GUIDELINES
for
EVALUATING and EXPRESSING
the
UNCERTAINTY
of
NIST MEASUREMENT RESULTS

Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results

based on

NIST Technical Note 1297 (1994 Edition)

by

Barry N. Taylor and Chris E. Kuyatt
Physics Laboratory, NIST

(Supersedes NIST Technical Note 1297, January 1993)

and

The Guide to the Expression of Uncertainty in Measurements (**GUM**) 1995 version

OUTLINE

1. History: The GUM and TN1297
2. Terminology: Measurand, Influence Quantities,
   Measurement Equation, Error, Uncertainty,
   Sources of Uncertainty, Types of Uncertainty,
   Systematic and Random Errors,
   Standard Uncertainty, Expanded Uncertainty,
   Confidence Intervals, Coverage Factors,
   Degrees of Freedom, etc.
3. Probability Distributions, Variance, Standard Deviation
4. Calculation of Uncertainty Components
5. Taylor Series and Propagation of Uncertainties
6. NIST Policy on reporting uncertainties

SECTION – 1

The GUM & TN1297

BRIEF HISTORY

BRIEF HISTORY

- In 1977, the International Committee for Weights and Measures (CIPM) requested the International Bureau for Weights and Measures (BIPM) that it should investigate the question of reaching an international consensus on expressing uncertainty in measurement.
- The request was initiated by then CIPM member and NBS Director **E. Ambler.**
- A Working Group on the Statement of Uncertainties was convened in 1980 by the BIPM.
- This group made several recommendations concerning the statement of uncertainties in measurements.
- The CIPM approach is founded on Recommendation INC-1 (1980) of this Working Group.

GUIDE
to the expression of
UNCERTAINTY *in* **MEASUREMENT**

- In 1985 CIPM requested the ISO to develop a broadly applicable guidance document based on Recommendation INC-1 (1980).

- This led to the development of the ***Guide*** *(the* ***GUM*** ***or*** ***Guide to the expression of uncertainty in measurement****)*.

- It is at present the most complete reference on the general application of the CIPM approach to expressing measurement uncertainty.

- Its development is giving further impetus to the worldwide adoption of that approach.

NIST Ad Hoc Committee on Uncertainty

- In July 1992 Director Lyons appointed a NIST Ad Hoc Committee on Uncertainty Statements and charged it with recommending a NIST policy on this important topic. This action was motivated in part by
  - ✓ the emerging international consensus on the approach to expressing uncertainty in measurement recommended by the CIPM,

and in part due to
  - ✓ the realization that there has never been a uniform approach at NIST to the expression of uncertainty.

- The movement toward the international adoption of the CIPM approach for expressing uncertainty is driven to a large extent by the global economy and marketplace; allowing measurements performed in different countries and sectors to be more easily understood, interpreted, and compared.

AD HOC COMMITTEE RECOMMENDATIONS

- The Ad Hoc Committee carefully reviewed the needs of NIST customers regarding statements of uncertainty and the compatibility of those needs with the CIPM approach and concluded that the CIPM approach could be used to provide quantitative expressions of measurement uncertainty that would satisfy the requirements of NIST customers.

- The Ad Hoc Committee then recommended a specific policy for the implementation of that approach at NIST.

- In October 1992, this policy on expressing measurement uncertainty was instituted at NIST. It is set forth in "Statements of Uncertainty Associated With Measurement Results," Appendix E, NIST Technical Communications Program, Subchapter 4.09 of the Administrative Manual (*reproduced as Appendix C of* ***Technical Note 1297***)

TECHNICAL NOTE 1297

Technical Note 1297 *was prepared by Dr. Taylor and Dr. Kuyatt, both of whom served on the NIST Ad Hoc committee.*

- It provides a helpful discussion of the CIPM approach and, with its aid, it is expected that the NIST policy can be implemented without excessive difficulty.

- Because NIST statements of uncertainty resulting from the policy will be uniform among themselves and consistent with current international practice, the policy will help our customers increase their competitiveness in the national and international marketplaces.

APPLICATIONS OF TN 1297

The guidance given in TN1297 is intended to be applicable to most, if not all, NIST measurement results, including results associated with

- international comparisons of measurement standards,
- basic research,
- applied research and engineering,
- calibrating client measurement standards,
- certifying standard reference materials, and
- generating standard reference data.

- Since the *Guide* itself is intended to be applicable to similar kinds of measurement results, it may be consulted for additional details.

SECTION – 2

TERMINOLOGY

THE MEASURAND

In general, the result of a measurement is only an approximation or estimate of the value of the specific quantity subject to measurement, that is, the **measurand**, and thus the result is complete only when accompanied by a quantitative statement of its uncertainty.

MEASURAND

Value of the specific quantity subject to measurement

Influence Quantity

Any quantity that is not the measurand but that affects the result of the measurement

Examples

a) temperature of a micrometer used to measure length;
b) frequency in the measurement of the amplitude of an alternating electric potential difference;
c) bilirubin concentration in the measurement of haemoglobin concentration in a sample of human blood plasma.

MEASURAND -- Examples

- Vapor pressure of a given sample of water at 20 deg C
- The length of a steel bar at 25 deg C and 101.325 Pa

The measurand should be defined with sufficient completeness with respect to the required accuracy.

Incomplete definition of a measurand can give rise to a component of uncertainty large enough that it must be included in the evaluation of the uncertainty of the measurement result.

Measurement Process and Measurement Result

Measurement Process

A set of operations having the object of determining a value of a quantity.

Measurement Result

Value attributed to a measurand, obtained by measurement.

A complete statement of the result of a measurement includes information about the uncertainty of measurement

UNCERTAINTY of MEASUREMENT

Parameter associated with the result of a measurement, that characterizes the **dispersion** of the values that could reasonably be attributed to the measurand

The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.

It is understood that the **result of the measurement** (after corrections) is the **best estimate** of the value of the measurand.

Error and Relative Error of Measurement

Error of Measurement

Result of a measurement minus the value of the measurand

Relative Error of Measurement

Error of measurement divided by the (absolute) value of the measurand

Random and Systematic Error

Random Error

Result of measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under **repeatability conditions**

