Evaluating Measurement Uncertainty

Antonio Possolo

October 30th, 2013

NIST

SIM METROLOGY SCHOOL

Uncertainty Evaluation in a Nutshell

- Measurement uncertainty
  - Reflects incomplete knowledge about value of measurand
  - Expressed most completely by probability distribution
- Experimental data may be used alone or combined with other information to
  - Estimate measurand
  - Evaluate measurement uncertainty — Bottom Up *vs.* Top Down

Measurement equation	*Observation equation*
Monte Carlo / Gauss	Mathematical Statistics
NIST Uncertainty Machine	Expert collaboration

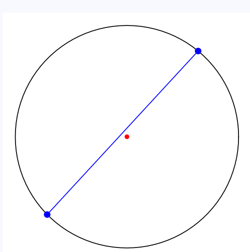
Uncertainty Evaluation — Area of Circle

MEASUREMENT EQUATION & GAUSS'S FORMULA

- Measure diameter D, compute radius $R = D/2$, estimate measurand $A = \pi R^2$

GAUSS'S FORMULA

$$\frac{u(A)}{A} \approx 2\frac{u(R)}{R}$$



$D = 5.6922\,\text{m}$

$R = 2.8461\,\text{m}$ $u(R) = 0.005\,\text{m}$

$A = 25.447\,\text{m}^2$ $u(A) \approx 0.089\,\text{m}^2$

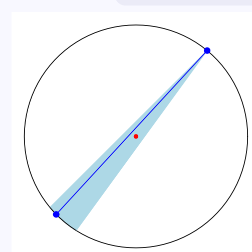
Uncertainty Evaluation — Area of Circle

MEASUREMENT EQUATION & MONTE CARLO METHOD

- Measure diameter D, compute radius $R = D/2$, estimate measurand $A = \pi R^2$

MONTE CARLO METHOD

- Specify probability model to describe uncertainty when defining diameter
- Simulate process many times
- Assess dispersion of area estimates



Diameter has Gaussian azimuth with no bias and uncertainty $\pi/30$

$A = 25.447\,\text{m}^2$ $u(A) = 0.099\,\text{m}^2$

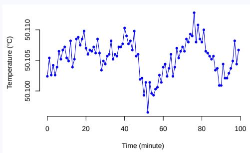
Uncertainty Evaluation — Thermal Bath

OBSERVATION EQUATION & STATISTICAL MODEL

- Measure temperature of thermal bath every minute for 100 minutes
- Determine if bath is thermally stable, and if so evaluate uncertainty of average temperature

ARIMA STATISTICAL MODEL

- Identify best ARIMA model
- Fit best model to data



Best model is stationary:
bath thermally stable

$t = 50.105\,°\mathrm{C}$ $u(t) = 0.001\,°\mathrm{C}$

Outline

- Uncertainty Evaluation in a Nutshell
  - Area of Circle — Gauss's Formula
  - Area of Circle — Monte Carlo Method
  - Thermal Bath — Statistical Model
- Probability Distributions & Random Variables
- Measurement Uncertainty & Measurement Error
- Measurement Models
  - Measurement Equations
  - Observation Equations
- Evaluation of Measurement Uncertainty
  - *NIST Uncertainty Machine*
- Load Cell Calibration

Uncertainty — Meaning

MEANING

- Uncertainty is the condition of being *uncertain* (unsure, doubtful, not possessing complete or fully reliable knowledge)
  - Also a qualitative or quantitative expression of the degree or extent of such condition

  It is a subjective condition because it pertains to the perception or understanding that ***you*** *have of the object of interest*

Uncertainty — Interpretation (GUM)

INTERPRETATION (GUM)

- *The uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurand* — GUM [3.3.1]
  - State of knowledge described most completely by probability distribution over set of possible values of measurand

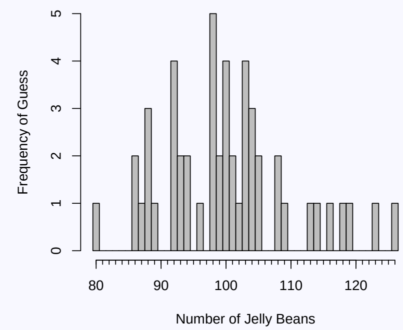
    Probability distribution expresses how well one believes one knows the measurand's true value

