Decision Risk Analysis for Alternative Calibration Scenarios[1]

Howard Castrup, Ph.D.
President
Integrated Sciences Group
Bakersfield, CA 93306
hcastrup@isgmax.com

Abstract

The results of calibration of unit-under-test (UUT) attributes and estimates of measurement process uncertainty are employed in the calculation of measurement decision risk within the context of four measurement scenarios. The decision risks of interest are those that are relevant to meeting Z540.3 requirements as well as internal quality control criteria. They include unconditional false accept risk (*UFAR*), conditional false accept risk (*CFAR*) and false reject risk (*FRR*). *UFAR* is computed both as a program-level and bench-level control metric. Examples are given to illustrate concepts and procedures.

Risk Analysis Alternatives

It should be said that alternative methods of risk analysis have been made available [4, 5]. For each, the measurement scenarios and equations presented in the present paper are applicable.

Calibration Scenarios

This paper discusses information obtained from measurements made during calibration and the application of this information to measurement decision risk analysis in the context of four calibration scenarios: [9]

Table 1. Calibration Scenarios

Scenario	Description
1	The measurement reference (MTE) measures the value of a passive attribute of the unit under test (UUT).
2	The UUT measures the value of a passive reference attribute of the MTE.
3	The UUT and MTE each provide an “output” or “stimulus” for comparison using a comparator.
4	The UUT and MTE both measure the value of an attribute of a common artifact that provides an output or stimulus.

The information obtained includes an observed value, referred to as a “measurement result” or “calibration result,” and an estimated uncertainty in the measurement error. Each scenario is

[1] Presented at the NCSLI Workshop & Symposium, Orlando, August 2008.

characterized by a measurement equation that is applicable to the manner in which calibrations are performed and calibration results are recorded or interpreted.

The measurement scenarios turn out to be simple and intuitive. In each, the measurement result and the measurement error are separable, allowing the estimation of measurement uncertainty. It is assumed in each scenario that the measurement result is an estimate of the value of the bias of the UUT attribute.

Calibration Results

In each of the scenarios described above, a measurement result δ, a measurement uncertainty u_{cal} and a UUT bias $e_{UUT,b}$ comprise the variables of interest to measurement decision risk analysis. For each calibration scenario, the composition of ε_{cal} and the expression for δ are given in Ref [9].

Measurement Decision Risk Analysis

In calibrating a UUT to determine if it is in- or out-of-tolerance, we face two principal varieties of measurement decision risk. Name, False Accept Risk and False Reject Risk. The former can be expressed in two ways. First, there is the probability that a UUT attribute is both out-of-tolerance and observed to be in-tolerance. Alternatively, we could focus on the probability that a UUT attribute accepted as being in-tolerance will be out-of-tolerance. The first alternative is called "unconditional false accept risk" or *UFAR*. The second is called the "conditional false accept risk" or *CFAR*. False reject risk is sometimes denoted *FRR*.[2]

Denote the tolerance limits for UUT attribute bias $e_{UUT,b}$ as $-L_1$ and L_2. Then, given the notation employed in this paper, the UUT attribute is in-tolerance if $-L_1 \leq e_{UUT,b} \leq L_2$. Next, accounting for the possibility that test limits or "guardband" limits $-A_1$ and A_2 may be used to trigger a UUT attribute adjustment or other corrective actions, we say that the UUT attribute is observed to be in-tolerance if $-A_1 \leq \delta \leq A_2$.

Probability Expressions

We use standard probability notation in which $P(E)$ represents the probability that an event E will occur, $P(\bar{E})$ represents the probability that the event E will not occur, $P(E_1,E_2)$ represents the probability that events E_1 and E_2 will both occur and $P(E_2|E_1)$ represents the probability that event E_2 will occur given that event E_1 has occurred.

To offer a compact notation for the measurement decision risk probability functions, we define E_A as the event $-A_1 \leq \delta \leq A_2$ and E_L as the event $-L_1 \leq e_{UUT,b} \leq L_2$. Then, we can express *UFAR* as

$$\begin{aligned} PFA &= P(e_{UUT,b} < -L_1 \text{ or } e_{UUT,b} > L_2, -A_1 \leq \delta \leq A_2) \\ &= P(\bar{E}_L, E_A). \end{aligned}$$

[2] *UFAR* and *CFAR* are also referred to respectively as the "probability for a false accept" or *PFA* and "the "conditional probability for a false accept" or *CPFA*. False reject risk is sometimes denoted *PFR*.