Systematic Error

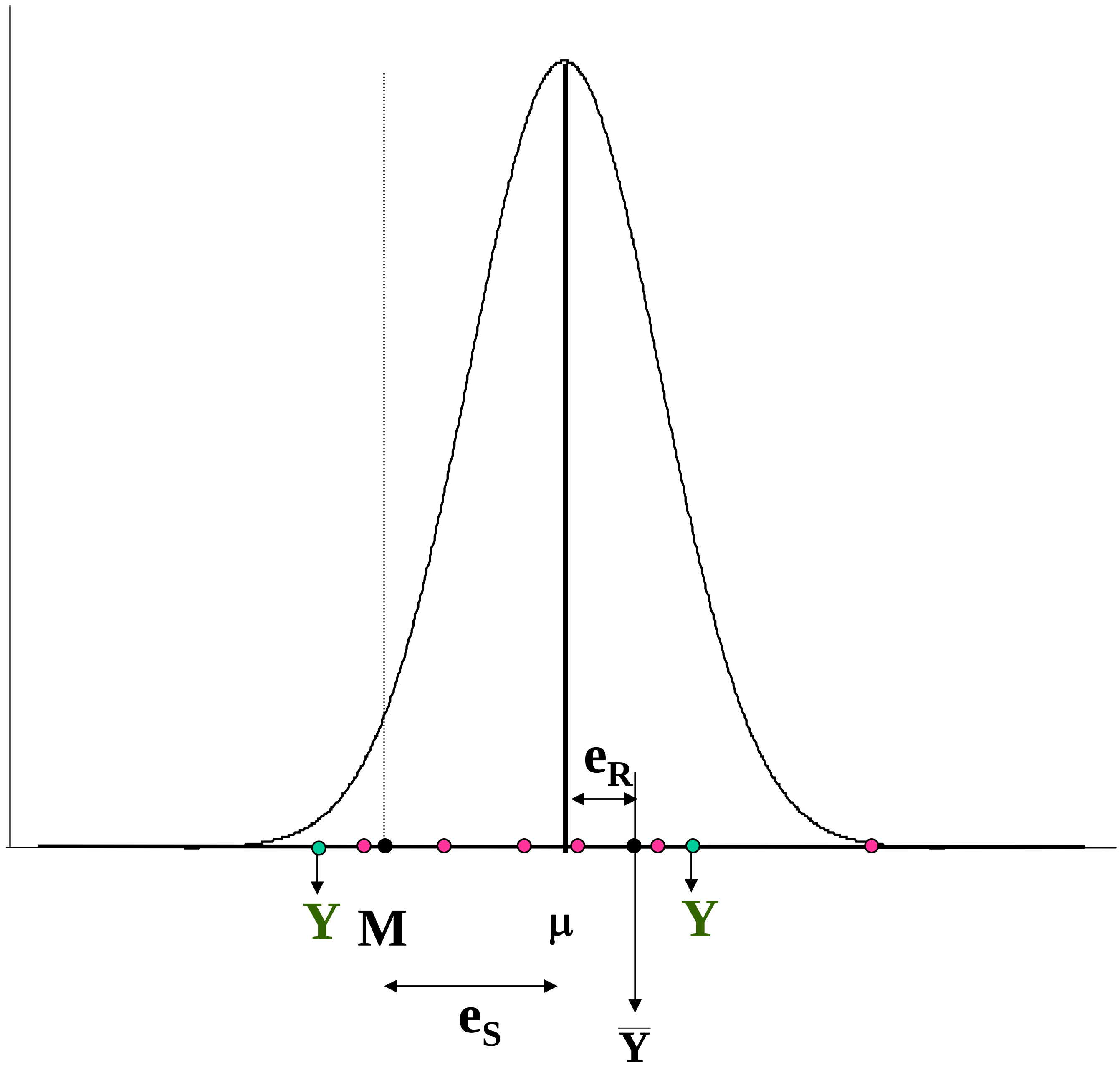
Mean that would result from an infinite number of measurements of the same measurand carried out under **repeatability conditions** minus the value of the measurand

Result of Measurement

= **Value of Measurand + Systematic Error** + **Random Error**

$$Y = M + (Y - m) + (m - M)$$

Measurand, Result, Random Error, Systematic Error



Accuracy & Precision

Accuracy:

Closeness of the agreement between the result of a measurement and the value of the measurand.

Precision:

Closeness of agreement between independent test results obtained under stipulated conditions.

- The term **precision** should not be used for "accuracy."

Accuracy & Precision

The terms accuracy and precision are terms that represent **qualitative concepts** and thus should be used with care.

- Because “accuracy” is a qualitative concept, one should not use it quantitatively, that is, associate numbers with it; numbers should be associated with measures of uncertainty instead. Thus one may write “the standard uncertainty is 2 μΩ” but not “the accuracy is 2 μΩ

- Similarly, the statement “the precision of the measurement results, expressed as the standard deviation obtained under repeatability conditions, is 2 μΩ” is acceptable, but the statement “the precision of the measurement results is 2 μΩ” is not.

Correction & Correction Factor

Correction

value added algebraically to the uncorrected result of a measurement to compensate for systematic error

The correction is equal to the negative of the estimated systematic error

Generally, a correction is applied to the raw result based on a best estimate of total systematic error. The correction is not perfect and there will be a residual systematic error present. Even though the best estimate of this residual systematic error is ZERO, there is uncertainty associated with it.

Correction & Correction Factor

Correction factor

numerical factor by which the uncorrected result of a measurement is multiplied to compensate for systematic error

Since the systematic error cannot be known perfectly, the compensation cannot be complete.

Repeatability & Reproducibility

Repeatability

closeness of the agreement between the results of successive measurements of the same measurand carried out under the **same conditions of measurement**

1. These conditions are called Repeatability Conditions
2. Repeatability conditions include:
   - the same measurement procedure
   - the same observer
   - the same measuring instrument, used under the same conditions
   - the same location
   - repetition over a short period of time
3. Repeatability may be expressed quantitatively in terms of the dispersion characteristics of the results

Repeatability & Reproducibility

Reproducibility

closeness of the agreement between the results of measurements of the same measurand carried out under **changed conditions of measurement**

1. **A valid statement of reproducibility requires specification of the conditions changes.**
2. The changed conditions **may include**:
   - principle of measurement
   - method of measurement
   - observer
   - measuring instrument
   - reference standard
   - location
   - conditions of use
   - time
3. Reproducibility may be expressed quantitatively in terms of the **dispersion characteristics** of the (*corrected*) results.

SECTION – 3

PROBABILITY DISTRIBUTIONS

Probability Distributions for Modeling Error

Uncertainty in Measurement arises due to the presence of error in the result whose magnitude is unknown (and not knowable exactly)

** Uncertainty and Error are not synonyms

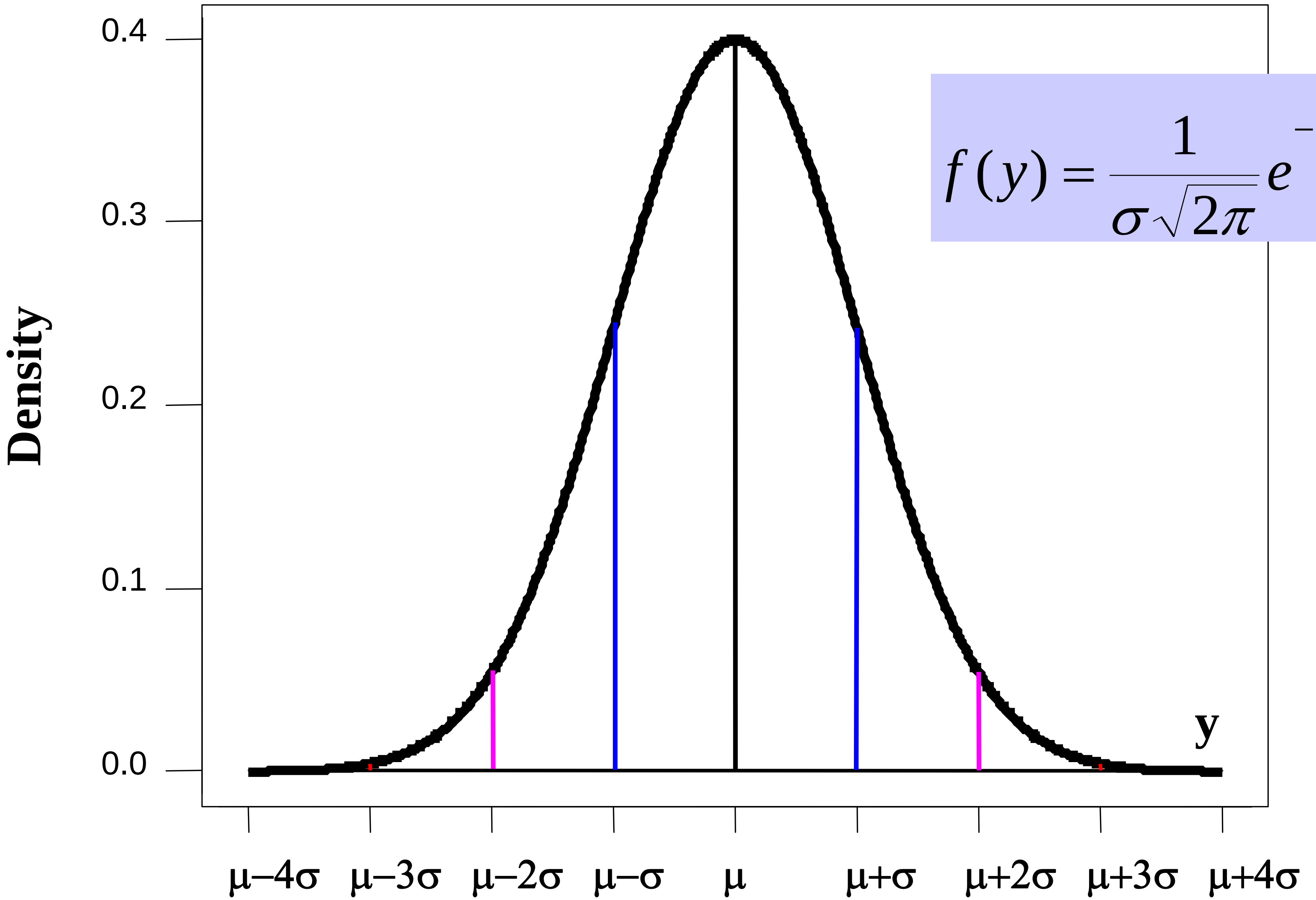
Expression of this uncertainty requires some knowledge concerning the **collection of plausible values of the error** along with their corresponding **probabilities**.

