An Examination of Measurement Decision Risk and Other Measurement Quality Metrics

Speaker/Author:
Dr. Howard Castrup
President, Integrated Sciences Group
14608 Casitas Canyon Rd
Bakersfield, CA 93306
hcastrup@isgmax.com

Abstract

Probability density functions are developed for false accept risk, false reject risk and other measurement quality metrics (MQMs). An examination of these functions casts doubt on whether the false accept risk definition commonly associated with the 2% risk requirement of Z540.3 [1] is an effective MQM from the equipment user's perspective. Also examined are various additional MQMs for consideration as well as pdfs representing post-test distributions. It is argued that use of the latter produces a definition of false accept risk that provides an MQM of relevance to users of tested or calibrated equipment and, as such, is better suited to the intent of Z540.3 than the prevailing definition.

Background

Work begun in the 1980s on establishing links between test and calibration support and product or "end item" performance [2] has recently been rejuvenated in efforts to develop measurement quality assurance guidelines and methods in support of NASA projects and NCSLI recommended practices.

As recent work has progressed, the subject of measurement quality metrics (MQMs) are being viewed in a new light — one that adds viable metrics to the list of traditional MQMs and provides visibility of their characteristics.

Introduction

One of the main thrusts of this paper is the examination of measurement decision risk (MDR) as expressed by false accept and false reject risks. These MQMs have long been applied as convenient metrics for measuring the quality of testing processes.[1] They provide quantifiable gauges of measurement quality relative to the specifications of attributes of items tested and have been well documented in the statistical and measurement science literature. In many statistics texts,[2] these risks are described in terms of the probabilities of *consumer* acceptance or rejection of items supplied by a *producer*. Such acceptance or rejection is based on applying hypothesis tests to fractions of defective items observed in samples taken on supplied item lots. In these

---

[1] Measurement decision risks and other such MQMs are applicable within the context of conformance testing, as defined in ASME B89.7.4.1-2005 [3]. Accordingly, unless otherwise indicated, both calibrations and tests are treated in this paper as conformance tests, and the terms "test," "tests" or "testing" are used for results obtained by either calibration labs or testing activities.

[2] See, for example, Brownlee [4].

treatments, the probability of accepting a relatively bad lot of items is called *consumers' risk* and the probability of rejecting a relatively good lot is called *producer's risk*.

A somewhat different approach to estimating consumer's risk was taken by Eagle, Grubbs and Coon in the mid 1950s [5, 6] and later applied to testing and calibration in a metrology context by Hayes [7]. With this approach, consumer's risk is defined as the joint probability that an item is both nonconforming and accepted by testing. This metric was later extended and applied by Castrup [8, 2, 9-13], Ferling [14, 15] , Ferling and Caldwell [16], Kuskey [17], Weber and Hillstrom [18], Hutchinson [19], Deaver [20]. Dobbert [21] Lira [22] and others.

Consumer's risk, as conceived in the work of Eagle, Grubbs and Coon is a meaningful MQM for consumers of goods, since both the manufactured quality of such goods and the efficacy of product screening initiate from the producer and benefit the consumer in terms of product quality. In other words, the joint probability of a product attribute being out-of-tolerance (OOT) and accepted as in-tolerance represents a measure of product quality of some relevance to both the producer and consumer.

In the context of a test and calibration support hierarchy, however, such as is maintained at separate levels with oversight by external assessment or auditing agencies, the *a priori* probability of conformance is a statistic of the workload submitted for testing, i.e., one that initiates from the consumer (equipment user) rather than from the producer (testing activity). Within this context, consumer's risk becomes a less informative MQM from the user's perspective than in the usual producer-consumer relationship.

To be of benefit to the equipment user, an MQM must be an unambiguous indicator of the quality of the product (a tested or calibrated attribute) delivered by the test or calibration activity, regardless of the *a priori* conformance probability. Likewise, the test or calibration activity needs MQMs that allow evaluating the efficacy of conformance testing regardless of this same probability.

It must be remarked that, if the producer determines equipment tolerances, testing procedures or calibration intervals, it is responsible in part for the *a priori* conformance probability and assumes to some extent the mantle of producer in the ordinary sense. If so, consumer's risk approaches the status it enjoys in the marketplace as a viable MQM. However, it is nevertheless afflicted by the same limitation as implied above, that is, it does not directly measure the *post-test* quality of the delivered product.

These considerations motivate the development of alternative MQMs, as presented in this paper and elsewhere [23, 24]. These MQMs are constructed from probability density functions (pdfs) for the probability distributions of both *a priori* and post-test attributes. The MQMs resulting from the use of these pdfs are metrics that provide specific information to equipment users and testing activities of direct value in assessing test and calibration support quality. In addition, they turn out to be important elements of end-to-end support for products or other "end items" [25, 26]. It should also be mentioned that, as the figures in Appendices A and B show, by focusing on pdfs for MQMs, we achieve graphical visibility of the probabilities of negative outcomes, such as false accept and false reject risks.

A variety of MQM pdfs are developed in this paper. The material is intended to provide methods for constructing MQMs and applying them to optimize test and calibration support end-to-end. While the requisite technical level presupposes familiarity with college-level algebra and basic calculus, with some exposure to mathematical statistics, a close following of the math is not necessary for an understanding of the premises, results and conclusions. To assist in this, examples are depicted of relevant probability functions which make use of the familiar normal distribution.

Relevant Circumstances and Outcomes

The circumstances of interest in developing MQMs for conformance testing are the occurrence or presence of the following:

- An out-of-tolerance (OOT) or "bad" UUT attribute value.
- An in-tolerance or "good" UUT attribute value.

The associated relevant outcomes are

- An attribute value obtained by measurement that is observed to be bad (OOT).
- An attribute value obtained by measurement that is observed to be good (in-tolerance).
- An attribute value accepted without adjustment or other renewal.
- An attribute value observed to require adjustment or other renewal. Sometimes referred to as a "rejected" attribute value.

Definitions

For discussion purposes, an attribute is assumed to have a specification consisting of a nominal, declared or expected value and either a single-sided tolerance limit or upper and lower tolerance limits. An attribute is also assumed to be accompanied by corresponding "acceptance" limits that bound a region of attribute values that are acceptable without adjustment or other renewal.

Definitive Numbers and Ratios

In what follows, we consider ratios of various numbers that will be used to develop probabilities and pdfs for various events and combinations of events. The numbers will be designated by the letter n with or without subscripts as the need arises. The subscript g will be used to designate a "good" attribute value and the subscript b will be used to designate a "bad" attribute value. The subscript a will be used to designate an accepted attribute and the subscript r will be used to designate a rejected value. Hence, n_g, n_b, n_a and n_r will, respectively, indicate the number of good, bad, accepted and rejected attributes, as shown in Table 1. Two letters together in a subscript indicate the occurrence of both corresponding events. Thus n_{ga}, for example, represent the number of occurrences of an attribute being both good and accepted. Note that such designations are symmetric, e.g., $n_{ag} = n_{ga}$.

Table 1.
Numbers Relevant to MDR

Quantity	Description
n	Number of UUT attributes in a tested lot
n_g	Number of in-tolerance UUT attributes in the lot of tested attributes
n_b	Number of OOT UUT attributes in the lot of tested attributes
n_a	Number of UUT attributes accepted without adjustment or other renewal by testing. Referred to as the "number accepted"
n_r	Number of UUT attributes shown by testing to require adjustment or other renewal. Referred to as the "number rejected"
n_{ga}	Number of UUT attributes in the lot of tested attributes that are both in-tolerance and accepted
n_{gr}	Number of UUT attributes in the lot of tested attributes that are both in-tolerance and rejected
n_{ba}	Number of UUT attributes in the lot of tested attributes that are both OOT and accepted
n_{br}	Number of UUT attributes in the lot of tested attributes that are both OOT and rejected

MQM Random Variables and Statistics

Random Variables

The random variables of interest in estimating MQMs are the bias of a UUT attribute, denoted x, the value of x estimated by measurement y, the bias uncertainty σ_x of the UUT attribute at the time of testing and the uncertainty σ_y of the measurement process.[3] The MQM random variables and distribution statistics employed in this paper are shown in Table 2.

Table 2
Basic MQM Variables and Distribution Statistics

Quantity	Description
x	The bias of a UUT attribute being tested or calibrated.
y	The value of x obtained by measurement.
σ_x	The *a priori* or "before-test" standard deviation of the population of values x.
σ_y	The standard deviation of the total combined measurement process error.
$-L_1$ (L_2)	The lower (upper) tolerance limit for x.
L	The interval $[-L_1, L_2]$.
$-A_1$ (A_2)	The lower (upper) "acceptance" limit for y.
A	The interval $[-A_1, A_2]$.

Measurement Scenarios

We consider two basic measurement scenarios; one in which the measurement system measures the value of a passive UUT attribute, such as a gauge block, and one in which the UUT attribute measures the value of a passive measurement reference.

[3] The subscripts x and y are generic placeholders representing the bias of a tested attribute and a measurement result, respectively. The specific subscripts used throughout the paper will vary with context.

Passive UUT Attribute

For the measurement of a passive UUT attribute, the true value of the attribute will be designated by the variable X, the nominal value will be denoted X_n, and an estimated value of X obtained by measurement will be designated by the letter Y. The measurement model is

$$Y = X + \varepsilon_m,$$

where ε_m is the error in the measurement process.[4] For passive attributes, the UUT attribute bias x is the deviation of the attribute's true value X from a nominal or declared value X_n, as shown in the expression

$$X = X_n + x.$$

Combining equations gives

$$Y = X_n + x + \varepsilon_m.$$

We now define a variable y as an estimate of the bias x such that

$$\begin{aligned} y &= Y - X_n \\ &= x + \varepsilon_m. \end{aligned}$$

From this expression, we see that, for a given UUT attribute value x, the variable y is distributed with mean value x and variance

$$\sigma_y^2 = \text{var}(y \mid x) = \text{var}(\varepsilon_m),$$

where var($y|x$) is verbally expressed as "variance of y given x."