Taking advantage of probability fundamentals, we can also write

$$PFA = P(E_A) - P(E_L, E_A) .$$

The first term on the right hand side of this expression is the probability that the UUT will be observed to be in-tolerance. The second term is the probability that the UUT will both be in-tolerance and observed to be in-tolerance.

Note that the conditional false accept risk can be written

$$\begin{aligned} CPFA = P(\bar{E}_L \mid E_A) &= \frac{PFA}{P(E_A)} \\ &= 1 - \frac{P(E_L, E_A)}{P(E_A)}, \end{aligned}$$

and that False Reject Risk can be written

$$PFR = P(E_L) - P(E_L, E_A) .$$

Risk Computations

The computation of measurement decision risk employs the probability density functions $f(\delta \mid e_{UUT,b})$ and $f(e_{UUT,b})$ [3, 4]. The standard deviation of the former is just u_{cal}, whereas the standard deviation of the latter is an estimate of the standard deviation of the population from which the UUT attribute value has been randomly selected prior to measurement. In this paper, δ turns out to be $N(e_{UUT,b}, u_{cal}^2)$.[3] The applicable distribution for the pdf $e_{UUT,b}$ may also be normal, but not necessarily so. For example, if the UUT tolerances are asymmetric, a skewed distribution, such as the lognormal or gamma, may be appropriate.

The relevant expressions are

$$\begin{aligned} P(E_L, E_A) &= \int_{-L_1}^{L_2} f(e_{UUT,b}) de_{UUT,b} \int_{-A_1}^{A_2} f(\delta \mid e_{UUT,b}) d\delta \\ &= \frac{1}{\sqrt{2\pi u}} \int_{-L_1}^{L_2} f(e_{UUT,b}) de_{UUT,b} \int_{-A_1}^{A_2} e^{-(\delta - e_{UUT,b})^2 / 2u_{cal}^2} d\delta , \end{aligned}$$

and

$$P(E_A) = \frac{1}{\sqrt{2\pi u}} \int_{-\infty}^{\infty} f(e_{UUT,b}) de_{UUT,b} \int_{-A_1}^{A_2} e^{-(\delta - e_{UUT,b})^2 / 2u_{cal}^2} d\delta .$$

The probability $P(E_L)$ is usually obtained from calibration history data. However, it can be expressed mathematically as

$$P(E_L) = \int_{-L_1}^{L_2} f(e_{UUT,b}) de_{UUT,b} .$$

---

[3] It is recognized that not all error sources may be normally distributed, such as the error due to resolution in digital readouts. However, the use of the normal distribution for the error ε is justified by the central limit theorem.

It should be said that alternative methods of risk analysis have been made available [4, 5]. For each, the measurement scenarios and equations presented in the present paper are applicable.

The relevant expressions are

$$P(E_L, E_A) = \int_{-L_1}^{L_2} f(e_{UUT,b}) de_{UUT,b} \int_{-A_1}^{A_2} f(\delta \mid e_{UUT,b}) d\delta$$
$$= \frac{1}{\sqrt{2\pi} u} \int_{-L_1}^{L_2} f(e_{UUT,b}) de_{UUT,b} \int_{-A_1}^{A_2} e^{-(\delta - e_{UUT,b})^2 / 2u_{cal}^2} d\delta \, ,$$

and

$$P(E_A) = \frac{1}{\sqrt{2\pi} u} \int_{-\infty}^{\infty} f(e_{UUT,b}) de_{UUT,b} \int_{-A_1}^{A_2} e^{-(\delta - e_{UUT,b})^2 / 2u_{cal}^2} d\delta \, .$$

The probability $P(E_L)$ is usually obtained from calibration history data. However, it can be expressed mathematically as

$$P(E_L) = \int_{-L_1}^{L_2} f(e_{UUT,b}) de_{UUT,b} \, .$$

If the applicable distribution for $e_{UUT,b}$ is normal, i.e., if $e_{UUT,b}$ is $N(0, u_{UUT,b}^2)$, where $u_{UUT,b}$ is the standard deviation of the UUT attribute value prior to calibration, then

$$f(e_{UUT,b}) = \frac{1}{\sqrt{2\pi} u_{UUT,b}} e^{-e_{UUT,b}^2 / 2u_{UUT,b}^2} \, .$$