Usually, errors that are large in magnitude are less likely to occur than errors of small magnitude.

Such information can be modeled using **probability distributions**.

Statistical Description of Error Distributions

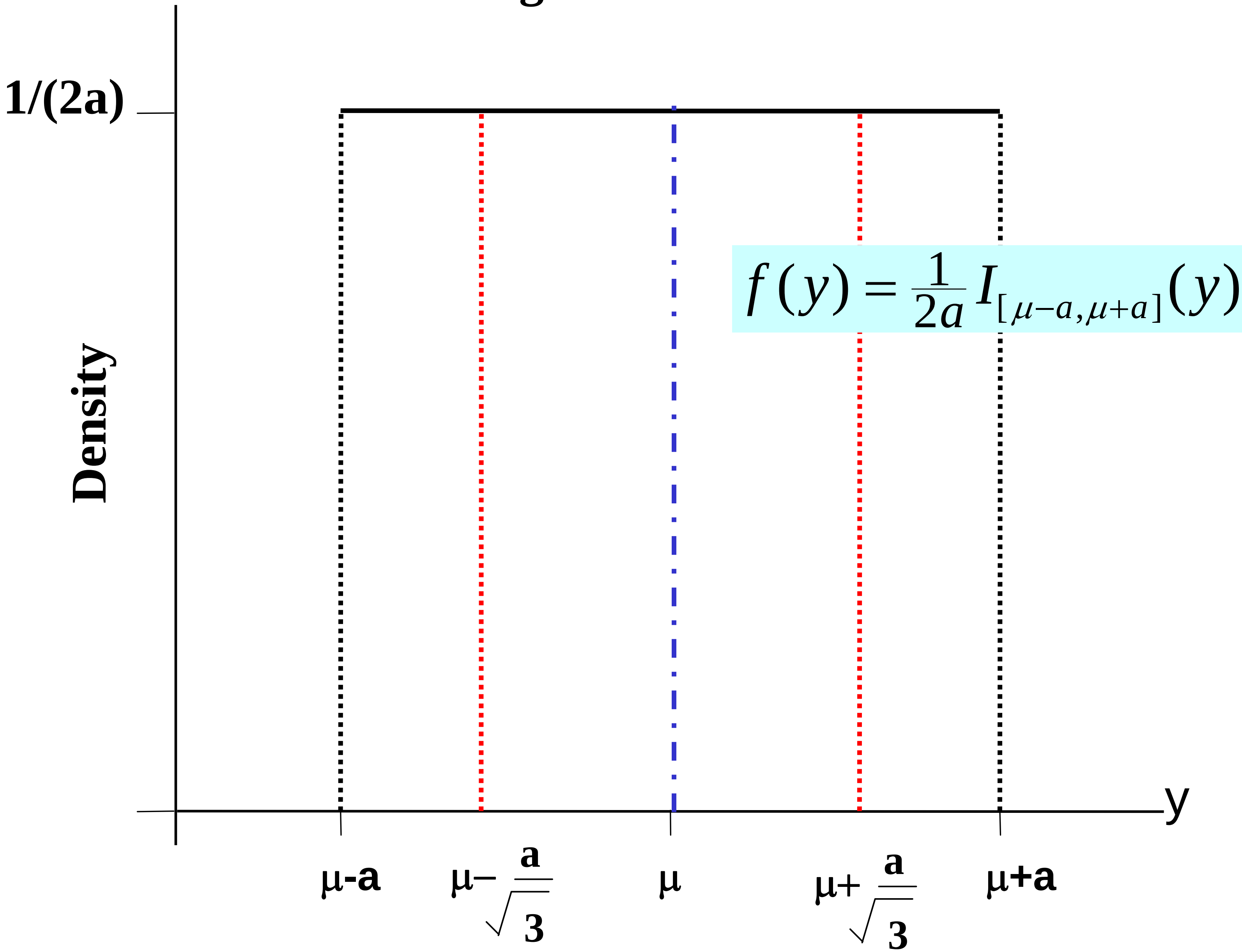
Normal Distribution



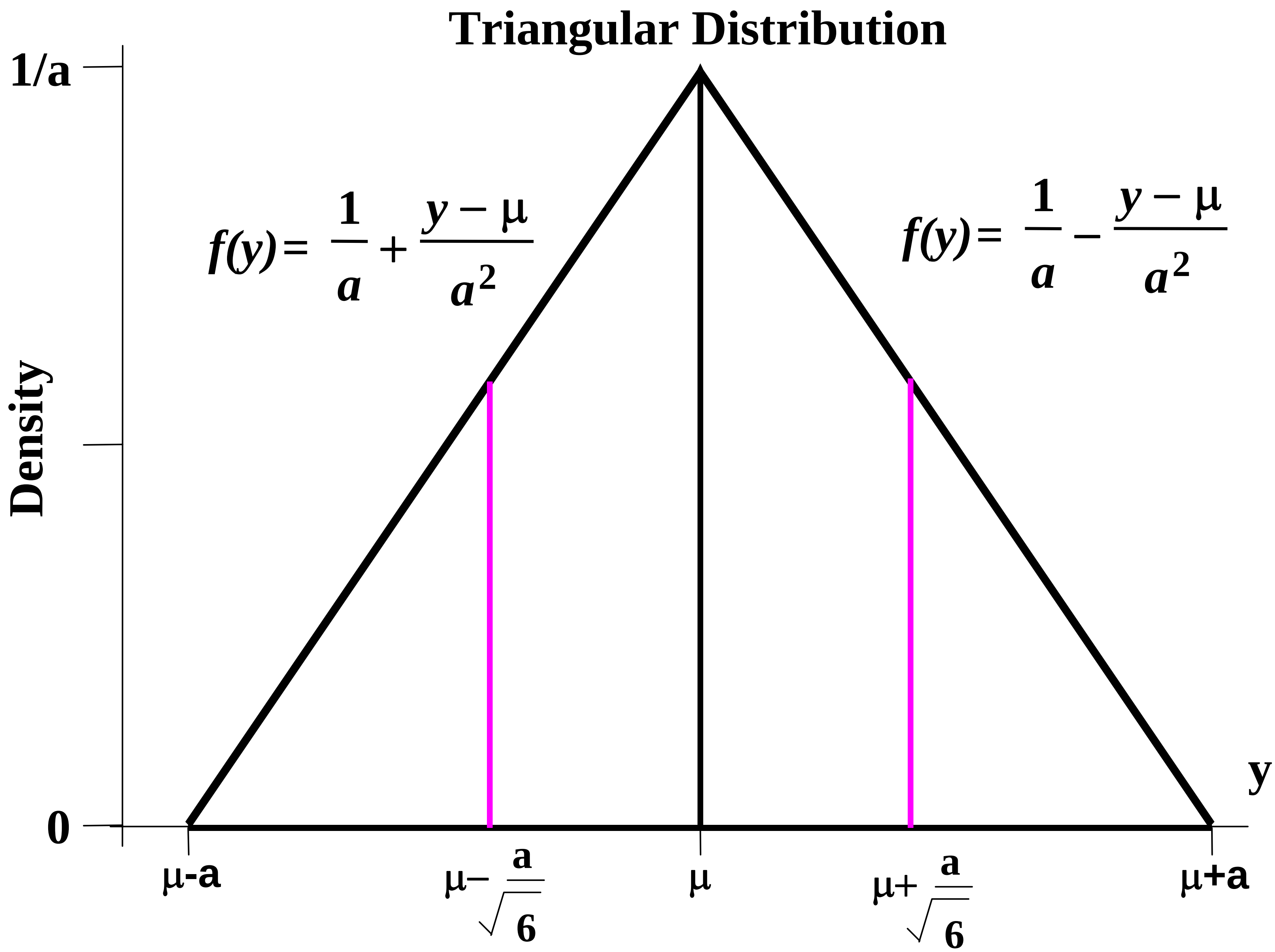
$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