Active UUT Attribute

In this case, measurements of the value of a passive reference are made using an active UUT attribute. The true value of the reference attribute is denoted Y, the nominal or declared value of the attribute is denoted Y_n, and the measured value of Y is denoted X. The measurement model is given by

$$X = Y + \varepsilon_m,$$

where ε_m is the measurement process error. The true value of the measurement reference can be written

$$Y = Y_n + e_Y,$$

where e_Y is the bias of the measurement reference. Substituting gives

$$X = Y_n + e_Y + \varepsilon_m.$$

The quantity obtained by measurement is the estimated bias y obtained from

$$\begin{aligned} y &= X - Y_n \\ &= e_y + \varepsilon_m. \end{aligned}$$

Note that, in this case, ε_m is composed of the bias x of the UUT measuring attribute combined with other measurement process errors. Denoting the latter as ε_p, we have

[4] For consistency with other publications, Greek characters are used to represent combined errors, while Latin characters are used to represent individual errors.

$$\varepsilon_m = x + \varepsilon_p,$$

and

$$y = e_y + \varepsilon_p + x.$$

From this, we see that, given a specific value of x, the population of y is distributed with mean value x and variance

$$\sigma_y^2 = \mathrm{var}(y \mid x) = \mathrm{var}(e_y + \varepsilon_p) .$$

In General

In either the passive or active UUT attribute case, the "before test" UUT attribute bias x is assumed to be distributed with a pdf $f(x)$, with variance σ_x^2, and the variable y is assumed to be distributed with a pdf $f(y|x)$, with variance σ_y^2.

Variations on the above measurement models are possible. In-depth descriptions, covering a variety of measurement models or "scenarios" are provided in a recent paper [13], in NCSLI RP-18 [27], in NASA 8730.19 [28] and in NASA 8730.19-4 [29].

Tolerance Limits and Acceptance Limits

In the following, a UUT attribute is considered in-tolerance if its bias lies within an interval $L = [-L_1, L_2]$ and OOT otherwise. Similarly, a UUT attribute is accepted without renewal if its observed bias lies within an interval $A = [-A_1, A_2]$, referred to as the *acceptance region*. If it lies outside this interval, it is either renewed or discarded, with renewal consisting or either adjustment or repair. Attributes whose biases are observed to lie within A are labeled "accepted" and those observed to lie outside A are labeled "rejected."[5] Definitions for tolerance limits and acceptance limits are given in Table 2.

Probability Functions

The probability of the occurrence of events or circumstances will be treated using the "experimental" definition of probability in which the probability of the occurrence of an event E is given by the relation

$$P(E) = \lim_{n\to\infty} \frac{n_E}{n} ,$$

where n_E is the number of occurrences of E out of n experimental trials. For example, suppose that, out of n tests, n_a attributes have been observed to be acceptable without renewal. Then

$$P(y \in A) = \lim_{n\to\infty} \frac{n_a}{n} ,$$

where the notation $P(y \in A)$ denotes the probability that observed y will lie within the acceptance region A. Similarly, the *a priori* UUT attribute in-tolerance probability $P(x \in L)$ is given by

$$P(x \in L) = \lim_{n\to\infty} \frac{n_g}{n} .$$

[5] It is important to note that a rejected attribute is not necessarily an OOT attribute, merely one that is observed to require a renewal action, such as adjustment or repair.

Table 3 shows the basic probability functions used in estimating various MQMs. The ratios in column 1 are shown without the associated limit functions.

Table 3.
Basic Probability Functions Used in Estimating MQM

Ratio	Probability	Description
n_g / n	$P(x \in L)$	The *a priori* or "before-test" probability that a UUT attribute is in-tolerance.
n_b / n	$P(x \notin L)$	The *a priori* UUT attribute OOT probability.
n_a / n	$P(y \in A)$	The probability of accepting attributes without adjustment or other renewal. Referred to as the "acceptance probability."
n_r / n	$P(y \notin A)$	The probability that tested attributes are observed to require adjustment or other renewal. Referred to as the "rejection probability."

Measurement Quality Metrics

MQMs are expressed as probabilities that can be used as measures of the quality of conformance testing. Twelve such probabilities are defined below.

Probability that Accepted Attributes are Good (In-tolerance)

This is the ratio of n_{ga} / n_a. This ratio can be expressed in terms of two probability functions of Table 3:

$$\frac{n_{ga}}{n_a} = \frac{P(x \in L, y \in A)}{P(y \in A)} . \tag{1}$$

This ratio can be expressed as a single probability function, rather than a ratio of probability functions. To do this, we employ a relation that is fundamental to probability theory: Let E_1 and E_2 represent two distinct events. The probability that both E_1 and E_2 will occur can be written

$$P(E_1, E_2) = P(E_1 \mid E_2) P(E_2) , \tag{2}$$

where the notation $P(E_1 \mid E_2)$ is the probability that E_1 will occur given that E_2 has occurred and $P(E_2)$ is the probability that E_2 will occur. Using this rule, the ratio in Eq. (1) can be written

$$\frac{n_{ga}}{n_a} = P(x \in L \mid y \in A) . \tag{3}$$

Probability that Good Attributes will be Accepted

The ratio for this is n_{ga} / n_g. From Table 3, we have, with the aid of Eq. (2),

$$\frac{n_{ga}}{n_g} = \frac{P(x \in L, y \in A)}{P(x \in L)} = P(y \in A \mid x \in L) . \tag{4}$$

Probability that Rejected Attributes are Good

The ratio for this is n_{gr} / n_r. From Table 3, and Eq. (2), we have

$$\frac{n_{gr}}{n_r} = \frac{P(x \in L, y \notin A)}{P(y \notin A)} = P(x \in L \mid y \notin A) . \tag{5}$$

Probability that Good Attributes will be Rejected

The ratio for this is n_{gr} / n_g. From Table 3, and Eq. (2), we have

$$\frac{n_{gr}}{n_g} = \frac{P(x \in L, y \notin A)}{P(x \in L)} = P(y \notin A \mid x \in L) . \tag{6}$$

Probability that Rejected Attributes are Bad

The ratio for this is n_{br} / n_r. From Table 3, and Eq. (2), we have

$$\frac{n_{br}}{n_r} = \frac{P(x \notin L, y \notin A)}{P(y \notin A)} = P(x \notin L \mid y \notin A) . \tag{7}$$

Probability that Bad Attributes will be Rejected

The ratio for this is n_{br} / n_b. From Table 3, and Eq. (2), we have

$$\frac{n_{br}}{n_b} = \frac{P(x \notin L, y \notin A)}{P(x \notin L)} = P(y \notin A \mid x \notin L) . \tag{8}$$

Probability that Accepted Attributes will be Bad

The ratio for this is n_{ba} / n_a. Again, using Table 3, we have, with the aid of Eq. (2),

$$\frac{n_{ba}}{n_a} = \frac{P(x \notin L, y \in A)}{P(y \in A)} = P(x \notin L \mid y \in A) . \tag{9}$$

Probability that Bad Attributes will be Accepted

The ratio for this is n_{ba} / n_b. Again, using Table 3, we have, with the aid of Eq. (2),

$$\frac{n_{ba}}{n_b} = \frac{P(x \notin L, y \in A)}{P(x \notin L)} = P(y \in A \mid x \notin L) . \tag{10}$$

The above twelve probabilities are summarized in Table 4.

Table 4.
Measurement Quality Metrics

Ratio	Metric	Description
n_{ga} / n	$P(x \in L, y \in A)$	The probability that UUT attributes are both in-tolerance and accepted without renewal.
n_{gr} / n	$P(x \in L, y \notin A)$	The probability that UUT attributes are both in-tolerance and rejected.
n_{ba} / n	$P(x \notin L, y \in A)$	The probability that UUT attributes are both OOT and accepted without renewal.
n_{br} / n	$P(x \notin L, y \notin A)$	The probability that UUT attributes are both OOT and rejected.
n_{ga} / n_a	$P(x \in L \mid y \in A)$	The probability that accepted attributes are good.
n_{ga} / n_g	$P(y \in A \mid x \in L)$	The probability that good attributes will be accepted.
n_{gr} / n_r	$P(x \in L \mid y \notin A)$	The probability that rejected attributes are good.
n_{gr} / n_g	$P(y \notin A \mid x \in L)$	The probability that good attributes will be rejected.
n_{br} / n_r	$P(x \notin L \mid y \notin A)$	The probability that rejected attributes are bad.

Ratio	Metric	Description
n_{br} / n_b	$P(y \notin A \mid x \notin L)$	The probability that bad attributes will be rejected.
n_{ba} / n_a	$P(x \notin L \mid y \in A)$	The probability that accepted attributes are be bad.
n_{ba} / n_b	$P(y \in A \mid x \notin L)$	The probability that bad attributes will be accepted.

General Probability Density Functions

The foregoing twelve MQMs can be developed using the pdfs defined in the next two sections. General expressions for these functions are given below. Appendix A presents special cases where *x* and *y* are normally distributed.

The Fundamental pdfs

The fundamental pdfs of use in developing the MQMs of Tables 3 and 4 are shown in Table 5.

Table 5

The Fundamental Probability Density Functions

Function	Description
$f(x)$	The pdf for attribute bias, represented by the random variable *x*.
$f(x, y)$	The joint pdf for the bias *x* of an attribute and the observed bias *y* obtained by measurement.
$f(y \mid x)$	The pdf for observing a value *y* given a UUT attribute bias *x*.
$f(y)$	The pdf for observed UUT attribute biases.

From the definitions in Table 5, the pdf for biases obtained by measurement is expressed as

$$f_t(y) = \int_{-\infty}^{\infty} f_t(y \mid x) f_{bt}(x) dx\,, \tag{11}$$

where the subscript *t* designates "conformance testing," and the subscript *bt* designates an *a priori* or "before test" quantity. The probability of accepting UUT attribute values without renewal is computed as the integral of $f_t(y)$ over *A*

$$P(y \in A) = \int_{-A_1}^{A_2} f_t(y) dy = \int_{-A_1}^{A_2} \int_{-\infty}^{\infty} f_t(y \mid x) f_{bt}(x) dx\, dy\,. \tag{12}$$

The pdf $f_{bt}(x)$ is related to the before-test in-tolerance probability $P(x \in L)$ according to

$$P(x \in L) = \int_{-L_1}^{L_2} f_{bt}(x) dx\,. \tag{13}$$

The MQM pdfs

In addition to the pdfs of Table 5, six pdfs are defined below that can be used to generate the twelve MQM probabilities of Table 4. These pdfs are summarized in Table 6.