Letting $x = e_{UUT,b}$ and $y = \delta$ for simplicity of notation, applying Eq. (65) in Eq. (62) yields

$$P(E_L, E_A) = \frac{1}{2\pi u_{UUT,b} u_{cal}} \int_{-L_1}^{L_2} e^{-x^2 / 2u_{UUT,b}^2} dx \int_{-A_1}^{A_2} e^{-(y-x)^2 / 2u_{cal}^2} dy$$
$$= \frac{1}{\sqrt{2\pi} u_{UUT,b}} \int_{-L_1}^{L_2} \left[\Phi\left(\frac{A_2 - x}{u_{cal}} \right) - \Phi\left(-\frac{A_1 + x}{u_{cal}} \right) \right] e^{-x^2 / 2u_{UUT,b}^2} dx$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-L_1/u_{UUT,b}}^{L_2/u_{UUT,b}} \left[\Phi\left(\frac{A_1 - u_{UUT,b}\zeta}{u_{cal}} \right) + \Phi\left(\frac{A_2 + u_{UUT,b}\zeta}{u_{cal}} \right) - 1 \right] e^{-\zeta^2/2} d\zeta \, ,$$

where the Φ is the normal distribution function. It should be noted that the integral must be solved for graphically or by numerical iteration.

The probability P(E_A) can be solved for in closed form by applying Eq. (65) in Eq. (63) and completing the integration

$$P(E_A) = \frac{1}{2\pi u_{UUT,b} u_{cal}} \int_{-\infty}^{\infty} e^{-x^2/2u_{UUT,b}^2} dx \int_{-A_1}^{A_2} e^{-(y-x)^2/2u_{cal}^2} dy$$
$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{-A_1}^{A_2} e^{-y^2/2\sigma^2} dy$$
$$= \Phi\left(\frac{A_2}{\sigma}\right) + \Phi\left(\frac{A_1}{\sigma}\right) - 1,$$

where

$$\sigma = \sqrt{u_{UUT,v}^2 + u_{cal}^2} \; .$$

Likewise, $P(E_L)$ can be expressed as

$$P(E_L) = \Phi\left(\frac{L_2}{u_{UUT,b}}\right) + \Phi\left(\frac{L_1}{u_{UUT,b}}\right) - 1 \, .$$

Measurement Decision Risk Analysis Summary

FAR and the *CFAR* are given by

$$UFAR = P(E_A) - P(E_L, E_A)$$

and

$$CFAR = 1 - \frac{P(E_L, E_A)}{P(E_A)} \, ,$$

respectively, where E_L is the event in which the UUT is in-tolerance and E_A is the event that it is accepted without adjustment or other correction. *FRR* is

$$FRR = P(E_L) - P(E_L, E_A) \, .$$

The pdf $f(\delta \,|\, e_{UUT,b})$ for each of the four calibration scenarios is taken to be

$$f(\delta \,|\, e_{UUT,b}) = \frac{1}{\sqrt{2\pi} u} e^{-(\delta - e_{UUT,b})^2/2u_{cal}^2} \, .$$

The applicable distribution for $e_{UUT,b}$ may be normal, but not necessarily so. If $e_{UUT,b}$ is $N(0, u_{UUT,b}^2)$, where $u_{UUT,b}$ is the standard deviation of the UUT attribute value prior to calibration, then

$$f(e_{UUT,b}) = \frac{1}{\sqrt{2\pi} u_{UUT,b}} e^{-e_{UUT,b}^2/2u_{UUT,b}^2} \, ,$$

$$P(E_L, E_A) = \frac{1}{\sqrt{2\pi}} \int_{-L_1/u_{UUT,b}}^{L_2/u_{UUT,b}} \left[\Phi\left(\frac{A_1 - u_{UUT,b}\zeta}{u_{cal}}\right) + \Phi\left(\frac{A_2 + u_{UUT,b}\zeta}{u_{cal}}\right) - 1 \right] e^{-\zeta^2/2} d\zeta \, ,$$

$$P(E_A) = \Phi\left(\frac{A_2}{\sigma}\right) + \Phi\left(\frac{A_1}{\sigma}\right) - 1,$$

where

$$\sigma = \sqrt{u_{UUT,v}^2 + u_{cal}^2},$$

and

$$P(E_L) = \Phi\left(\frac{L_2}{u_{UUT,b}}\right) + \Phi\left(\frac{L_1}{u_{UUT,b}}\right) - 1.$$

Nomenclature

The nomenclature used in this paper is summarized in Table 2.

