around a quantity a and bounded by limits $a \pm L$, the two-sided error limits $\pm L$ are given by $L = PL$.

Risk Analysis and Uncertainty Estimates

In conventional analyses of consumer and producer risk, expressions for false accept and false reject risk for normally distributed attributes are computed from joint probabilities functions (NASA 1994, Hayes 1955 - Deaver 1993). These probability functions are obtained by numerically integrating cumulative normal distribution functions over ranges of subject attribute or measurement attribute values.

This paper describes how this numerical integration can be avoided under certain conditions by substituting cumulative distribution functions developed from a simple convolution of subject parameter and measurement parameter attributes.

Matching Attributes Risk Analysis

Single-Sided Risk

Suppose we have two variables x and w distributed according to

$$f(x) = \frac{1}{\sqrt{2\pi}u_x} e^{-(x-a)^2/2u_x^2} \text{ and } \qquad g(w) = \frac{1}{\sqrt{2\pi}u_w} e^{-(w-b)^2/2u_w^2} \; . \tag{37}$$

The quantities u_x and u_w are the uncertainty estimates for the values of x and w. These estimates are based on error source uncertainties for x and w that are, in turn, composed of appropriate combinations of process uncertainty estimates.

Imagine that the variable w represents an artifact dimension and that the variable x represents a companion artifact dimension. For instance, x could be the outside diameter of a bolt and w the inside diameter of a nut. The quantity (w - x) is the nominal "play" between nut and bolt. Obviously, the nut will not fit the bolt if $w < x$.

The risk that this will happen is given by

$$\begin{aligned} P(w < x) &= \int_{-\infty}^{\infty} dx\, f(x) \int_{-\infty}^{x} dw\, g(w) \\ &= \frac{1}{\sqrt{2\pi}u_x} \int_{-\infty}^{\infty} \Phi\left(\frac{x-b}{u_w}\right) e^{-(x-a)^2/2u_x^2} dx \\ &= 1 - \frac{1}{\sqrt{2\pi}u_x} \int_{-\infty}^{\infty} \Phi\left(\frac{b-x}{u_w}\right) e^{-(x-a)^2/2u_x^2} dx \; . \end{aligned} \tag{38}$$

We now express the same risk through the use of the distribution of the variable $z = x - w$. By convolving x and w, we obtain this distribution as

$$h(z) = \frac{1}{\sqrt{2\pi}u} e^{-[z-(b-a)]^2/2u^2} \; , \tag{39}$$

where

$$u^2 = u_x^2 + u_w^2 \; . \tag{40}$$

Since P($w < x$) = P(z < 0), we have

$$\begin{aligned} P(w < x) = P(z < 0) &= \int_{-\infty}^{0} h(z) dz \\ &= 1 - \Phi\left(\frac{b-a}{\sigma}\right). \end{aligned} \tag{41}$$

Equating Eq. (41) with Eq. (38) gives the result

$$\frac{1}{\sqrt{2\pi}u_x}\int_{-\infty}^{\infty}\Phi\left(\frac{b-x}{u_w}\right)e^{-(x-a)^2/2u_x^2}dx = \Phi\left(\frac{b-a}{u}\right). \tag{42}$$

Equivalence with Uncertainty Analysis

From Eq. (42), we surmise that setting a confidence level for the risk management of two normally distributed variables with means a and b and uncertainties u_x and u_w is equivalent to setting a confidence level for a normally distributed variable z with mean b - a and variance $u^2 = u_x^2 + u_w^2$.

To see how this works, suppose that a manufacturer of ordnance claims that the inside diameter of his cannons follows a $N(a,u_x^2)$ distribution. If we want to make cannon balls for these cannons, we need to be confident that our cannon balls will not get stuck in his barrels. In other words, we need to control to some small amount α the probability that the diameter w of one of our cannon balls will be larger than one of his inside cannon diameters. Suppose that we settle on α = 0.01, i.e., a 99% chance that this won't happen.