$P(x \in L, y \in A)$ and $P(x \notin L, y \in A)$

The pdf for these metrics is the joint pdf

$$f(x, y \in A) = f(x \mid y \in A)P(y \in A) = \frac{f(y \in A \mid x) f_{bt}(x)}{P(y \in A)} P(y \in A)$$
$$= f_{bt}(x) f(y \in A \mid x) = f_{bt}(x) \int_{-A_1}^{A_2} f_t(y \mid x) dy. \tag{14}$$

$P(x \in L, y \notin A)$ and $P(x \notin L, y \notin A)$
The pdf for these metrics is the joint pdf

$$f(x, y \notin A) = f_{bt}(x) \left[\int_{-\infty}^{-A_1} f_t(y \mid x) dy + \int_{A_2}^{\infty} f_t(y \mid x) dy \right]$$
$$= f_{bt}(x) \left[1 - \int_{-A_2}^{-A_1} f_t(y \mid x) dy \right] = f_{bt}(x) - f(x, y \in A). \tag{15}$$

$P(x \in L \mid y \in A)$ and $P(x \notin L \mid y \in A)$
The pdf for these metrics is the conditional pdf

$$f(x \mid y \in A) = \frac{f(x, y \in A)}{P(y \in A)}.$$

Comparison with Eq. (14) shows that

$$f(x \mid y \in A) = \frac{f_{bt}(x) \int_{-A_1}^{A_2} f_t(y \mid x) dy}{P(y \in A)}, \tag{16}$$

where $P(y \in A)$ is given in Eq. (12).

$P(x \in L \mid y \notin A)$ and $P(x \notin L \mid y \notin A)$
The pdf for these metrics is the conditional pdf

$$f(x \mid y \notin A) = \frac{f(x, y \notin A)}{P(y \notin A)}.$$

From Eq. (15), we see that this can be expressed as

$$f(x \mid y \notin A) = \frac{f_{bt}(x)}{1 - P(y \in A)} \left[1 - \int_{-A_1}^{A_2} f_t(y \mid x) dy \right]. \tag{17}$$

$P(y \in A \mid x \in L)$ and $P(y \notin A \mid x \in L)$
The pdf for these metrics is the conditional pdf

$$f(y \mid x \in L) = \frac{f(x \in L, y)}{P(x \in L)} = \frac{\int_{-L_1}^{-L_2} f_t(y \mid x) f_{bt}(x) dx}{P(x \in L)} \tag{18}$$

where $P(x \in L)$ is given in Eq. (13).

$P(y \in A \mid x \notin L)$ and $P(y \notin A \mid x \notin L)$

The pdf for these metrics is the conditional pdf

$$f(y \mid x \notin L) = \frac{f(x \notin L, y)}{P(x \notin L)} = \frac{1}{1 - P(x \in L)}\left[f_t(y) - \int_{-L_1}^{-L_2} f_t(y \mid x) f_{bt}(x) dx \right]. \tag{19}$$

Table 6

The MQM Probability Density Functions

pdf	Description	Associated MQMs
$f(x, y \in A)$	The joint pdf for UUT attribute biases *x* and values of *y* observed to lie within the acceptance region.	$P(x \in L, y \in A)$ and $P(x \notin L, y \in A)$
$f(x, y \notin A)$	The joint pdf for UUT attribute biases *x* and values of *y* observed to lie outside the acceptance region.	$P(x \in L, y \notin A)$ and $P(x \notin L, y \notin A)$
$f(x \mid y \in A)$	The pdf for accepted UUT attribute biases.	$P(x \in L \mid y \in A)$ and $P(x \notin L \mid y \in A)$
$f(x \mid y \notin A)$	The pdf for rejected UUT attribute biases.	$P(x \in L \mid y \notin A)$ and $P(x \notin L, y \notin A)$
$f(y \mid x \in L)$	The pdf for accepting or rejecting in-tolerance UUT attribute biases.	$P(y \in A \mid x \in L)$ and $P(y \notin A \mid x \in L)$
$f(y \mid x \notin L)$	The pdf for accepting or rejecting OOT UUT attribute biases.	$P(y \in A \mid x \notin L)$ and $P(y \notin A \mid x \notin L)$

False Accept Risk

The estimation of false accept risk has recently become a subject of some importance in the wake of the publication of the Z540.3 standard [1]. The definition of false accept risk that was assumed during the development of the standard is the UFAR probability function defined by Eagle, Grubbs, Coon and Hayes [5-7].[6]

It has been pointed out in the literature that the consumer's risk probability function comes in at least two forms; consumer's risk from the perspective of the consumer and consumer's risk from the perspective of the producer [8].

Because of the inherent potential for confusion, the consumer's version was dubbed *conditional false accept risk* (CFAR) and the producer's version was labeled *unconditional false accept risk* (UFAR) [27-29]. The latter has also been referred to as the *probability of a false accept* or "PFA" [30].

Unconditional False Accept Risk (UFAR)

UFAR is defined as the joint probability of a UUT attribute being both out-of-tolerance (OOT) and observed during conformance testing to be in-tolerance. It is expressed as

$$\text{UFAR} = P(x \notin L, y \in A)\,. \tag{20}$$

[6] In addition to applying this definition, the Z540.3 Handbook [30] includes the CFAR definition, and what have come to be called the *Bayesian* and the *Confidence Level* risk analysis definitions [12, 13, 28, 29].

From Table 6, we see that UFAR can be developed using the joint pdf $f(x,y\in A)$ given in Eq. (14) and repeated here for convenience

$$f(x, y \in A) = f(y \in A \mid x) f_{bt}(x) = f_{bt}(x) \int_{-A_1}^{A_2} f_t(y \mid x) dy \,.$$

Joint pdfs with a single random variable, such as $f(x,y\in A)$, are not true pdfs in the usual sense, since integration over all values of the random variable, x in this case, do not yield unity.[7] For example, integrating $f(x,y\in A)$ from -∞ to ∞ yields

$$\begin{aligned}\int_{-\infty}^{\infty} f(x \mid y \in A) dx &= \int_{-\infty}^{\infty} \int_{-A_1}^{A_2} f(y \mid x) f_{bt}(x) \, dy \, dx = \int_{-\infty}^{\infty} \int_{-A_1}^{A_2} f(x, y) \, dy \, dx \\ &= \int_{-A_1}^{A_2} f(y) \, dy = P(y \in A),\end{aligned} \tag{21}$$

where $P(y\in A)$ is the probability that the UUT attribute will be accepted without renewal.

At this point, we return to remarks made in this paper's introduction and observe that, for a pdf to be useful as a tool in assessing the state of a UUT attribute from the standpoint of the UUT user, i.e., following conformance testing, it must be a true pdf that yields meaningful probability estimates in the context of usage. Since $P(y\in A)$ is a probability that has significance with respect to planning for renewal expenses and workflow management, the pdf of Eq. (14) shows that UFAR is a metric directly relevant to the conformance testing activity. It does not provide an MQM by which the attribute user can directly establish reliance on the in-tolerance probability of an attribute following testing.

The UFAR pdf of Eq. (14) is given in Eq. (A-8) of Appendix A for normally distributed x and y. Figure A-1 of Appendix A provides a plot of this pdf.

Conditional False Accept Risk (CFAR)

Conditional false accept risk (CFAR) is defined as the probability that accepted attributes are OOT. CFAR is given by

$$\begin{aligned}\text{CFAR} = P(x \notin L \mid y \in A) &= \frac{P(x \notin L, y \in A)}{P(y \in A)} \\ &= \frac{P(y \in A) - P(x \in L, y \in A)}{P(y \in A)} = 1 - \frac{P(x \in L, y \in A)}{P(y \in A)}.\end{aligned} \tag{22}$$

By comparison with Table 4, we see that the pdf for CFAR is $f(x|y\in A)$, given in Eq. (16) and repeated here for convenience

$$f(x \mid y \in A) = \frac{f_{bt}(x) \int_{-A_1}^{A_2} f_t(y \mid x) dy}{P(y \in A)} \,. \tag{23}$$

[7] The requirement that the integration equal unity is a necessary, if not sufficient, condition for pdfs. See, for example, Harris, pp 68ff [31].

To examine the validity of this function as a true pdf, we integrate over x from $-\infty$ to ∞ to get

$$\int_{-\infty}^{\infty} f(x \mid y \in A) dx = \frac{\int_{-\infty}^{\infty} f_{bt}(x) \int_{-A_1}^{A_2} f_t(y \mid x) dy\, dx}{P(y \in A)} . \quad (24)$$

By Eq. (21), we see that the numerator on the RHS of this relation is just $P(y \in A)$, so that

$$\int_{-\infty}^{\infty} f(x \mid y \in A) dx = \frac{P(y \in A)}{P(y \in A)} = 1 .$$

So, $f(x|y \in A)$ satisfies the necessary condition for a true pdf. It is given in Eq. (A-10) of Appendix A for normally distributed x and y. Figure A-2 of Appendix A provides a plot of this pdf.

False Reject Risk

When the term "false accept risk" became favored over the term "consumer's risk" [8, 14], the term "false reject risk" (FRR) began to be seen in conformance testing circles as preferable to the term "producer's risk."[8] FRR is defined as the joint probability that attributes being tested will be both in-tolerance and rejected:

$$\text{FRR} = P(x \in L, y \notin A) . \quad (25)$$

Accordingly, the expression given in Table 6 for the pdf used to compute FRR is $f(x, y \notin A)$, expressed as

$$f(x, y \notin A) = f(x \mid y \notin A) P(y \notin A) = f(y \notin A \mid x) f_{bt}(x) . \quad (26)$$

As expected, integration of x over L yields

$$\int_{-L_1}^{L_2} f(x, y \notin A) dx = P(x \in L, y \notin A) = \text{FRR} . \quad (27)$$

Note that integrating $f(x|y \notin A)$ over x from $-\infty$ to ∞ yields $P(y \notin A)$ rather than unity:

$$\int_{-\infty}^{\infty} f(x, y \notin A)\, dx = P(y \notin A) . \quad (28)$$

From this, we see that, as with the pdf $f(x, y \in A)$ in Eq. (14), $f(x, y \notin A)$ does not meet the necessary condition of a true pdf.

Moreover, $P(y \notin A)$ is a probability that has significance with respect to costs incurred by the testing facility due to unnecessary adjustment or other corrective action and to direct and indirect administrative costs associated with OOT reports. In addition, false rejects can significantly impact perceptions of in-tolerance probability at the time of test and, consequently lead to costs due to unnecessarily reduced calibration intervals. Hence, FRR is a metric that is directly relevant to the testing facility in the form of cost of labor and parts and is relevant to the attribute

[8] Recently, the term *probability of a false reject* or "PFR" has also come into use [30].

user in the form of costs due to unnecessarily frequent calibration and corresponding equipment downtime. It is not, however, a metric that provides information to the equipment user regarding the quality of testing as embodied in the post-test UUT attribute bias distribution.