Statistical Description of Error Distributions

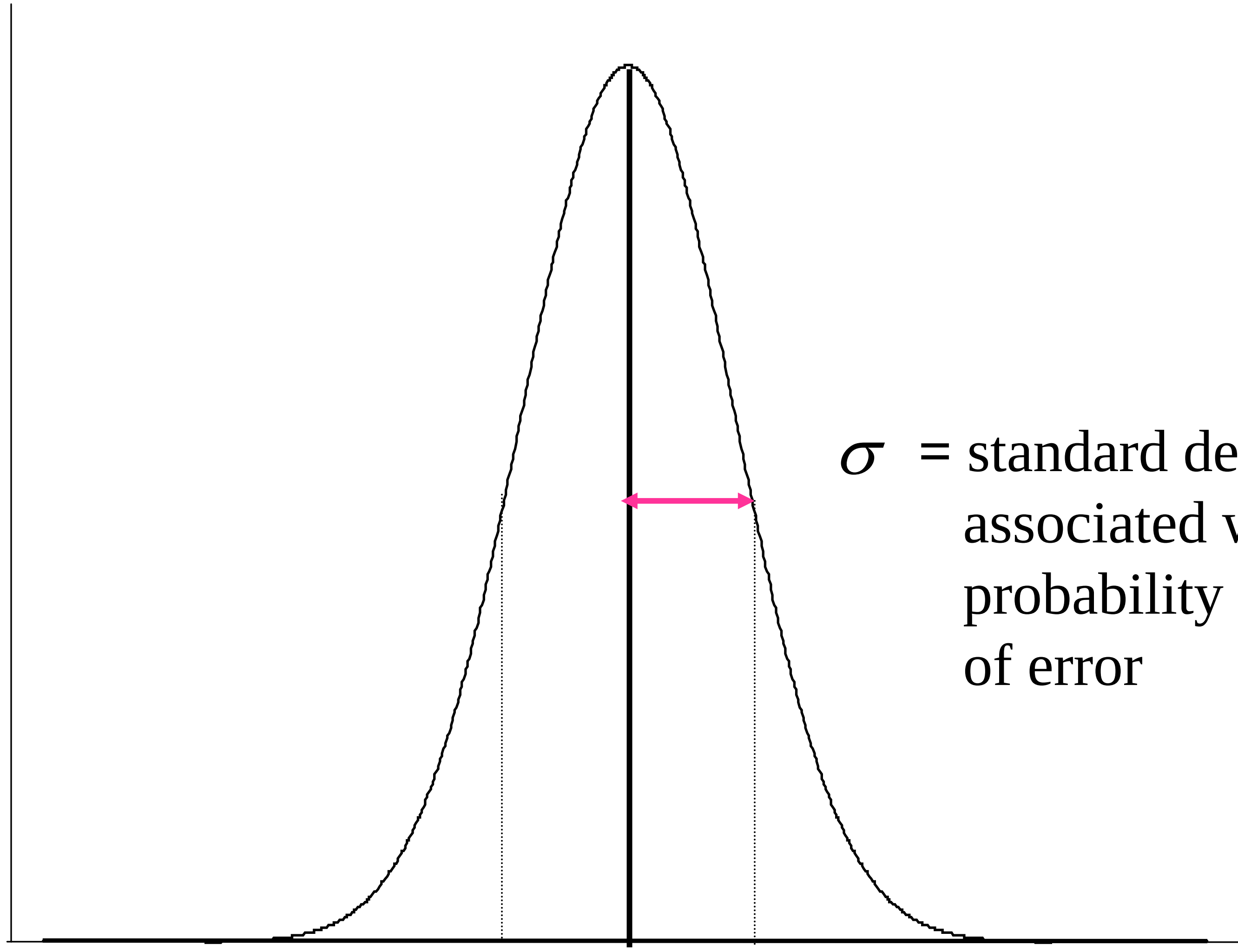
Rectangular Distribution



Statistical Description of Error Distributions



Expressing Uncertainty due to a Component of Error



σ = standard deviation associated with the probability distribution of error

Mean, Variance, & Standard Deviation of a Probability Distribution

$$\mu = \int_{-\infty}^{\infty} y \, \mathrm{f}(y) \, dy$$

$$\sigma^2 = \int_{-\infty}^{\infty} (y - \mu)^2 \, \mathrm{f}(y) \, dy$$

$$\sigma = \sqrt{\int_{-\infty}^{\infty} (y - \mu)^2 \, \mathrm{f}(y) \, dy}$$

Mean, Variance, & Standard Deviation of a Set of Repeat Measurements $y_1, y_2, \cdots, y_n$

Sample mean:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Sample variance

$$s_Y^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}$$

Sample
Standard deviation

$$s_Y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}}$$

$\bar{y}$ estimates μ and s_Y estimates σ

SECTION – 4

CLASSIFICATION & EVALUATION of UNCERTAINTY COMPONENTS

STANDARD UNCERTAINTY

Basic to the CIPM approach is representing each component of uncertainty that contributes to the uncertainty of a measurement result by an **estimated standard deviation**, termed **standard uncertainty.**

Thus, the **standard uncertainty** associated with the i^{th} component of error is the estimated standard deviation of the distribution of the collection of plausible values of that component or error. It is denoted by the (CIPM suggested) symbol u_i and is equal to the positive square root of the estimated variance u_i^2

Classification of Components of Uncertainty

The components of uncertainty, in the CIPM approach, may be grouped into two categories according to the method used to estimate their numerical values:

Type A: those which are evaluated by **statistical methods**,

Type B: those which are evaluated by **other means**.

There is not always a simple correspondence between the classification of uncertainty components into **categories A and B** and the commonly used classification of **error components** as "**random**" and "**systematic**."

ERROR versus UNCERTAINTY

- Error and Uncertainty are not synonyms.

- Error refers to the actual difference between the reported result and the actual value of the measurand

- Uncertainty refers to the estimate of the standard deviation of the probability distribution describing the collection of plausible errors.

- The result of a particular measurement can *unknowably* be very close to the unknown value of the measurand, and thus have negligible error, even though it may have a large uncertainty.

TYPE-A EVALUATION of an UNCERTAINTY COMPONENT

An uncertainty component in category A is represented by a **statistically estimated standard deviation** s_i, equal to the positive square root of the statistically estimated variance s_i^2, and the associated number of degrees of freedom ν_i.

For such a component the standard uncertainty is $u_i = s_i$.

The evaluation of uncertainty by the **statistical analysis** of series of observations is termed a **Type A evaluation (of uncertainty)**.

Type A Evaluation of Standard Uncertainty

- **A Type A evaluation of standard uncertainty may be based on any valid statistical method for treating data.**
- Examples are
  - -- Calculating the standard deviation of the mean of a series of independent observations;
  - -- Using the method of least squares to fit a curve to data in order to estimate the parameters of the curve and their standard deviations;
  - -- Carrying out an analysis of variance (ANOVA) in order to identify and quantify random effects in certain kinds of measurements.

TYPE-B EVALUATION of an UNCERTAINTY COMPONENT

An uncertainty component in category B is represented by a quantity u_j , which may be considered an approximation to the standard deviation of the probability distribution model for the corresponding component of error; this probability distribution model is arrived at based on all the available information.

For such a component the standard uncertainty is simply u_j.

The evaluation of uncertainty by means other than the statistical analysis of series of observations is termed a **Type B evaluation (of uncertainty)**.

Type B Evaluation of Standard Uncertainty

A Type B evaluation of standard uncertainty is usually based on scientific judgment using all the relevant information available, which may include

- previous measurement data,
- experience with, or general knowledge of, the behavior and property of relevant materials and instruments,
- manufacturer's specifications,
- data provided in calibration and other reports, and
- uncertainties assigned to reference data taken from handbooks.