Outline

Probability Distributions — Motivation

- Measurand: Number of jelly beans in the jar

- Frequency distribution of students' estimates
  - Depicts dispersion of measurement results
  - Characterizes students' collective uncertainty



Probability Distributions & Random Variables

CONCEPTS

- Probability distribution specifies probability of unknown value of a quantity Y being in any given subset of its range $\mathcal{B}$

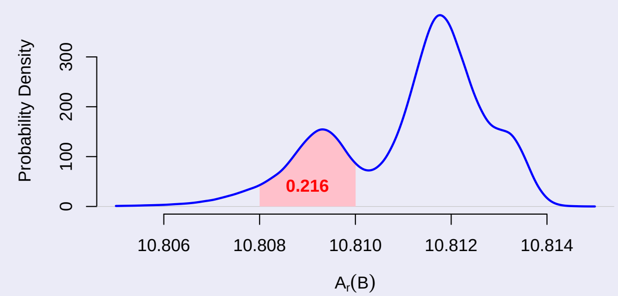
  $$\Pr(Y \in B), \text{ for } B \text{ a subset of } \mathcal{B}$$

- Random variable is a mathematical model for unknown value of a quantity that has associated a probability distribution
  - All quantities about whose values there is uncertainty can be modeled as random variables
    - Even if the quantity value is fixed (but unknown)
    - Irrespective of whether they relate to *chance* events

Probability Distributions & Random Variables

EXAMPLE

EXAMPLE — ATOMIC WEIGHT OF BORON



- Random variable $A_r(B)$ describes atomic weight of boron in sample known to come from one of main commercial sources in US, Turkey, Chile, Argentina, or Russia
- Probability density gives probability of $A_r(B)$'s value being in any given interval as area under curve

Outline

Measurement Uncertainty

VIM / GUM DEFINITION

DEFINITION

- *Measurement uncertainty* is a non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used — VIM 2.26
  - For scalar measurands, that *parameter* typically chosen to be standard deviation of probability distribution describing dispersion of values
  - *For the many measurands that are not scalars (spectra, maps, shapes, genomes, etc.) use suitable generalization*

Measurement Uncertainty — Sources

SOURCES / EFFECTS

- Principles and methods of measurement
- Definition of measurand
- Environmental conditions
- Instrument calibration
- Calibration standards and corrections
- Temporal drifts of relevant attributes of measurand, instruments, procedures
- Differences between operators and between laboratories
- Statistical models and methods for data reduction

Measurement Error

- Difference between measured quantity value and “true” (or, reference) quantity value
  - Not observable, and generally unknown
  - If measurand has conventional value (speed of light), then that difference can be computed and measurement error becomes known
- Contemporary approach emphasizes uncertainty, as opposed to error
  - Facilitates use of uniform methods for uncertainty analysis

Measurement Uncertainty — Evaluation

- Evaluated so as to be fit for purpose
- Sources of uncertainty may be evaluated based on:
  - Experimental data (Type A evaluations)
  - Other sources of information (Type B evaluations)

    Elicitation of expert opinion

    — structured procedure to do Type B evaluations, for example using Sheffield Elicitation Framework (SHELF)

Measurement Error — Fixed (Systematic)