The FRR pdf is given in Eq. (A-11) of Appendix A for normally distributed x and y. Figure A-3 of Appendix A provides a plot of this pdf.

Afterword

As pointed out earlier, a given MQM is relevant to the equipment user only if its associated pdf is useful for making decisions or estimating other probabilities of significance to equipment management.

With this in mind, we recall that UFAR, the common definition of false accept risk has no direct relevance to the consumer. Its continued use in calibration and testing is due to the inertia of historical precedent rather than to its utility as a viable risk control metric in the context of the objectives of Z540.3 and other quality control standards or guidelines.[9] Given today's pressure to provide quality that is acceptable for intended usage, it would seem prudent to seek other metrics, particularly those generated from post-test pdfs.

Expressions for post-test pdfs can be found in the literature [2, 34, 28]. The fundamental relation, developed in Appendix B of this paper, is

$$f_{pt}(x) = f_t(x \mid \text{accepted}) P_t(\text{accepted}) + f_s(x \mid \text{renewed}) P_s(\text{renewed}), \tag{29}$$

where the subscript "*pt*" designates "post-test," t refers to the conformance test activity and s denotes servicing by an adjustment or repair activity. The term "renewed" applies to adjustment with or without repair. The appropriate component pdfs and probability functions are described in Appendix B.

The mathematical form of $f_{pt}(x)$, depends on whether attributes are "renewed" (adjusted or repaired) and on which activity (conformance testing, adjustment or repair) performs the renewal. The alternative "renewal policies" are accounted for in Appendix B. For each case, false accept risk is defined as

$$\text{FAR} = 1 - \int_{-L_1}^{L_2} f_{pt}(x)\, dx \tag{30}$$

Conclusion

A number of MQMs have been presented that supplement those commonly in use. The relevance of each MQM to testing activities and equipment users has been made visible by its associated pdf, as shown in the graphics of Appendix A. From the perspective of the user, it has been argued that the MQM of greatest value to users is false accept risk as defined in Eq. (30).

[9] See, for example, MIL-STD 45662A [32], Z540-1 [33] and ASME B89.7.4.1-2005 [3].

References

1. NCSLI, *Requirements for the Calibration of Measuring and Test Equipment*, ANSI/NCSL Z540.3-2006, 2006.
2. Castrup, H., "Calibration Requirements Analysis System," *Proc. 1989 NCSL Workshop and Symposium*, Denver, July 1989.
3. ASME, *Measurement Uncertainty and Conformance Testing: Risk Analysis*, B89.7.4.1-2005, ASME Technical Report, February 2006.
4. Brownlee, *Statistical Theory and Methodology in Science and Engineering*, 2nd Ed., John Wiley & Sons, Inc., New York, 1965.
5. Eagle, A., "A Method for Handling Errors in Testing and Measuring," *Industrial Quality Control*, March, 1954.
6. Grubbs, F. and Coon, H., "On Setting Test Limits Relative to Specification Limits," *Industrial Quality Control*, March, 1954.
7. Hayes, J., *Factors Affecting Measuring Reliability*, U.S. Naval Ordnance Laboratory, TM No. 63-106, October 1955.
8. Castrup, H., *Evaluation of Customer and Manufacturer Risk v. Acceptance Test In-Tolerance Level*, TRW Technical Report No. 99900-7871-RU-00, April 1978.
9. Castrup, H., "Analyzing Uncertainty for Risk Management," *Proc. 49th ASQC Annual Quality Congress*, Cincinnati, May 1995.
10. Castrup, H., "Uncertainty Analysis for Risk Management," *Proc. MSC*, Anaheim, January 27, 1995.
11. Castrup, H., "Risk-Based Control Limits," *Proc. Measurement Science Conference*, Anaheim, January 2001.
12. Castrup, H., "Risk Analysis Methods for Complying with Z540.3," *Proc. NCSLI Workshop & Symposium*, St. Paul, 2007.
13. Castrup, H., "Decision Risk Analysis for Alternative Calibration Scenarios," *Proc. NCSLI Workshop & Symposium*, Orlando, 2008.
14. 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.
15. Ferling, J., "The Role of Accuracy Ratios in Test and Measurement Processes," *Proceedings of the 1984 Measurement Science Conference*, Long Beach, CA, January 1984.
16. Ferling, J. and Caldwell, D., "Analytic Modeling for Electronic Test Equipment Adjustment Policies," Presentation, *1989 NCSL Workshop and Symposium*, Denver, CO, July 1989.
17. 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.
18. NBS Special Publication 673, *Economic Model of Calibration Improvements for Automatic Test Equipment*, Weber, S. and Hillstrom A., U.S. Department of Commerce, April 1984.
19. Hutchinson, Bill, "Setting Guardband Test Limits to Satisfy MIL-STD-45662A Requirements," *Proc. NCSL Workshop & Symposium*, 1991, pp. 305-309.

20. Deaver, D., "Guardbanding with Confidence," *Proc. NCSL Workshop & Symposium*, 1994.
21. Dobbert, M., "A Guard-Band Strategy for Managing False-Accept Risk," *Proc. NCSLI Workshop & Symposium*, Orlando, 2008.
22. Lira, I., "A Bayesian Approach to the Consumer's and Producer's Risks in Measurement," *Metrologia*, October 1999.
23. Castrup, H., "Test and Calibration Metrics", Unpublished Note, Presented to the NCSLI Metrology Practices Committee, Anaheim, January 2002.
24. Jackson, D., "Measurement Risk Analysis Methods," *Proc. Meas. Sci. Conf.*, Anaheim, January 2005.
25. Castrup, H. and Castrup, S., "Measurement Quality Assurance Principles and Methods," Presented at the Measurement Science Conference, March 2009, Anaheim.
26. Mimbs, S., "Measurement Assurance for End-Item Users," Presented at the Measurement Science Conference, March 2009, Anaheim.
27. NCSLI, *Estimation and Evaluation of Measurement Decision Risk*, Recommended Practice, RP-18, 2009.
28. NASA, NASA-HDBK-8730.19, *NASA Measurement Quality Assurance Handbook*, Chapter 2, 2009.
29. NASA, NASA-HDBK-8730.19-4, Annex 4, *Estimation and Evaluation of Measurement Decision Risk*, 2009.
30. NCSLI, *Handbook for the Application of ANSI/NCSL Z540.3-2006*, February 2009,
31. Harris, B., *Theory of Probability*, Addison-Wesley, London, 1966.
32. MIL-STD 45662A, *Calibration Systems Requirements*, U.S. Dept. of Defense, 1 August 1988.
33. ANSI/NCSL Z540-1:1994, *Calibration Laboratories and Measuring and Test Equipment – General Requirements*, July 1994.
34. JPL, *Metrology – Calibration and Measurement Processes Guidelines*, NASA Reference Publication 1342, June 1994.
35. Abramowitz, M, and Stegun, I, *Handbook of Mathematical Functions*, U.S. Dept. of Commerce Applied Mathematics Series **55**, 1972. Also available from www.dlmf.nist.gov.
36. ISG, RiskGuard 2.0, Freeware, www.isgmax.com, © Integrated Sciences Group, 2003-2009.

Appendix A – pdfs for Normally Distributed Variables

The general expressions for the pdfs developed in this paper are specialized in this appendix to apply to cases where x is $N(0, \sigma_x^2)$ and y is $N(x, \sigma_y^2)$. It is not suggested here that the use of the normal distribution applies in all situations. Rather, we employ the normal distribution to demonstrate the development of explicit pdfs and to graphically show differences between MQM pdfs.

The Basic Normal Probabilities and pdfs

The "before test" UUT attribute bias pdf $f(x)$ for normally distributed x is given by

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

For a normally distributed estimate y of the value of a given UUT attribute bias x, the pdf $f(y \mid x)$ becomes

$$f(y\mid x)=\frac{1}{\sqrt{2\pi}\sigma_y}e^{-(y-x)^2/2\sigma_y^2}\,. \tag{A-2}$$

By Eqs. (11), (A-1 and (A-2), the pdf $f(y)$ for values of x obtained by measurement is given by

$$\begin{aligned}f(y)&=\int_{-\infty}^{\infty}f(y\mid x)f(x)dx\\&=\frac{1}{2\pi\sigma_x\sigma_y}\int_{-\infty}^{\infty}e^{-(y-x)^2/2\sigma_y^2}e^{-x^2/2\sigma_x^2}dx=\frac{1}{\sqrt{2\pi}}e^{-y^2/2\sigma^2}\,,\end{aligned} \tag{A-3}$$

where

$$\sigma^2=\sigma_x^2+\sigma_y^2\,. \tag{A-4}$$

Hereafter, it will be convenient to use a function defined as

$$\varphi(x,Q,c)=\Phi\left(\frac{Q_1+x}{c}\right)+\Phi\left(\frac{Q_2-x}{c}\right)-1\,, \tag{A-5}$$

where Q represents a range of values $[-Q_1, Q_2]$ and Φ is the cumulative normal distribution function.[10]

The probabilities $P(y\in A)$ and $P(x\in L)$ are computed using Eq. (A-3) in Eq. (12) and Eq. (A-1) in Eq. (13)

$$P(y\in A)=\int_{-A_1}^{A_2}f(y)dy=\frac{1}{\sqrt{2\pi}\sigma}\int_{-A_1}^{A_2}e^{-y^2/\sigma^2}dy=\varphi(0,A,\sigma)\,, \tag{A-6}$$

and

$$P(x\in L)=\int_{-L_1}^{L_2}f(x)dx=\frac{1}{\sqrt{2\pi}\sigma_x}\int_{-L_1}^{L_2}e^{-x^2/2\sigma_x^2}dx=\varphi(0,L,\sigma_x)\,. \tag{A-7}$$

Normally Distributed MQM pdfs

In the following expressions, the general MQM pdfs presented in this paper are used to develop pdfs for normally distributed x and y. Using Eqs. (A-1), (A-2) and (A-5) in Eq. (14), we have

[10] See, for example, Abramowitz and Stegun [35]. Φ is also accessible as a "workbook" function in popular spreadsheet applications.

$$f(x, y \in A) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-x^2/2\sigma_x^2} \int_{-A_1}^{A_2} e^{-(y-x)^2/2\sigma_y^2} dy$$
$$= \frac{1}{2\pi\sigma_x} e^{-x^2/2\sigma_x^2} \int_{-(A_1+x)/\sigma_y}^{(A_2-x)/\sigma_y} e^{-\zeta^2/2} d\zeta = \frac{\varphi(x, A, \sigma_y)}{\sqrt{2\pi}\sigma_x} e^{-x^2/2\sigma_x^2}. \quad \text{(A-8)}$$

Using Eqs. (A-1), (A-2) and (A-5) in Eq. (15) gives

$$f(x, y \notin A) = \frac{1 - \varphi(x, A, \sigma_y)}{\sqrt{2\pi}\sigma_x} e^{-x^2/2\sigma_x^2}. \quad \text{(A-9)}$$

By Eqs. (16), (A-5) and (A-8), we have the conditional pdf

$$f(x \mid y \in A) = \frac{f(x, y \in A)}{P(y \in A)} = \frac{\varphi(x, A, \sigma_y)}{P(y \in A)} \frac{e^{-x^2/2\sigma_x^2}}{\sqrt{2\pi}\sigma_x}, \quad \text{(A-10)}$$

where $P(y \in A)$ is given in Eq. (A-6).