Evaluation of Uncertainty

Type A evaluations of uncertainty based on limited data are not necessarily more reliable than soundly based Type B evaluations.

Combined Standard Uncertainty

5.1 The **combined standard uncertainty** of a measurement result, suggested symbol u_c, is taken to represent the estimated standard deviation of the *error in the* result.

It is obtained by combining the individual standard uncertainties u_i (and covariances as appropriate), whether arising from a Type A evaluation or a Type B evaluation, using the **usual method for combining standard deviations**.

This method is summarized in Appendix A [Eq. (A-3)], and is often called the *law of propagation of uncertainty* or the "root-sum-of-squares" (square root of the sum-of-the squares) or "RSS" method of combining uncertainty components estimated as standard deviations.

NOTE – The NIST policy also allows the use of established and documented methods equivalent to the "RSS" method, such as the numerically based "bootstrap" (see Appendix C).

Identification of uncertainty components

Table D.1 – Uncertainty Budget: End-Gauge Calibration

Source of uncertainty	Standard uncertainty (nm)
Calibration of standard end gauge	25 (B)
Measured difference between end gauges:	
repeated observations	5.8 (A)
random effects of comparator	3.9 (A)
systematic effects of comparator	6.7 (B)
Thermal expansion of standard end gauge	1.7 (B)
Temperature of test bed:	
mean temperature of bed	5.8 (A)
cyclic variation of temperature of room	10.2 (B)
Difference in expansion coefficients of end gauges	2.9 (B)
Difference in temperatures of end gauges	16.6 (B)
Combined standard uncertainty: $u_c(l)$ = 34 nm	

CONFIDENCE INTERVALS

5.4 In many practical measurement situations, the probability distribution characterized by the measurement result y and its combined standard uncertainty $u_c(y)$ is approximately normal (Gaussian).

When this is the case and $u_c(y)$ itself has negligible uncertainty (see Appendix B for further discussion), $u_c(y)$ defines an interval

$$y - u_c(y) \text{ to } y + u_c(y)$$

about the measurement result y within which the value of the measurand Y estimated by y is **believed** to lie with a level of confidence of approximately 68 percent. This is commonly written as $Y = y \pm u_c(y)$

NOTE – If $u_c(y)$ has non-negligible uncertainty, the level of confidence will differ from 68 percent. A procedure is given in Appendix B of TN1297 as a simple expedient approach for approximating the level of confidence in these cases. More on this later.

EXPANDED UNCERTAINTY

6.1 Although the combined standard uncertainty u_c is used to express the uncertainty of many NIST measurement results, for some commercial, industrial, and regulatory applications of NIST results (e.g., when health and safety are concerned), what is often required is a measure of uncertainty that defines **an interval about the measurement result *y* within which the value of the measurand *Y* is confidently believed to lie**.

The measure of uncertainty intended to meet this requirement is termed **expanded uncertainty**, suggested symbol U, and is obtained by multiplying $u_c(y)$ by a **coverage factor**, suggested symbol k.

Thus $U = k\, u_c(y)$ and it is confidently believed that $y - U \leq Y \leq y + U$, which is commonly written as $Y = y \pm U$.

COVERAGE FACTORS

6.2 In general, the value of the coverage factor k is chosen on the basis of the desired level of confidence to be associated with the interval defined by $U = k\, u_c$.

Typically, k is in the range 2 to 3.

When the normal distribution applies and u_c has negligible uncertainty, $U = 2u_c$ (i.e., $k = 2$) defines an interval having a level of confidence of approximately 95 percent, and $U = 3u_c$ (i.e., $k = 3$) defines an interval having a level of confidence greater than 99 percent.

NOTE – For a quantity z described by a normal distribution with mean μ_z and standard deviation σ, the interval $\mu_z \pm k\,\sigma$ encompasses 68.27, 90, 95.45, 99, and 99.73 percent of the distribution for $k = 1$, $k = 1.645$, $k = 2$, $k = 2.576$, and $k = 3$, respectively (see the last line of Table B.1 of Appendix B).

RELATIVE UNCERTAINTIES

If $u(x_i)$ is a standard uncertainty, then $u(x_i)/|x_i|$, $x_i \neq 0$, is the corresponding ***relative standard uncertainty***;

if $u_c(y)$ is a combined standard uncertainty, then $u_c(y)/|y|$, $y \neq 0$, is the corresponding ***relative combined standard uncertainty***; and

if $U = k\, u_c(y)$ is an expanded uncertainty, then $U/|y|$, $y \neq 0$, is the corresponding ***relative expanded uncertainty***.

SECTION – 5

TAYLOR SERIES

and

PROPAGATION OF UNCERTAINTY

Law of Propagation of Uncertainty

A.1 In many cases a measurand Y is not measured directly, but is determined from N other quantities $X_1, X_2, \ldots, X_N$ through a functional relation f:

$$Y = f(X_1, X_2, \ldots, X_N) \qquad \textbf{(A-1)}$$

Included among the quantities X_i are corrections (or correction factors) as described in subsection 5.2, as well as quantities that take into account other sources of variability, such as different observers, instruments, samples, laboratories, and times at which observations are made (e.g., different days).

Thus the function f should express **not simply a physical law** but a measurement process, and in particular, it should contain all quantities that can contribute a significant uncertainty to the measurement result.

Law of Propagation of Uncertainty

A.2 An estimate of the measurand or *output quantity* Y, denoted by y, is obtained from Eq. (A-1) using *input estimates* $x_1, x_2, \ldots, x_N$ for the values of the N input quantities $X_1, X_2, \ldots, X_N$.

Thus the *output estimate* y, which is the result of the measurement, is given by

$$y = f(x_1, x_2, \ldots, x_N) \,. \qquad \text{(A-2)}$$

Law of Propagation of Uncertainty

A.3 The combined standard uncertainty of the measurement result *y*, designated by $u_c(y)$ and taken to represent the estimated standard deviation of the result, is the positive square root of the estimated variance $u_c^2(y)$ obtained from (A-3)

$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) \quad + \quad 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \qquad \text{(A-3)}$$

- Equation (A-3) is based on a first-order Taylor series approximation of $Y = f(X_1, X_2, \ldots, X_N)$ and is conveniently referred to as the *law of propagation of uncertainty*.

- The partial derivatives $\partial f/\partial x_i$ (often referred to as *sensitivity coefficients*) are equal to $\partial f/\partial X_i$ evaluated at $X_i = x_i$;

- $u(x_i)$ is the standard uncertainty associated with the input estimate x_i; and $u(x_i, x_j)$ is the estimated covariance associated with x_i and x_j.

Law of Propagation of Uncertainty

Sometimes the propagation of uncertainty equation is written as follows:

$$u_c^2(y) = \sum_{i=1}^{N} c_{x_i}^2 u^2(x_i) \quad + \quad 2\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} c_{x_i,x_j} \; u(x_i, x_j)$$

Here, the coefficient c_{x_i} is the sensitivity coefficient associated with x_i that accounts for the uncertainty in x_i , and

the coefficient c_{x_i,x_j} is the sensitivity coefficient that accounts for the correlation between x_i and x_j

The propagation of uncertainty formula is based on the following general result:

If $Y = c_1 X_1 + c_2 X_2 + \cdots + c_N X_N$ then

$$u^2(Y) = \sum_{i=1}^{N} c_i^2 u^2(X_i) + 2\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} c_i \, c_j \, u(X_i, X_j)$$

EXAMPLE EVALUATION – TYPE A UNCERTAINTY

A.4 As an example of a Type A evaluation, consider an input quantity X_i whose value is estimated from n independent observations $X_{i,k}$ of X_i obtained under the same conditions of measurement. In this case the input estimate x_i is usually the sample mean

$$x_i = \bar{X}_i = \frac{1}{n}\sum_{k=1}^{n} X_{i,k} \qquad \text{(A-4)}$$

and the standard uncertainty $u(x_i)$ to be associated with x_i is the estimated standard deviation of the mean

$$u(x_i) = s(\bar{X}_i) = \left(\frac{1}{n(n-1)}\sum_{k=1}^{n}(X_{i,k} - \bar{X}_i)^2\right)^{1/2} \qquad \text{(A-5)}$$

EXAMPLE EVALUATION – TYPE B UNCERTAINTY

A.5 As an example of a Type B evaluation, consider an input quantity X_i whose value is estimated from an assumed rectangular probability distribution of lower limit a_- and upper limit a_+. In this case the input estimate is usually the expectation of the distribution

$$x_i = \frac{a_+ + a_-}{2} \qquad \text{(A-6)}$$

and the standard uncertainty $u(x_i)$ to be associated with x_i is

$$u(x_i) = a/\sqrt{3} \qquad \text{(A-7)}$$

the positive square root of the variance of the distribution where $a = (a_+ - a_-)/2$.