To control risks to α, we need to know the uncertainty in our manufacturing process. Assume that the combined cannon ball manufacturing error is a $N(0,u_w^2)$ variable and that the degrees of freedom for u_x are large enough that we can use Eq. (36) with $P = 1 - \alpha$ to set an upper confidence limit for cannon ball diameters

$$L = \Phi^{-1}(1-\alpha)u\ . \tag{43}$$

We next move the mean diameter of our cannon ball manufacturing process to the point $b = a - L$ and get a $1 - \alpha$ probability that cannon balls will not be too big. To verify this, we express the probability that a cannon's inside diameter w will be less that a cannon ball diameter x

$$\begin{aligned} P(w<x) &= \int_{-\infty}^{\infty}dx\,f(x)\int_{-\infty}^{x}dw\,g(w) \\ &= 1-\frac{1}{\sqrt{2\pi}u_x}\int_{-\infty}^{\infty}\Phi\left(\frac{x-b}{u_w}\right)e^{-(x-a)^2/2u_x^2}dx\ . \end{aligned}$$

We now substitute from Eq. (42) to get

$$P(w<x) = 1-\Phi\left(\frac{b-a}{u}\right) = \Phi\left(\frac{L}{u}\right).$$

Substituting for L from Eq. (43) gives

$$P(w<x) = \Phi\left[\frac{\Phi^{-1}(1-\alpha)u}{u}\right] = \Phi\left[\Phi^{-1}(1-\alpha)\right] = 1-\alpha \quad \text{(QED)}\ .$$

Two-Sided Risk

The foregoing dealt with analyzing the risk of whether a given attribute was smaller (or larger) than another attribute. We now look at the problem of analyzing the probability of whether a given attribute lies within certain limits that bound another.

Let x represent the value of the second attribute and let w represent the value of the first. If the limits L_1 and L_2 bound the acceptance region for the first attribute then the second attribute is acceptable if

$$x - L_1 \le w \le x + L_2\ .$$

Suppose that we want to implement the following probabilistic constraints:

$$P(w < x - L_1) = \alpha_L \qquad \text{and} \qquad P(w > x + L_2) = \alpha_U\ .$$

If x and w are $N(a,u_x^2)$ and $N(b,u_w^2)$ variables, respectively, then

$$
\begin{aligned}
P(w < x - L_1) &= \int_{-\infty}^{\infty} dx\, f(x) \int_{-\infty}^{x-L_1} g(w)dw \\
&= 1 - \frac{1}{\sqrt{2\pi}u_x} \int_{-\infty}^{\infty} \Phi\left(\frac{x - L_1 - b}{u_w}\right) e^{-(x-a)^2/2u_x^2} dx \qquad (44) \\
&= \alpha_L ,
\end{aligned}
$$

and

$$
\begin{aligned}
P(w > x + L_2) &= \int_{-\infty}^{\infty} dx\, f(x) \int_{x+L_2}^{\infty} g(w)dw \\
&= 1 - \frac{1}{\sqrt{2\pi}u_x} \int_{-\infty}^{\infty} \Phi\left(\frac{x + L_2 - b}{u_w}\right) e^{-(x-a)^2/2u_x^2} dx \qquad (45) \\
&= \alpha_U ,
\end{aligned}
$$

We return to Eq. (42) and rewrite it as

$$
\frac{1}{\sqrt{2\pi}u_x} \int_{-\infty}^{\infty} \Phi\left(\frac{x - b}{u_w}\right) e^{-(x-a)^2/2u_x^2} dx = 1 - \Phi\left(\frac{b-a}{u}\right), \qquad (46)
$$

where, as before, $u^2 = u_x^2 + u_w^2$. Substituting Eq. (46) in Eqs. (44) and (45) gives

$$
\alpha_L = \Phi\left(\frac{L_1 + (b-a)}{u}\right) \text{and} \qquad \alpha_U = \Phi\left(\frac{L_2 - (b-a)}{u}\right) . \qquad (47)
$$