Eqs. (17) and (A-2) yield

$$f(x \mid y \notin A) = \frac{f(x, y \notin A)}{P(y \notin A)} = \frac{1 - \varphi(x, A, \sigma_y)}{1 - P(y \in A)} \frac{e^{-x^2/2\sigma_x^2}}{\sqrt{2\pi}\sigma_x}, \quad \text{(A-11)}$$

where use was made of the relation $P(y \notin A) = 1 - P(y \in A)$.

Combining Eqs. (18), (A-1) and (A-2) gives

$$f(y \mid x \in L) = \frac{f(x \in L, y)}{P(x \in L)} = \frac{1}{P(x \in L)} \int_{-L_1}^{L_2} f(y \mid x) f(x) dx$$
$$= \frac{1}{P(x \in L)} \frac{1}{2\pi\sigma_x\sigma_y} \int_{-L_1}^{L_2} e^{-(y-x)^2/2\sigma_y^2} e^{-x^2/2\sigma_x^2} dx.$$

if we define

$$\Sigma = \sigma_x\sigma_y / \sigma \quad \text{(A-12)}$$

and

$$\beta = (\sigma_x / \sigma)^2, \quad \text{(A-13)}$$

then, using Eq. (A-5), the expression for $f(y|x \in L)$ can be written

$$f(y \mid x \in L) = \frac{1}{P(x \in L)} \frac{e^{-y^2/2\sigma^2}}{2\pi\sigma\Sigma} \int_{-L_1}^{L_2} e^{-(x-\beta y)^2/2\Sigma^2} dx$$
$$= \frac{\varphi(\beta y, L, \Sigma)}{P(x \in L)} \frac{e^{-y^2/2\sigma^2}}{\sqrt{2\pi}\sigma} = \frac{f(y)}{P(x \in L)} \varphi(\beta y, L, \Sigma). \quad \text{(A-14)}$$

Using Eqs. (A-1) and (A-2) in Eq. (19), we have

$$f(y \mid x \notin L) = \frac{f(x \notin L, y)}{1 - P(x \in L)} = \frac{1}{1 - P(x \in L)}\left[f(y) - f(x \in L, y) \right]$$
$$= \frac{1 - \varphi(\beta y, L, \Sigma)}{1 - P(x \in L)} f(y) = \frac{1 - \varphi(\beta y, L, \Sigma)}{1 - P(x \in L)} \frac{e^{-y^2/2\sigma^2}}{\sqrt{2\pi}\sigma}, \tag{A-15}$$

where β and Σ are given in Eqs. (A-12) and (A-13).

Plots of pdfs for Normally Distributed x and y

Shown in the following plots are the six pdfs for cases where L = [-10,10], $A = L$, the UUT attribute before test in-tolerance probability is 0.85 — corresponding to a standard deviation σ_x of 6.947 — and the total measurement process error has a standard deviation σ_y of 1.428, which yields 92% confidence limits (expanded uncertainties) of ±2.5.

These parameters represent a kind of common baseline for cases where the ratio σ_x/σ_y is 4:1. Using the prescription of Z540.3, this corresponds to a TUR of 3.57:1. The unconditional false accept risk (UFAR) is $P(x \notin L, y \in A)$ = 1.9292%, the conditional false accept risk (CFAR) is 2.2926% and the false reject risk (FRR) is 2.7817%.[11]

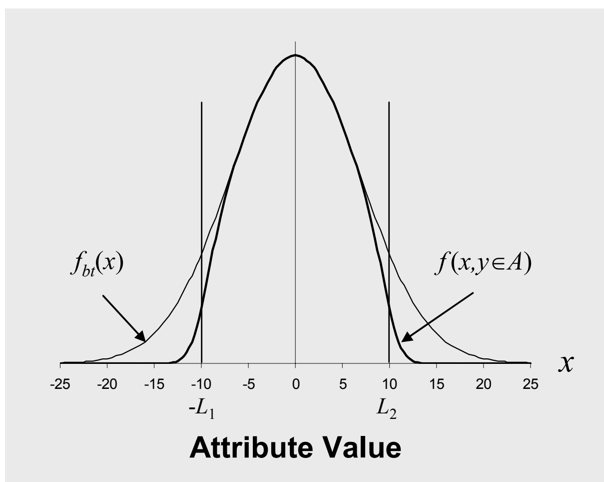

Figure A-1. The pdf $f(x, y \in A)$ for the event that an attribute will have the value x and will be accepted by testing vs. the before-test pdf $f_{bt}(x)$. The probability $P(x \in L, y \in A)$ is computed by integrating the pdf over $[-L_1, L_2]$. UFAR, defined as the probability $P(x \notin L, y \in A)$, is equal to the combined area under the dark-lined curve outside $[-L_1, L_2]$.

[11] UFAR, CFAR and FRR were computed using the RiskGuard freeware application [36].

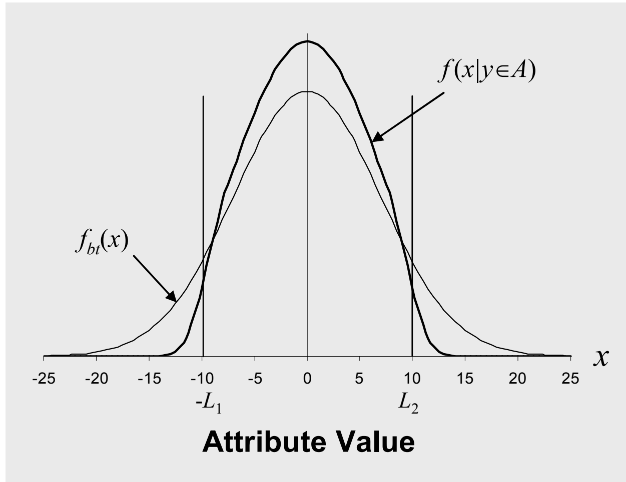

Figure A-2. The pdf $f(x|y\in A)$ for attribute values that have been accepted without renewal vs. the before-test pdf $f_{bt}(x)$. The probability $P(x\in L|y\in A)$ is computed by integrating the pdf over $[-L_1, L_2]$. CFAR, defined as the probability $P(x\notin L|y\in A)$, is equal to the combined area under the dark-lined curve outside $[-L_1, L_2]$.

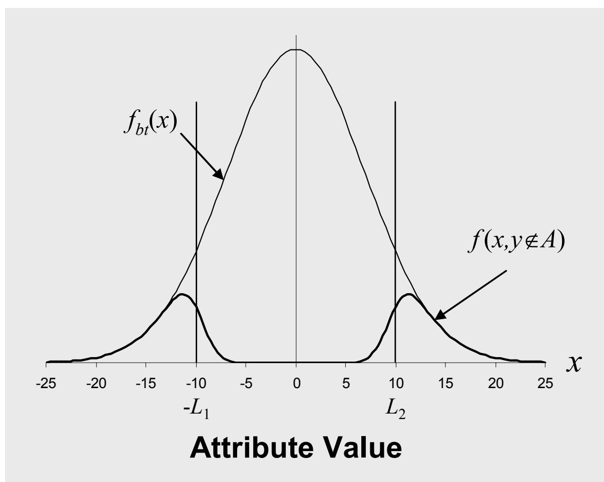

Figure A-3. The pdf $f(x, y\notin A)$ for an attribute having the value x and being rejected vs. the before-test pdf $f_{bt}(x)$. FRR, defined as the probability $P(x\in L, y\notin A)$, is computed by integrating the pdf from $[-L_1, L_2]$. The probability $P(x\notin L, y\notin A)$ is equal to the combined area under the dark-line curve outside $[-L_1, L_2]$.

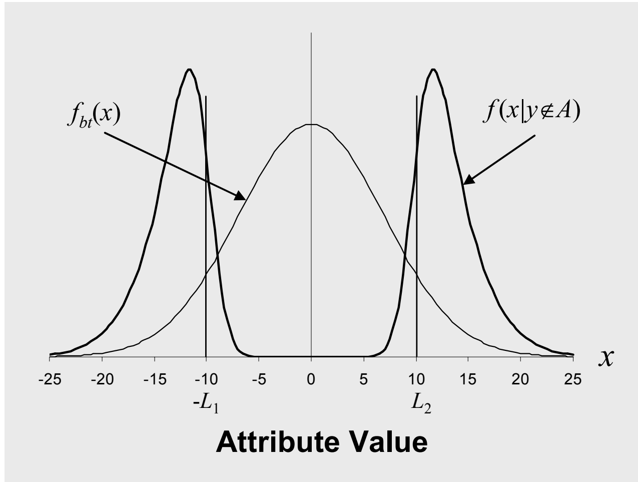

Figure A-4. The conditional pdf $f(x|y\notin A)$ for attribute values that have been rejected by testing vs. the before-test pdf $f_{bt}(x)$. The probability $P(x\in L|y\notin A)$ that rejected attributes are in-tolerance is equal to the combined area under the dark-lined curve between the limits $-L_1$ to L_2. The probability $P(x\notin L|y\notin A)$ that rejected attributes are OOT is equal to the combined area under the dark-line curve outside $[-L_1, L_2]$.