NOTE – When x_i is obtained from an assumed distribution, the associated variance is appropriately written as $u^2(X_i)$ and the associated standard uncertainty as $u(X_i)$, but for simplicity, $u^2(x_i)$ and $u(x_i)$ are used. Similar considerations apply to the symbols $u_c^2(y)$ and $u_c(y)$.

THE FOUR STEP PROCEDURE

B.3 The four-step procedure for calculating k_p is as follows:

1) Obtain y and $u_c(y)$ as indicated in Appendix A.

2) Estimate the effective degrees of freedom ν_{eff} of $u_c(y)$ from the **Welch-Satterthwaite** formula (B-1).

3) Obtain the t-factor $t_p(\nu_{eff})$ for the required level of confidence p from a table of values of $t_p(\nu)$ from the t-distribution, such as Table B.1 of Appendix-B.

4) Take $k_p = t_p(\nu_{eff})$ and calculate $U_p = k_p u_c(y)$.

SECTION – 6

NIST POLICY

on

REPORTING UNCERTAINTY

NIST POLICY ON REPORTING UNCERTAINTY

The stated NIST policy regarding reporting uncertainty is (see Appendix C):

- **Report *U* together with the coverage factor *k* used to obtain it**, or **report u_c.**
- When reporting a measurement result and its uncertainty, **include the following information** in the report itself or by referring to a published document:
  - **A list of all components of standard uncertainty**, together with **their degrees of freedom where appropriate**, and the resulting **value of u_c**. The components should be identified according to the method used to estimate their numerical values:
    - **Type A**: those which are evaluated by statistical methods,
    - **Type B**: those which are evaluated by other means.
  - **A detailed description of how each component of standard uncertainty was evaluated**.

NIST POLICY ON REPORTING UNCERTAINTY

- To be consistent with current international practice, the value of k to be used at NIST for calculating U is, **by convention, k = 2**. Values of k other than 2 are only to be used for specific applications dictated by established and documented requirements.

- If a value other than 2 is chosen for k, a description of how k was chosen is to be given.

- It is often desirable to provide a probability interpretation, such as a level of confidence, for the interval defined by U or u_c.

- When this is done, the basis for such a statement must be given.

NIST POLICY ON REPORTING UNCERTAINTY

- An example of the use of a value of *k* other than 2 is taking *k* equal to a *t*-factor obtained from the *t*-distribution when u_c has low degrees of freedom in order to meet the dictated requirement of providing a value of $U = k\,u_c$ that defines an interval having a level of confidence close to 95 percent.

(The 4-step procedure is expected to provide a value of *k* such that $U = k\,u_c$ defines an interval having a level of confidence close to 95 percent)

EXCEPTIONS TO NIST POLICY

6.6 The NIST policy provides for exceptions as follows (see Appendix C):

It is understood that **any valid statistical method that is technically justified under the existing circumstances may be used to determine the equivalent of u_i , u_c, or U**.

Further, it is recognized that international, national, or contractual agreements to which NIST is a party may occasionally require deviation from NIST policy.

In both cases, the report of uncertainty must document what was done and why.

EXAMPLES

7.3 NIST policy on reporting uncertainty refers to the desirability of providing a probability interpretation, such as a level of confidence, for the interval defined by U or u_c.

The following examples show how this might be done when the numerical result of a measurement and its assigned uncertainty are reported, assuming that the published detailed description of the measurement provides a sound basis for the statements made.

(In each of the three cases, the quantity whose value is being reported is assumed to be a nominal 100 g standard of mass m_s .)

EXAMPLE - 1

m_s = (100.021 47 ± 0.000 70) g,

where the number following the symbol ± is the numerical value of an expanded uncertainty $U = ku_c$, with U determined from a combined standard uncertainty (i.e., estimated standard deviation) u_c = 0.35 mg and a coverage factor k = 2.

If it can be assumed that the possible estimated values of the standard are approximately normally distributed with approximate standard deviation u_c, the unknown value of the standard can be believed to lie in the interval defined by U with a level of confidence of approximately 95 percent.

EXAMPLE - 2

m_S = (100.021 47 ± 0.000 79) g,

where the number following the symbol ± is the numerical value of an expanded uncertainty $U = ku_c$, with U determined from a combined standard uncertainty (i.e., estimated standard deviation) u_c = 0.35 mg and a coverage factor k = 2.26 based on the t-distribution for $\nu = 9$ degrees of freedom.

It defines an interval within which the unknown value of the standard is believed to lie with a level of confidence of approximately 95 percent.

SECTION – 7

AN EXAMPLE

End-Gauge Calibration

The Measurement Problem

- The length of a nominally 50 mm end gauge is determined by comparing it with a known standard of the same nominal length.

- The direct output of the comparison of the two end gauges is the difference *d* in their lengths:

$$d = l\,(1 + \alpha\,\theta) - l_S\,(1 + \alpha_S\,\theta_S\,)$$

l = **the measurand**, i.e., the length at 20 deg C of the end gauge being calibrated

l_S = the length of the standard at 20 deg C, as given in its calibration certificate

α = coefficient of thermal expansion of the gauge being calibrated

α_S = coefficient of thermal expansion of the standard

θ = deviation in temperature for the test gauge from 20 deg C

θ_S = deviation in temperature for the standard from 20 deg C

Measurement Equation

The measurement equation is
$$l = \frac{l_S(1+\alpha_S\theta_S) + d}{(1+\alpha\theta)}$$

It may be re-expressed as
$$l \approx \big(l_S(1+\alpha_S\theta_S) + d\big)(1-\alpha\theta)$$
$$\approx l_S + d + l_S(\alpha_S\theta_S - \alpha\theta)$$
$$= l_S + d - l_S(\delta\alpha\,\theta - \alpha_S\,\delta\theta)$$

where $\delta\alpha = \alpha - \alpha_S$ $\qquad \delta\theta = \theta - \theta_S$

- $\delta\alpha$ and $\delta\theta$ are **estimated** to be zero, but their uncertainties are not zero
- d is estimated using $\bar{d}$ = the arithmetic mean of n = 5 independent repeated observations.
- The best estimate of l is $l_S + \bar{d}$

Propagation of Uncertainty

Recall

$$u_c^2(y) = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) \quad + \quad 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}u(x_i, x_j)$$

Assuming that the component input quantities are **uncorrelated**, we get

$$u_c^2(l) = \left(\frac{\partial f}{\partial l_S}\right)^2 u^2(l_S) + \left(\frac{\partial f}{\partial d}\right)^2 u^2(d) + \left(\frac{\partial f}{\partial \alpha_S}\right)^2 u^2(\alpha_S)$$
$$+ \left(\frac{\partial f}{\partial \theta}\right)^2 u^2(\theta) + \left(\frac{\partial f}{\partial \delta\alpha}\right)^2 u^2(\delta\alpha) + \left(\frac{\partial f}{\partial \delta\theta}\right)^2 u^2(\delta\theta)$$

Propagation of Uncertainty *(continued)*

$$\frac{\partial f}{\partial l_S} = c_{l_S} = 1 - (\delta\alpha\,\theta + \alpha_S\,\delta\theta) = 1 \qquad \frac{\partial f}{\partial d} = c_d = 1$$

$$\frac{\partial f}{\partial \alpha_S} = c_{\alpha_S} = -l_S\,\delta\theta = 0 \qquad \frac{\partial f}{\partial \theta} = c_\theta = -l_S\,\delta\alpha = 0$$

$$\frac{\partial f}{\partial \delta\alpha} = c_{\delta\alpha} = -l_S\,\theta \qquad \frac{\partial f}{\partial \delta\theta} = c_{\delta\theta} = -l_S\,\alpha_S$$