Eq. (47) can be used to solve for tolerance limits for *w*:

$$
\begin{aligned}
b_L &= a - L_1 + \Phi^{-1}(\alpha_L)u \\
b_U &= a + L_2 - \Phi^{-1}(\alpha_U)u .
\end{aligned} \qquad (48)
$$

To illustrate the use of these expressions, we continue with the cannon/cannon ball problem. In this particular application, we don't want cannon balls to be too small ($w < x - L_1$), because this reduces their range. On the other hand, we don't want cannon balls too large ($w > x + L_2$), because this gets them stuck in the cannon. The first alternative reduces cannon performance, while the second eliminates it. Accordingly, we want to assign different levels of acceptable risk to each.

Suppose that, because of engineering considerations L_1 = 0.100 cm and L_2 = 0.000 cm for a cannon of diameter a = 10 cm. Suppose further that we settle on risks of α_L = 0.05 and α_U = 0.01. Then, since Φ^{-1}(0.05) = -1.64485, and Φ^{-1} (0.01) = -2.32635, we have

$$
b_L = [9.900 \text{ cm} - (1.64485)u]
$$

and

$$
b_U = [10.000 \text{ cm} + (2.32635)u] .
$$

Once the uncertainty *u* in the cannon ball manufacturing process is determined, the tolerance limits follow immediately.

Conclusion

In employing uncertainty estimates in risk analyses, it is important to identify all relevant sources of measurement error. This is facilitated by first developing an error model. The error model provides a framework for analyzing total uncertainty and assists in identifying error sources. Following the identification of error sources, each source is decomposed into its constituent process error components and the uncertainty in each component is estimated. The component uncertainties for each source are then combined to yield an estimate of the uncertainty due to the error in the source. Source uncertainties are next combined to yield a total measurement uncertainty.

A total uncertainty estimate can serve as figure of merit for a measurement process. In addition, a total uncertainty estimate can be employed in measurement decision risk management. The uncertainty estimate a key ingredient in the computation of false accept and false reject risk, and can also be used to determine tolerance limits that control risks to meet quality objectives.

References

ISO/TAG 4/WG 3, *Guide to the Expression of Uncertainty in Measurement*, June 1992.

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

Eagle, A.R., "A Method for Handling Errors in Testing and Measuring," *Industrial Quality Control*, 10-15, March 1954.

Castrup, H., "Practical Methods for Analysis of Uncertainty Propagation," *Proc. 38th Annual Instrumentation Symposium*, Las Vegas, April 1992.

Castrup, H., "Uncertainty Analysis for Risk Management," *Proc. 1995 Measurement Science Conference*, Anaheim, January 1995.

ISG, *AccuracyRatio 1.1*, Integrated Sciences Group, Bakersfield, CA 93306, 1992-1994.

ISG, *UncertaintyAnalyzer 1.0*, Integrated Sciences Group, Bakersfield, CA 93306, 1994.

Metrology — Calibration and Measurement Processes Guidelines, NASA Reference Publication 1342, Jet Propulsion Laboratory, June 1994.

Himelblau, et al., *Guidelines for Dynamic Data Acquisition and Analysis*, Jet Propulsion Laboratory, Pasadena, California, 26 April 1993.

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

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

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.

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.

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.

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

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

Castrup, H., "Calibration Requirements Analysis System," *Proc. 1989 NCSL Workshop and Symposium*, Denver, July 1989.

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

Deaver, D., "How to Maintain Your Confidence," *Proc. 1993 NCSL Workshop and Symposium*, Albuquerque, July 1993.

Deaver, D., "Guardbanding with Confidence," *Proc. 1994 NCSL Workshop and Symposium*, Chicago, July - August 1993.

ISG, 1995 In-Progress R&D, Integrated Sciences Group, Bakersfield, CA 93306.