Table 2. Nomenclature

Quantity	**Description**
UUT	Unit Under Test.
Attribute	A measurable property of a device, substance or other quantity.
MTE	Measuring or Test Equipment. The measurement reference.
ε_m	The total error in the measurement of the value of an attribute.
δ	The result of a calibration.
ε_{cal}	The error in δ.
u_{cal}	The uncertainty in ε_{cal}.
$e_{UUT,b}$	The bias of a calibrated UUT attribute. The quantity estimated by δ.
$e_{MTE,b}$	The bias of the MTE attribute used to calibrate the UUT attribute.
$\varepsilon_{UUT,m}$	The error in measurements made with the UUT attribute or the error in measuring the UUT attribute's value with a comparator.
$\varepsilon_{MTE,m}$	The error in measurements made with the MTE attribute or the error in measuring the MTE attribute's value with a comparator.
x_n	The nominal value of the UUT attribute.
x_{true}	The true value of the UUT attribute.
y_n	The nominal value of the MTE attribute.
y_{true}	The true value of the MTE attribute.
x_c	The value of the UUT attribute indicated by a measurement taken with a comparator.
$e_{c,b}$	The bias in a comparator indication.
A_1 and A_2	Lower and upper limits bounding acceptable values of δ.
L_1 and L_2	Tolerance limits for the UUT attribute, i.e., lower and upper limits bounding acceptable values of $e_{UUT,b}$.

Quantity	Description
UFAR	Unconditional False Accept Risk. The probability that an out-of-tolerance UUT attribute will be observed to be in-tolerance.
FRR	False Reject Risk. The probability that an in-tolerance UUT attribute will be observed to be out-of-tolerance.
$f(e_{UUT,b})$	The probability density function for the bias of the UUT attribute as received for calibration
$f(\delta \vert e_{UUT,b})$	The probability density function for obtaining a calibration result δ, given a UUT attribute bias $e_{UUT,b}$.
E_L	The event that the UUT attribute is in-tolerance.
E_A	The event that the UUT attribute is observed to be in-tolerance.
$P(E_L)$	The probability that the event E_L will occur.
$P(E_A)$	The probability that the event E_A will occur.
$P(E_L, E_A)$	The joint probability that the events E_L and E_A will both occur.

References

[1] NCSLI, *Determining and Reporting Measurement Uncertainties*, Recommended Practice RP-12, NCSL International, Under Revision.

[2] NASA, *Measurement Uncertainty Analysis Principles and Methods*, KSC-UG-2809, National Aeronautics and Space Administration, November 2007.

[3] ISG, *Analytical Metrology Handbook*, Integrated Sciences Group, www.isgmax.com/analytical_metrology.htm, 2006.

[4] Castrup, H., "Risk Analysis Methods for Complying with Z540.3," *Proc. NCSLI Workshop & Symposium*, St. Paul, 2007.

[5] ISG, RiskGuard 2.0, www.isgmax.com/risk_freeware.htm, © 2007, Integrated Sciences Group.

[6] ISG, UncertaintyAnalyzer 3.0, www.isgmax.com/unc_features.htm, © 2005-2006 Integrated Sciences Group.

[7] ISG, Uncertainty Sidekick Pro, www.isgmax.com/sidekickpro_features.htm, © 2006-2007, Integrated Sciences Group.

[8] ISG, Uncertainty Sidekick, www.isgmax.com/uncertainty_freeware.htm, © 2005-2007, Integrated Sciences Group.

[9] Castrup, H. and Castrup, S., "Uncertainty Analysis for Alternative Calibration Scenarios," *Proc. NCSLI Workshop & Symposium*, Orlando, 2008.