Hence we have

$$u_c^2(l) = u^2(l_S) + u^2(d) + l_S^2\,\theta^2\,u^2(\delta\alpha) + l_S^2\,\alpha_S^2\,u^2(\delta\theta)$$

We now discuss the evaluation of each uncertainty component

$u(l_S)$ = Uncertainty of the Calibration of the Standard

The calibration certificate gives as the expanded uncertainty of the standard, the value U = 0.075 μm, and states that it was obtained using a coverage factor of *k* = 3. Hence the **standard uncertainty** is

$$u(l_s) = \frac{0.075\,\mu\text{m}}{3} = 0.025\,\mu\text{m} = 25\,\text{nm}$$

u(d) = Uncertainty of the Measured Difference in Lengths

A single measured value of *d*, denoted by d_i, is related to *d* as follows:

$$d_i = d + e_{\text{comparator}} + e_{\text{repeatability}}$$

So we have

$$\bar{d} = d + e_{\text{comparator}} + \bar{e}_{\text{repeatability}}$$

Error associated with measurement of *d* is

$$\bar{d} - d = e_{\text{comparator}} + \bar{e}_{\text{repeatability}}$$

Hence the uncertainty of the measured difference in lengths is

$$u^2(e_{\text{comparator}}) + u^2(\bar{e}_{\text{repeatability}})$$

$u^2(\bar{e}_{\text{repeatability}})$ = Uncertainty of the Measured Difference in Lengths Due to Repeatability Error

Repeatability error, using past data, is known to have a probability distribution whose mean is zero and standard deviation is estimated to be 13 nm based on 25 repeats in an auxiliary experiment. So

$$s_{\text{repeatability}} = 13 \text{ nm}$$

Since is calculated as the arithmetic mean of n = 5 repeat observations, the standard uncertainty due to repeatability error is

$$u^2(\bar{e}_{\text{repeatability}}) = \frac{s^2_{\text{repeatability}}}{5}$$

$$u(\bar{e}_{\text{repeatability}}) = \frac{s_{\text{repeatability}}}{\sqrt{5}} = \frac{13\,\text{nm}}{\sqrt{5}} = 5.8\,\text{nm}$$

$u^2(e_{\text{comparator}})$ = Uncertainty of the Measured Difference in Lengths Due to Calibration Error

According to the calibration certificate of the comparator used to compare l with l_S, the calibration uncertainty has a component due to random errors, which we write as $u(d_1)$, and a component due to systematic errors, which we write as $u(d_2)$.

The uncertainty "due to random errors" is quoted as $\pm$ 0.1 μm at a 95 % confidence level, and is based on 6 replicate measurements. So the standard uncertainty, using a t-factor $t_{95}(5) = 2.57$ for $\nu = 6 - 1 = 5$ df, we have

$$u(d_1) = (0.01\,\mu\text{m}) / 2.57 = 3.9\text{ nm}$$

The uncertainty "due to systematic errors" is quoted as $\pm$ 0.02 μm at the "three-sigma" level. The standard uncertainty from this cause is therefore taken to be

$$u(d_2) = (0.02\,\mu\text{m}) / 3 = 6.7\text{ nm}$$

u(d) = Uncertainty of the Measured Difference in Lengths

$$u^2(d) = u^2(e_{\text{comparator}}) + u^2(\bar{e}_{\text{repeatability}})$$

$$= u^2(d_1) + u^2(d_2) + u^2(\bar{e}_{\text{repeatability}})$$

$$= (3.9\,\text{nm})^2 + (6.7\,\text{nm})^2 + (5.8\,\text{nm})^2 = 93\,\text{nm}^2$$

$$u(d) = 9.7\,\text{nm}$$

$u^2(\alpha_S)$ = Uncertainty of the Thermal Expansion Coefficient

The coefficient of thermal expansion of the standard end gauge is given as $\alpha_S = 11.5 \times 10^{-6}$ deg C^{-1} with an uncertainty represented by a rectangular distribution with bounds $\pm\ 2 \times 10^{-6}$ deg C^{-1}.

The standard uncertainty is therefore

$$u(\alpha_S) = (2 \times 10^{-6}\ {}^\circ C^{-1}) / \sqrt{3} = 1.2 \times 10^{-6}\ {}^\circ C^{-1}$$

Because the sensitivity coefficient corresponding to α_S is 0, this uncertainty contributes nothing to the uncertainty $u(d)$.

$u^2(\theta)$ = Uncertainty of the Deviation of the Temperature of the End Gauge

The temperature of the test bed is reported as $(19.9 \pm 0.5)^\circ C$

The temperature at the time of the individual observations was not recorded. The stated maximum offset, Δ = 0.5 deg C, is said to represent the amplitude of an approximately cyclical variation of the temperature under a thermostatic system, not the uncertainty of the mean temperature.

The value of the mean temperature deviation

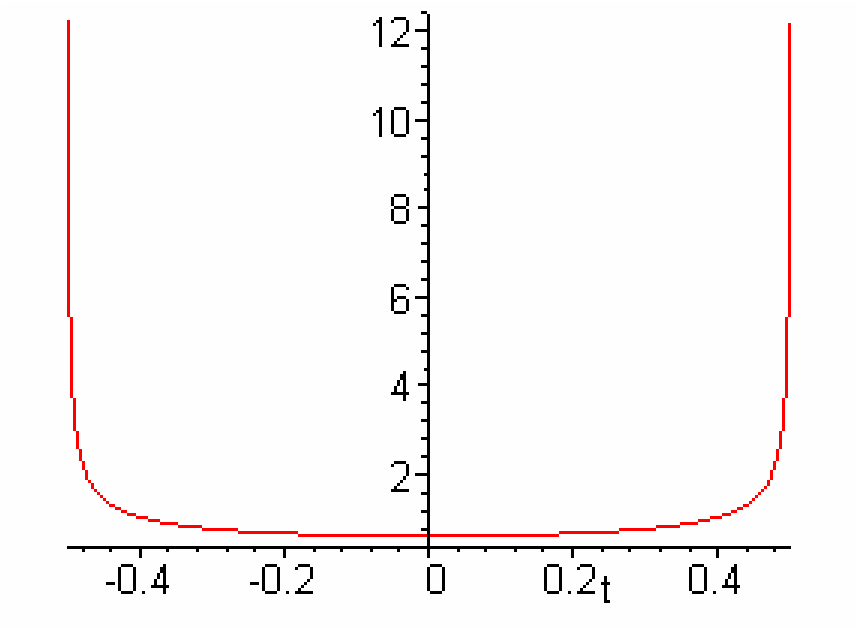
$$\bar{\theta} = 19.9 \text{ deg C} - 20.0 \text{ deg C} = -0.1 \text{ deg C},$$

is reported as having a standard uncertainty itself due to the uncertainty in the mean temperature of the test bed of

$$u(\bar{\theta}) = 0.2^\circ C$$

$u^2(\theta)$ = Uncertainty of the Deviation of the Temperature of the End Gauge

The cyclic variation of temperature over time produces a U-shaped Distribution shown below.



$$f := t \to \frac{1}{\pi \sqrt{\frac{1}{4} - t^2}}$$

The standard deviation of the above distribution is sqrt(1/8) = 0.35

$u^2(\theta)$ = Uncertainty of the Deviation of the Temperature of the End Gauge

The measurement model for the temperature is $\theta = \theta_{\text{average}} + e_\theta$

where θ_{average} is the mean temperature of the test bed and e_θ is the deviation from the mean due to the cyclic variation

$$u^2(\theta) = u^2(\bar{\theta}) + u^2(e_\theta) = (0.2\,^\circ C)^2 + (\sqrt{1/8})^2 = 0.165\,^\circ C^{\,2}$$

which gives $u(\theta) = 0.41\,^\circ C$

However, $u(\theta)$ is not involved in the calculation of $u_c(l)$ because the sensitivity coefficient associated with it is zero. The calculations are shown here for illustration only.