FIXED (*Systematic*) MEASUREMENT ERROR

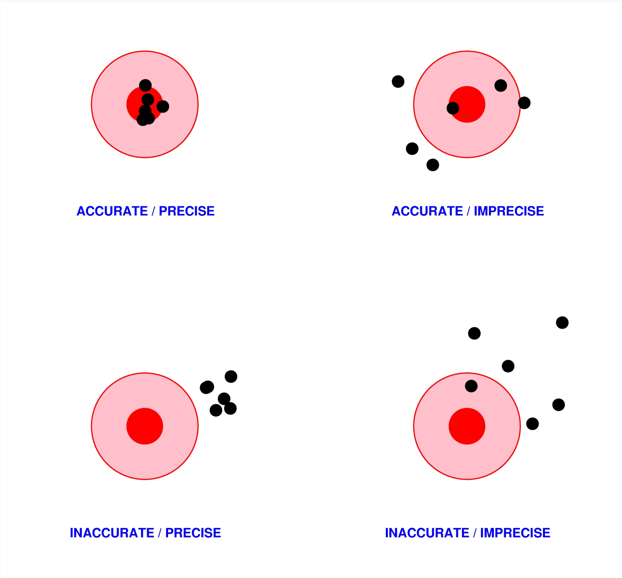
- Component of measurement error that, in replicate measurements, either remains constant or varies in a predictable manner
  - Incomplete extraction when measuring mass fractions of multiple PCB congeners
  - Wrong value of thermal expansion coefficient in length measurements
- Evaluated either by Type A or Type B methods

Measurement Error — Variable (Random)

VARIABLE (*Random*) MEASUREMENT ERROR

- Component of measurement error that in replicate measurements varies in an unpredictable manner
  - Manifest in dispersion of multiple measured values obtained under *repeatability conditions*
    - Same measurement procedure, same operators, same measuring system, same operating conditions and same location, made on same object over short period of time
- Evaluated either by Type A or Type B methods

Accuracy & Precision — Illustration



Accuracy & Precision — Definition

MEASUREMENT ACCURACY (VIM 2.13)

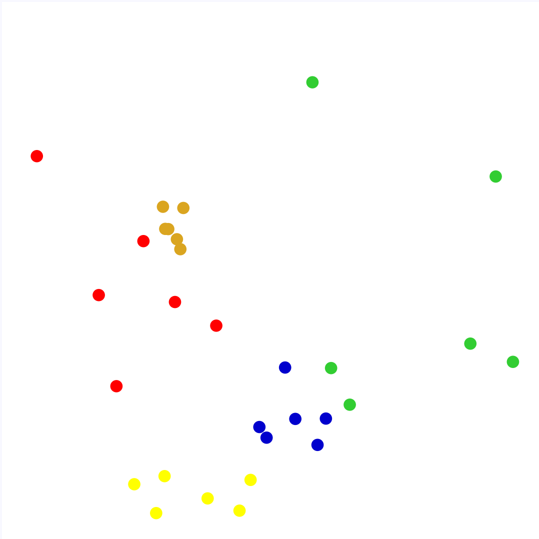
- Closeness of agreement between a measured quantity value and a true quantity value of a measurand

How to evaluate if true quantity is unknown?

MEASUREMENT PRECISION (VIM 2.15)

- Closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions

Accuracy & Precision — Where's the Target?



Measurement Uncertainty — Evaluation

BOTTOM-UP / TOP-DOWN

- Bottom-up assessments — uncertainty budgets for individual labs or measurement methods
- Top-down assessments — via interlaboratory and multiple method studies
  - *Often reveal unsuspected uncertainty components*

Dark uncertainty
— Thompson & Ellison (2011)

Types of Measurement Uncertainty Evaluations

TYPE A

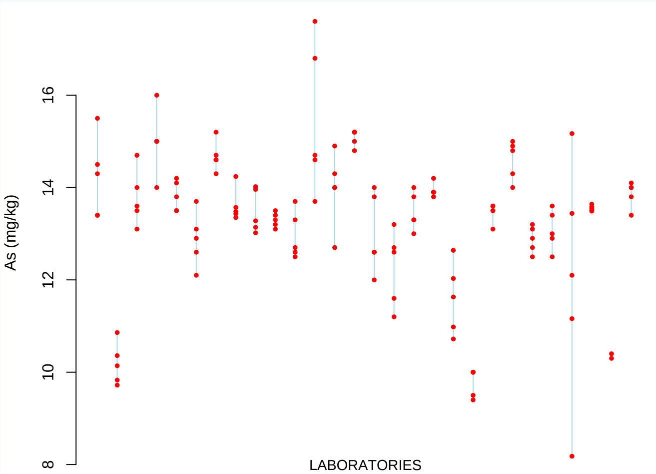
- Based on statistical scatter of measured values obtained under comparable measurement conditions (*repeatability*, *reproducibility*)