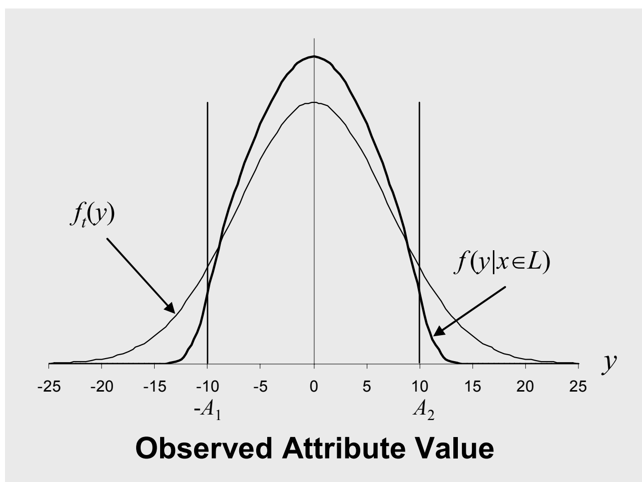

Figure A-5. The conditional pdf $f(y|x\in L)$ for observed values of in-tolerance attributes vs. the pdf $f_t(y)$. The probability $P(y\in A|x\in L)$ that in-tolerance attributes will be accepted is equal to the area under the dark-line curve between the limits $-A_1$ and A_2. The probability $P(y\notin A|x\in L)$ that in-tolerance attributes will be rejected is equal to the combined area under the dark-line curve outside $[-A_1, A_2]$.

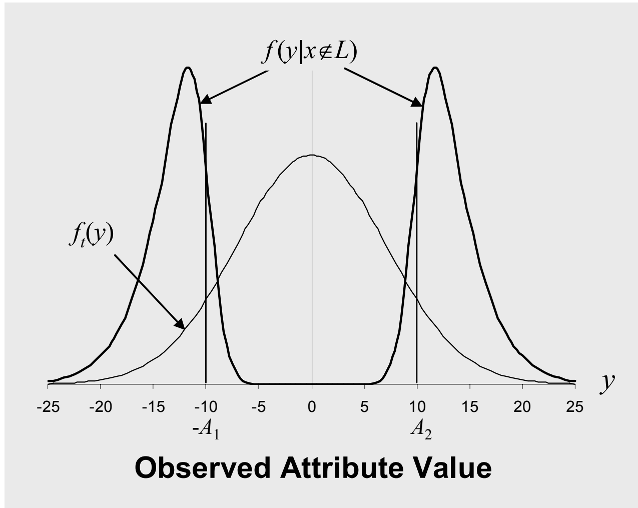

Figure A-6. The conditional pdf $f(y|x\notin L)$ for observed values of OOT attributes vs. the pdf $f_t(y)$. The probability $P(y\in A|x\notin L)$ that out-of-tolerance attributes will be accepted is equal to the area under the curve between the limits $-A_1$ and A_2. The probability $P(y\notin A|x\notin L)$ that out-of-tolerance attributes will be rejected is equal to the combined area under the curve outside $[-A_1, A_2]$ The similar appearance of $f(y|x\notin L)$ and $f(x|y\notin A)$ in Figure A-4 is due to the fact that σ and Σ are approximately equal to σ_{bt} and $P(y\in A) \cong P(x\in L)$.

Appendix B – The Post-Test Distribution

As stated in this paper's Introduction, pdfs of post-test distributions can be applied to developing MQMs that are important measures of the quality of end-to-end support to products or other "end items" [25]. Such pdfs are constructed in this appendix.

We begin with a breakdown of the pdfs of the conformance testing process, as depicted in Figure B-1. Table B-1 describes the pdf subscripts in Figure B-1 and in other graphics in this appendix.

Table B-1.
pdf Subscript References

Subscript	Reference
bt	before-test
pt	post-test
t	test
tp	test process
a	adjustment
ap	adjustment process
r	repair
rp	repair process
s	service

Each of the pdfs in Figure B-1 is described in the following sections and summarized in Table B-2. It should be noted that the diagram in Figure B-1 applies to attributes submitted for conformance testing. Attributes of items marked for repair of functional failures prior to submission for conformance testing are not included.

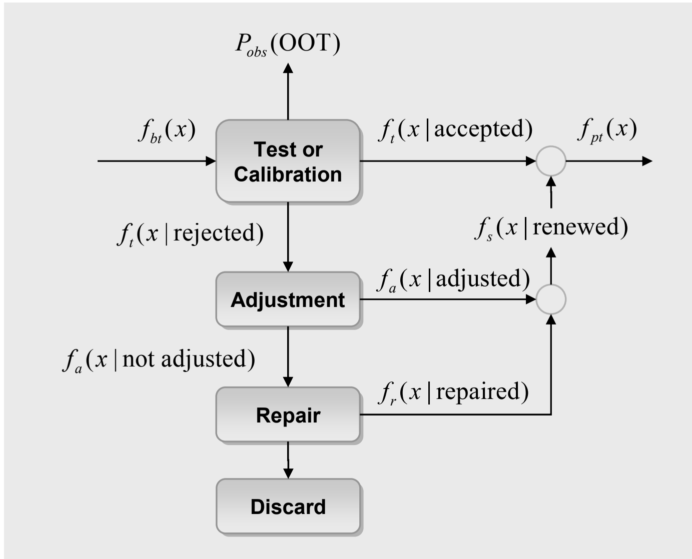

Figure B-1. The Conformance Testing Process. The distribution of the UUT attribute value x at the time of test or calibration is represented by the pdf $f_{bt}(x)$. The distribution of values emerging from test or calibration is represented by $f_{pt}(x)$. The percentage of instances in which the attribute is found out-of-tolerance is represented by the probability $P_{obs}(\text{OOT}) = P(y \notin A)$.

Table B-2

Conformance Testing Probability Density Functions

pdf	Applicable To
$f_t(x\|\text{accepted})$	Attributes accepted without adjustment or repair.
$f_t(x\|\text{rejected})$	Attributes observed to need either adjustment or repair.
$f_a(x\|\text{adjusted})$	Attributes that are adjusted to center spec.
$f_a(x\|\text{not adjusted})$	Attributes that are deemed not adjustable and are submitted for repair.
$f_r(x\|\text{repaired})$	Attributes that are repaired and adjusted to center spec.
$f_r(x\|\text{not repaired})$	Attributes that are discarded as unrepairable.
$f_s(x\|\text{renewed})$	"Serviced" attributes, i.e., those set to center spec either through normal adjustment or through adjustment following repair.
$f_{pt}(x)$	Attributes emerging from the conformance testing process.

Renewal Policies

As used in this paper, a renewal policy represents the practice followed in adjusting or otherwise correcting UUT attributes following conformance testing. Three alternative policies, *renew as-needed*, *renew always* and *no renewal* are discussed. Applicable activities, decisions and criteria are summarized in Table B-3.

Table B-3
Renewal Policy Alternatives

Activity	Decision	Criterion
Conformance Testing	Accept	The observed UUT attribute bias lies within the interval A = [$-A_1$, A_2].
	Reject	The observed UUT attribute bias lies outside the interval A.
Attribute Adjustment	Adjust	The UUT attribute bias is observed to lie within an interval R = [$-R_1$, R_2] and is considered to be correctable by adjustment.
	Don't Adjust	The UUT attribute bias is observed to lie outside the interval R.
Repair	Repair	The UUT attribute bias is correctable by repair and subsequent adjustment.
	Discard	The UUT attribute is not correctable by repair or is beyond economical repair.

Renew As-Needed

The *renew as-needed* policy applies to conformance testing, as depicted in Figure B-1, in which "failed" attributes are renewed and accepted following renewal or are discarded if renewal is not considered feasible.

Accepted Attributes

The acceptance segment of Figure B-1 involves the pdfs $f_t(x|\text{accepted})$ and $f_t(x|\text{rejected})$ defined in Table B-2. From Eq. (16), $f_t(x|\text{accepted})$ is

$$f_t(x\,|\,\text{accepted}) = f_t(x\,|\,y \in A) = \frac{f_t(y \in A\,|\,x) f_{bt}(x)}{P_t(\text{accepted})}. \tag{B-1}$$

For normally distributed variables, Eqs. (20), (21) and (24) yield

$$f_{bt}(x) = \frac{1}{\sqrt{2\pi}\sigma_{bt}} e^{-x^2/2\sigma_{bt}^2}, \tag{B-2}$$

and

$$f_t(y \in A\,|\,x) = \frac{1}{\sqrt{2\pi}\sigma_t} \int_{-A_1}^{A_2} e^{-(y-x)^2/2\sigma_t^2}\, dy = \varphi(x, A, \sigma_t), \tag{B-3}$$

where σ_{bt} and σ_t are, respectively, the standard deviations for the UUT attribute *a priori* bias and total combined measurement errors of the conformance testing process.

Rejected Attributes

From Eq. (17), $f_t(x|\text{rejected})$ is

$$f_t(x\,|\,\text{rejected}) = f_t(x\,|\,y \notin A) = \frac{f_t(y \notin A\,|\,x) f_{bt}(x)}{P_t(\text{rejected})}. \tag{B-4}$$

This pdf is given in Eq. (29) for normally distributed x and y. The probability $P_t(\text{rejected})$ is just

$$P_t(\text{rejected}) = 1 - P_t(\text{accepted}), \tag{B-5}$$

where

$$P_t(\text{accepted}) = P_t(y \in A) = \int_{-\infty}^{\infty} \int_{-A_1}^{A_2} f_t(y \mid x) f_{bt}(x)\, dy dx \,.$$

Adjusted Attributes

Noting that, in the adjustment segment of Figure B-1, attribute adjustments are made by the adjustment activity and that attribute biases are adjusted to zero (center spec), the appropriate pdf for adjusted attributes is expressed as

$$f_a(x \mid \text{adjusted}) = f_a(x) \,. \tag{B-6}$$

For a normally distributed combined measurement error, this pdf is given by

$$f_a(x) = \frac{1}{\sqrt{2\pi}\sigma_{ap}} e^{-x^2/2\sigma_{ap}^2} \,. \tag{B-7}$$

The quantity σ_{ap} represents the standard deviation of biases following adjustment, defined as

$$\sigma_{ap} = \sqrt{\sigma_a^2 + \sigma_{rb}^2} \,, \tag{B-8}$$

where σ_a is the standard uncertainty of the total measurement error in the measurement system employed in making the adjustment and σ_{rb} represents any uncertainty due to errors arising from inadvertent responses to physical adjustments, referred to as *rebound error* [2].[12] More will be said later about this error. For now, note that, if the same measurement system is used for the adjustment of rejected attributes as is used in conformance testing, then $\sigma_a = \sigma_t$.