$u^2(\delta\alpha)$ = Uncertainty of the Difference in Expansion Coefficients

The estimated bounds on the variability of $\delta\alpha$ are $\pm$ 1 x 10^{-6} deg C^{-1} with an equal probability of $\delta\alpha$ having any value within these bounds. Hence, a rectangular distribution may be used to obtain the standard uncertainty.

$$u(\delta\alpha) = (1 \times 10^{-6}\ {}^{\circ}C^{-1}) / \sqrt{3} = 0.58 \times 10^{-6}\ {}^{\circ}C^{-1}$$

$u^2(\delta\theta)$ = Uncertainty of the Difference in Temperatures of the Gauges

The standard gauge and the test gauge are expected to be at the same temperature, but the temperature difference could lie with equal probability anywhere in the estimated interval (– 0.05 ^oC , +0.05 ^oC).

Thus, using a rectangular distribution to model the difference in temperature, we get

$$u(\delta\theta) = (0.05\,^\circ C)/\sqrt{3} = 0.029\,^\circ C$$

$u_c^2(l)$ = Combined Standard Uncertainty of l

$$u_c^2(l) = (25\,\text{nm})^2 + (9.7\,\text{nm})^2$$

$$+(0.05\,\text{m})^2(-0.1\,^\circ\text{C})^2(0.58\times10^{-6}\,^\circ\text{C})^2$$

$$+(0.05\,\text{m})^2(11.5\times10^{-6}\,^\circ\text{C}^{-1})^2(0.029\,^\circ\text{C})^2$$

$$= 1002\,\text{nm}^2$$

$$u_c(l) = 32\,\text{nm}$$

Final Result

The calibration certificate for the standard gauge gives l_S = 50.000 623 mm as its length at 20 deg C.

The arithmetic mean $\bar{d}$ of the n=5 repeat observations of the difference in lengths between the standard gauge and the test gauge is 215 nm.

Hence the measurement result is

$$l = l_S + \bar{d} = 50.000623\,\text{mm} + 0.000215\,\text{mm} = 50.000838\,\text{mm}$$

The final statement is
"l = 50.000838 mm with a combined standard uncertainty u_c = 32 nm"

The corresponding relative combined standard uncertainty is given by
$u_c/l = 6.4 \times 10^{-7}$

Calculation of Expanded Uncertainty

Use the 4-step procedure discussed earlier. For this, we need the degrees of freedom values for each component of uncertainty.

For type-A evaluations, the particular statistical method used will also provide a value for the degrees of freedom.

For type-B evaluations, the degrees of freedom is often taken to be infinity (i.e., the uncertainty component is regarded as exactly known), or, as discussed earlier, one may use

$$\nu_i = \frac{1}{2}\left(\frac{u(x_i)}{\Delta u(x_i)}\right)^2$$

where $\Delta u(x_i)/u(x_i)$ is the estimated relative error in the value $u(x_i)$.

Calculation of Expanded Uncertainty

1) **DF for the Uncertainty of the calibration of the standard:**

The calibration certificate states that the effective degrees of freedom of the combined standard uncertainty from which the quoted expanded uncertainty was obtained is $\nu_{eff}\ (l_S) = 18$.

Calculation of Expanded Uncertainty

2) DF for the Uncertainty of the measured difference in lengths:

Although $\bar{d}$ was calculated based on n = 5 repeat measurements, its standard uncertainty was calculated from a pooled experimental standard deviation based on 25 observations. Hence $\nu(\bar{d})$ = 25 – 1 = 24.

The degrees of freedom for $u(d_1)$, the uncertainty due to random errors in the comparator, is $\nu(d_1)$ = 6 – 1 = 5, because d1 was obtained from six repeated measurements.

The $\pm$ 0.02 mm uncertainty for systematic errors of the comparator is thought to be reliable to 25 percent. Hence, the degrees of freedom associated with $u(d_2)$ is

$$\nu(d_2) = \tfrac{1}{2}\left(\frac{u(x_i)}{\Delta u(x_i)}\right)^2 = \tfrac{1}{2}(4)^2 = 8$$

Calculation of Expanded Uncertainty

Hence, the effective degrees of freedom associated with $u(d)$, denoted by, $\nu_{\mathrm{eff}}(d)$ is calculated as follows using the Welch-Satterthwaite formula:

$$\nu_{\mathrm{eff}}(d) = \frac{[u(d)]^4}{\dfrac{u^4(\bar{d})}{\nu(\bar{d})} + \dfrac{u^4(d_1)}{\nu(d_1)} + \dfrac{u^4(d_2)}{\nu(d_2)}}$$

$$= \frac{(9.7\,nm)^4}{\dfrac{(5.8\,nm)^4}{24} + \dfrac{(3.9\,nm)^4}{5} + \dfrac{(6.7\,nm)^4}{8}} = 25.6$$

Calculation of Expanded Uncertainty

3) DF for $u(\delta\alpha)$, the uncertainty of the difference in expansion coefficients

The estimated bounds of $\pm$ 1 x 10^{-6} $^{o}C^{-1}$ are deemed to have a relative error of 10 percent. Hence the degrees of freedom for $u(\delta\alpha)$ is taken to be

$$\nu(\delta\alpha) = \tfrac{1}{2}(1/0.1)^2 = 50$$

4) DF for $u(\delta\theta)$, the uncertainty of the difference in temperatures of the gauges

The estimated interval $\pm$ 0.05 ^{o}C for the temperature difference $\delta\theta$ is believed to have a relative error of 50 percent, which gives $\nu(\delta\theta) = 2$.

Calculation of Expanded Uncertainty

The calculation of the effective degrees of freedom for *u*(*l*) uses the Welch-Satterthwaite formula.

$$\nu_{\text{eff}}(l) = \frac{(32\,nm)^4}{\dfrac{(25\,nm)^4}{18} + \dfrac{(9.7\,nm)^4}{25.6} + \dfrac{(2.9\,nm)^4}{50} + \dfrac{(16.6\,nm)^4}{2}}$$

$$= 16.7 \approx 16 \quad \text{(truncated to the integer part)}$$

The t-factor with 16 df, for a 99 percent confidence level, is 2.92. Hence the expanded uncertainty (k = 2.92) is

$$U_{99} = (2.92)(32\,nm) = 93\,nm$$

Statement of the Final Result

l = (50.000 838 $\pm$ 0.000 093) mm, where the number following the symbol $\pm$ is the expanded uncertainty $U = k\, u_c$, with U determined from a combined standard uncertainty u_c = 32 nm and a coverage factor k = 2.92 based on the t-distribution with ν = 16 df, and defines an interval estimated to have a level of confidence of 99 percent.

The corresponding relative expanded uncertainty is $U/l = 1.9 \times 10^{-6}$

Uncertainty Budget

Standard Uncertainty component u(xi)		Source of Uncertainty	Type	Value of Standard Uncertainty u(xi)		ci = sensitivity coefficient	Contribution to u(l) from the component	df
u(lS)		Calibration of Standard End Gauge	B(syst)	25 nm		1	25	18
u(d)		Measured difference between end gauges		9.7 nm		1	9.7	25.6
	u(dbar)	repeated observations	A(rand)		5.8 nm			24
	u(d1)	random errors in comparator	A(syst)		3.9 nm			5
		systematic errors in comparator	B(syst)		6.7 nm			8
u(alpha_S)		Thermal expansion coefficient of standard gauge	B(syst)	1.2 x 10^(-6) /deg C		0	0	
u(theta)		Temperature of test bed		0.41 deg C		0	0	
	u(theta_bar)	mean temperature of test bed	A(rand)		0.2 deg C			
		cyclic variation of temperature of room	B(syst)		0.35 deg C			
u(delta_alpha)		Difference in expansion coefficients of gauges	B(syst)	0.58x10^(-6) /degC		-lS theta	2.9	50
u(delta_theta)		Difference in temperatures of gauges	B(rand)	0.029 deg C		-lS alpha_S	16.6	2