As indicated in Table B-3, the probability that an attribute will be adjusted is obtained from

$$P_a(\text{adjusted}) = P_a(y \in R) = \int_{-R_1}^{R_2} f_a(y \in R) dy = \int_{-\infty}^{\infty} \int_{-R_1}^{R_2} f_a(y \mid x) f_m(x \mid \text{rejected}) dy\, dx \,. \tag{B-9}$$

Substituting from Eq. (B-4) for $f_t(x|\text{rejected})$ yields

$$P_a(\text{adjusted}) = \int_{-\infty}^{\infty} \int_{-R_1}^{R_2} f_a(y \mid x) \frac{f_t(\zeta \notin A \mid x) f_{bt}(x)}{P_t(\text{rejected})} dy\, dx = \frac{\int_{-\infty}^{\infty} \int_{-R_1}^{R_2} f_a(y \mid x)\left[1 - f_t(\zeta \in A \mid x)\right] f_{bt}(x) dy\, dx}{1 - P_t(\text{accepted})} \tag{B-10}$$

$$= \frac{1}{1 - P_t(\text{accepted})} \left\{ \int_{-\infty}^{\infty} \int_{-R_1}^{R_2} f_a(y \mid x) f_{bt}(x) dy\, dx - \int_{-\infty}^{\infty} \int_{-R_1}^{R_2} f_a(y \mid x) f_t(\zeta \in A \mid x) f_{bt}(x) dy\, dx \right\}.$$

For normally distributed variables, this expression becomes

$$P_a(\text{adjusted}) = \frac{1}{1 - P_t(\text{accepted})} \left[\varphi(0, R, \sigma_{adj}) - \frac{1}{\sqrt{2\pi}\sigma_{bt}} \int_{-\infty}^{\infty} \varphi(x, A, \sigma_t) \varphi(x, R, \sigma_a) e^{-x^2/2\sigma_{bt}^2}\, dx \right], \tag{B-11}$$

where

[12] This error is, of course, taken to be zero for software corrections or documented correction factors.

$$\sigma_{adj} = \sqrt{\sigma_a^2 + \sigma_{bt}^2}\,,$$[13]

and where the function φ is defined in Eq. (A-5).

Unadjusted Attributes

Attribute biases observed by the adjustment activity to be outside the region R are submitted for repair. Accordingly, the pdf for attributes that are *not* adjusted and are submitted for repair is

$$f_a(x \mid \text{not adjusted}) = f_a(x \mid y \notin R), \tag{B-12}$$

where $f_a(x \mid y \notin R)$ is given by

$$f_a(x \mid y \notin R) = \frac{f_a(y \notin R \mid x) f_t(x \mid \text{rejected})}{P_a(y \notin R)}.$$

The first term in the numerator is expressed as

$$f_a(y \notin R \mid x) = \int_{-\infty}^{\infty} f_a(y \mid x)dy - \int_{-R_1}^{R_2} f_a(y \mid x)dy = 1 - \int_{-R_1}^{R_2} f_a(y \mid x)dy,$$

and first term in the denominator is given by

$$\begin{aligned} P_a(y \notin R) &= \int_{-\infty}^{\infty} f_a(y \notin R \mid x) f_t(x \mid \text{rejected})dx = \int_{-\infty}^{\infty} \left[1 - \int_{-R_1}^{R_2} f_a(y \mid x)dy\right] f_t(x \mid \text{rejected})dx \\ &= 1 - \frac{1}{1 - P_t(\text{accepted})} \int_{-\infty}^{\infty} \int_{-R_1}^{R_2} f_a(y \mid x) f_t(y \notin A \mid x) f_{bt}(x) dy\, dx. \end{aligned} \tag{B-13}$$

The pdf for unadjusted attributes can then be written

$$f_a(x \mid y \notin R) = \frac{f_t(y \notin A \mid x) f_{bt}(x)}{P_a(y \notin R)\left[1 - P_t(\text{accepted})\right]} \left[1 - \int_{-R_1}^{R_2} f_a(y \mid x)dy\right], \tag{B-14}$$

where use was made of Eq. (B-4).

Developing these equations for normally distributed variables is accomplished in the same way as was done for adjusted attributes above.

Repaired Attributes

As indicated in Table B-3, attributes are repaired if their observed biases are considered too large to be corrected by adjustment alone and they are not discarded as unrepairable. Denoting the probability for the discarding an attribute as P_d and recalling that the decision to submit an attribute for repair is made by the adjusting activity, the probability that an attribute will be repaired is expressed as

$$P_r(\text{repaired}) = \left[1 - P_a(\text{adjusted})\right](1 - P_d), \tag{B-15}$$

where P_a(adjusted) is given in Eq. (B-9).[14]

[13] The integral in Eq. (B-11) is computed numerically.

Noting that, in the repair segment of Figure B-1, as with adjustments made by the adjustment activity, attribute biases are adjusted to zero. Hence, the appropriate pdf for repaired attributes is expressed as

$$f_r(x \,|\, \text{repaired}) = f_r(x) \,. \tag{B-16}$$

For a normally distributed combined measurement error, this pdf becomes

$$f_r(x) = \frac{1}{\sqrt{2\pi}\sigma_{rp}} e^{-x^2/2\sigma_{rp}^2} \,.$$

The quantity σ_{rp} represents the standard deviation of the probability distribution of biases following repair, given by

$$\sigma_{rp} = \sqrt{\sigma_r^2 + \sigma_{rb}^2} \,, \tag{B-17}$$

where σ_r is the standard uncertainty of the total measurement error in the repair activity's measurement system and σ_{rb} again represents rebound error. Note that, if the same measurement system is used for the adjustment of repaired attributes as is used in conformance testing, then $\sigma_a = \sigma_t$.

The Renew As-Needed Post-Test Distribution

The pdf $f_{pt}(x)$ shown in Figure B-1 is obtained from

$$f_{pt}(x) = f_t(x \,|\, \text{accepted}) P_t(\text{accepted}) + f_s(x \,|\, \text{renewed}) P_s(\text{renewed}) \,, \tag{B-18}$$

where

$$f_t(x \,|\, \text{accepted}) = \frac{f_t(\text{accepted} \,|\, x) f_{pt}(x)}{P_t(\text{accepted})} \,,$$

and

$$f_s(x \,|\, \text{renewed}) = \frac{f_a(x \,|\, \text{adjusted}) P_a(\text{adjusted}) + f_r(x \,|\, \text{repaired}) P_r(\text{repaired})}{P_s(\text{renewed})} \,. \tag{B-19}$$

Substituting Eq. (B-19) in Eq. (B-18), yields

$$\begin{aligned} f_{pt}(x) = f_t(\text{accepted} \,|\, x) f_{bt}(x) + f_a(x \,|\, \text{adjusted}) P_a(\text{adjusted}) \\ + f_r(x \,|\, \text{repaired}) P_r(\text{repaired}) \,. \end{aligned} \tag{B-20}$$

Finally, using Eqs. (B-1), (B-6), (B-15) and (B-16) in Eq. (B-20) gives

$$f_{pt}(x) = f_t(y \in A \,|\, x) f_{bt}(x) + f_a(x) P_a(\text{adjusted}) + f_r(x) \left[1 - P_a(\text{adjusted})\right](1 - P_d) \,. \tag{B-21}$$

Renew Always

We see from Figure B-3 that, with the renew always policy, the $f(x|\text{accepted})$ term is omitted from Eq. (B-18). Consequently, by Eq. (B-21), we have

---

[14] An attribute is discarded if it is a parameter of a discarded item. Presumably, such an item would be considered unrepairable for reasons of functional failure or degradation of performance unrelated to the probability distributions discussed in this paper.

$$f_{pt}(x) = f_a(x \mid \text{adjusted}) P_a(\text{adjusted}) + f_r(x \mid \text{repaired}) P_r(\text{repaired}) \,, \tag{B-22}$$

and substituting from Eqs. (B-6), (B-15) and (B-16), yields

$$f_{pt}(x) = f_a(x) P_a(\text{adjusted}) + f_r(x) \left[1 - P_a(\text{adjusted})\right](1 - P_d) \,. \tag{B-23}$$

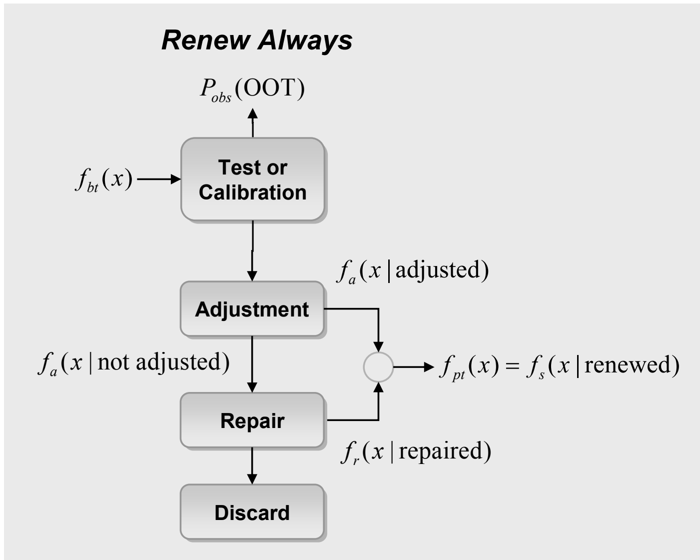

Figure B-2. The Renew Always Policy. With the renew always policy, $f_{pt}(x)$ is formed from $f_t(x|\text{adjusted})$ and $f_r(x|\text{repaired})$. Items with attributes that require renewal are discarded if an attribute cannot be renewed by adjustment or repair.

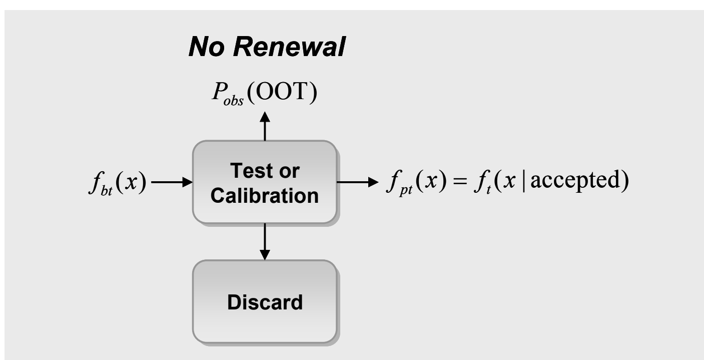

Figure B-3. The No Renewal Policy. With the no renewal policy, $f_{pt}(x)$ is equal to $f_t(x|\text{accepted})$. Items with attributes that require adjustment or repair are discarded.

No Renewal

With the no renewal policy,

$$f_{pt}(x) = f_t(x \mid \text{accepted}) \,, \tag{B-24}$$

where $f_t(x|\text{accepted})$ is given in Eqs. (2-10) and (2-11)

$$f_t(x \mid \text{accepted}) = \frac{f_t(\text{accepted} \mid x) f_{bt}(x)}{P_t(\text{accepted})} = \frac{f_t(y \in A \mid x) f_{bt}(x)}{P_t(y \in A)} \tag{B-25}$$

where $f_t(y \in A | x)$ is obtained by integrating Eq. (21) over A.

Comparison of Post-Test Distributions

Figure B-4 depicts the pdf $f_{pt}(x)$ for each of the renewal policies described above. The plots apply to cases where $f_{bt}(x)$ and $f_t(y|x)$ represent normally distributes variables. The pdfs shown for the renew as-needed and renew always policies are developed under the assumption that P_r(repaired) = 0 and that $\sigma_a = \sigma_t$. Given these conditions, we have

$$f_{pt}(x) = \begin{cases} f_t(y \in A | x) f_{bt}(x) + f_a(x)[1 - P_t(y \in A)] & \text{(renew as-needed)} \\ f_{tp}(x) & \text{(renew always)} \\ \dfrac{f_t(y \in A | x)}{P_t(y \in A)} f_{bt}(x). & \text{(no renewal)} \end{cases}$$

For normally distributed variables, these pdfs become

$$f_{pt}(x) = \begin{cases} \dfrac{\varphi(x, A, \sigma_t)}{\sqrt{2\pi}\sigma_{bt}} e^{-x^2/2\sigma_{bt}^2} + \dfrac{1 - \varphi(0, A, \sigma_t)}{\sqrt{2\pi}\sigma_{tp}} e^{-x^2/2\sigma^2}, & \text{(renew as-needed)} \\ \dfrac{1}{\sqrt{2\pi}\sigma_{tp}} e^{-x^2/2\sigma_{tp}^2}, & \text{(renew always)} \\ \dfrac{\varphi(x, A, \sigma_t)}{\varphi(0, A, \sigma_{obs})} \dfrac{e^{-x^2/2\sigma_{bt}^2}}{\sqrt{2\pi}\sigma_{bt}}, & \text{(no renewal)}, \end{cases}$$

where σ is defined in Eq. (A-6) and σ_{tp} is given in Eq. (B-8) with σ_a replaced by σ_t.

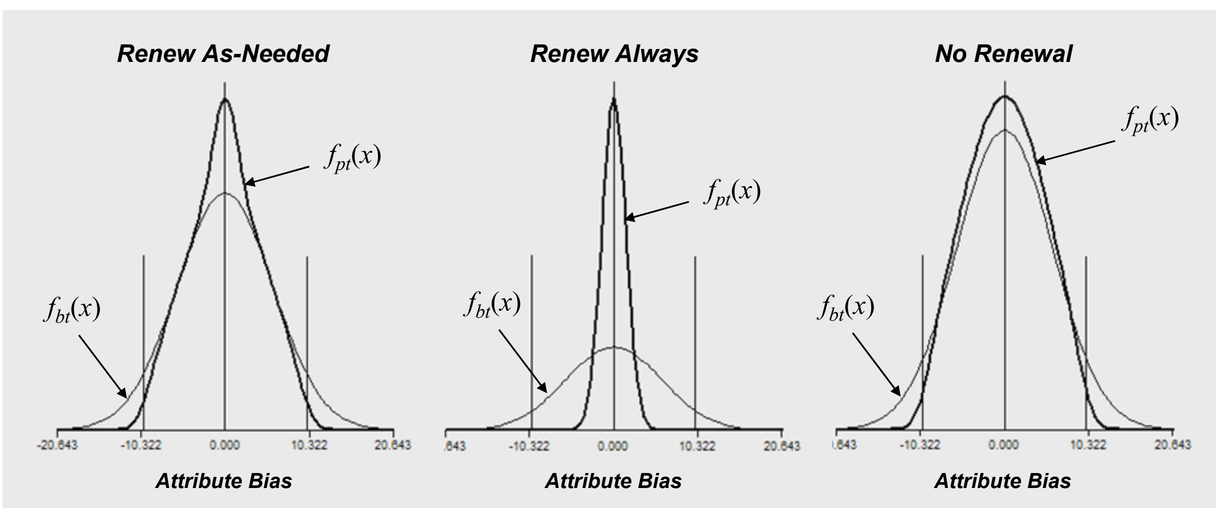

Figure B-4. Comparison of Renewal Policies. Before-test and post-test pdfs are shown for a case in which $L = 10$, P_{obs}(OOT) = 0.10 and the measurement process uncertainty is 25% of the before-test UUT attribute bias uncertainty. Rebound error is assumed to be zero.

Renewal Effects – Rebound Error Revisited

In the pdfs of Figure B-4, rebound error and uncertainty are zero. Figures B-5 through B-7 demonstrate the impact on the post-test distribution for the renew as-needed and renew always policies (there is no impact with the no renewal policy).

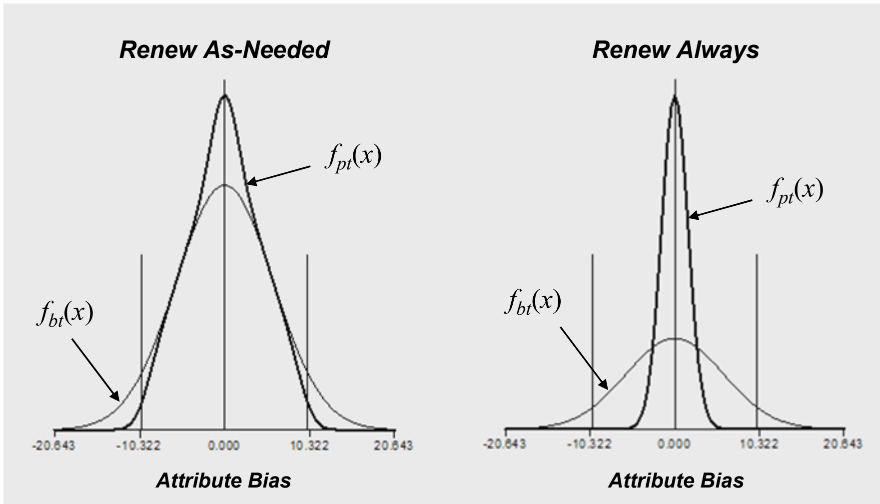

Figure B-5. Renewal Impact – Moderate Uncertainty Due to Rebound Error. Shown are cases where the uncertainty in rebound error is about half the standard uncertainty of the test process. Apparently, for small rebound error uncertainties, renewal has little negative effect on post-test distributions.

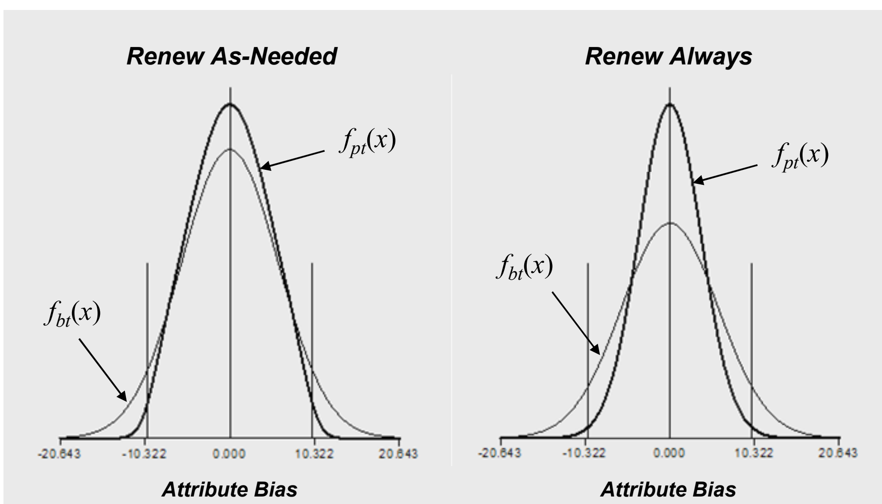

Figure B-6. Renewal Impact – Large Uncertainty Due to Rebound Error. Shown are cases where the uncertainty in rebound error is about twice the standard uncertainty of the test process. Comparison with Figure B-5 shows that rebound error has a more pronounced effect under the renew always policy than under the renew as-needed policy.

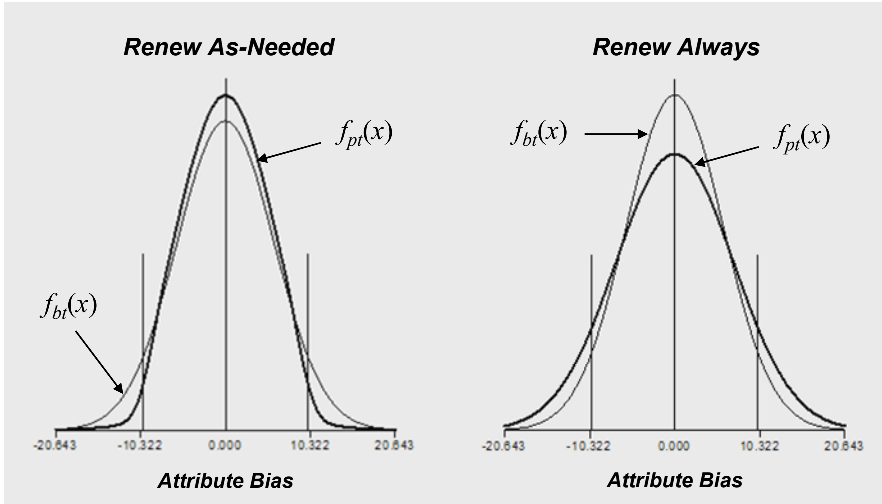

Figure B-7. Renewal Impact – Extreme Uncertainty Due to Rebound Error. Shown are cases where the uncertainty in rebound error is about four times the standard uncertainty of the test process. Notice that, for the renew always policy, if the uncertainty due to rebound error becomes excessive, the post-test UUT attribute bias distribution reflects a higher post-test OOT probability than is reflected in the before-test distribution. The upshot is that, in such cases, renewal has a negative rather than a beneficial effect.