TYPE B

- Based on other evidence, including information
  - Published in compilations of quantity values
  - Obtained from a calibration certificate, or associated with certified reference material
  - Obtained from experts

Uncertainty Components — Interlab Study

- Lab-specific
- Dispersion of measured values from different labs



Outline

Measurement Models & Uncertainty Evaluations

ORIENTATION

MEASUREMENT MODELS

- Measurement equations
- Observation equations

May have to use both in the course of an uncertainty evaluation

UNCERTAINTY EVALUATIONS

- Measurement equations: Gauss's formula or Monte Carlo Method — NIST Uncertainty Machine
- Observation equations: Statistical methods

Measurement Models

MEASUREMENT EQUATION

DEFINITION

- Measurand (*output quantity*) is known function of input quantities
- Estimates of the input quantities and characterizations of associated uncertainties are available

EXAMPLE

- Thermal expansion coefficient

$$\alpha = \frac{L_1 - L_0}{L_0(T_1 - T_0)}$$

Measurement Models

OBSERVATION EQUATION — DEFINITION & EXAMPLE (TEMPERATURE)

DEFINITION

- Measurand is function of parameters of statistical model for experimental data

EXAMPLE

- Replicated measurements of temperature τ: $t_1 = 99.49\,°C$, $t_2 = 100.42\,°C$, $t_3 = 99.56\,°C$, $t_4 = 100.95\,°C$
- Measurement model (*observation equation*) $t_i = \tau + \epsilon_i$, for $i = 1, \dots, 4$

  $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ non-observable measurement errors

Measurement Models

OBSERVATION EQUATION — EXAMPLE (WEIBULL LIFETIME)

- Lifetime W of mechanical part has Weibull probability distribution with shape α and scale β

$$\log W = \log \beta + \tfrac{1}{\alpha} \log(-\log U)$$

Non-observable "error" U has rectangular distribution on $(0, 1)$

- Measurand is expected lifetime $\eta = \beta\Gamma(1 + \frac{1}{\alpha})$

Model Selection

MEASUREMENT EQUATION *vs.* OBSERVATION EQUATION

MEASUREMENT EQUATION

- Measurand is known function of input quantities, and these do not depend on value of measurand
- Estimates of input quantities, and characterizations of associated uncertainties, are available

OBSERVATION EQUATION

- Observable quantities (*experimental data*) depend on value of measurand but relationship between them is stochastic, not deterministic
- Replicated values of same observable quantity typically vary (i.e., they are *mutually inconsistent*) even though measurand remains invariant

Measurement Uncertainty — Evaluation

MEASUREMENT EQUATION & GAUSS'S FORMULA

$$\boxed{Y = f(X_1, \dots, X_n)}$$

GAUSS'S (1823) FORMULA — GUM (10), (13)

$$u^2(y) \approx \sum_{j=1}^{n} c_j^2 u^2(x_j) + 2 \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} c_j c_k u(x_j) u(x_k) r(x_j, x_k)$$

c_j — Value at $(x_1, \dots, x_n)$ of first-order partial derivative of f with respect to jth argument

$r(x_j, x_k)$ — Correlation between X_j and X_k

Gauss's Formula

INGREDIENTS & ASSUMPTIONS

- Estimates, standard uncertainties, and correlations of input quantities
- Values of partial derivatives of f
- Measurement function f approximately linear in neighborhood of estimates of input quantities
- Uncertainty of input quantities small relative to size of that neighborhood
- Probabilistic interpretation of $y \pm ku(y)$ involves additional assumptions

Measurement Uncertainty — Evaluation

MEASUREMENT EQUATION & MONTE CARLO METHOD

$$Y = f(X_1, \dots, X_n)$$

MONTE CARLO METHOD — GUM-S1

- Apply perturbations to values of input quantities drawn from respective probability distributions
- Compute value of output quantity for each set of perturbed values of input quantities
- Values of output quantity are sample from corresponding probability distribution:
  - Their standard deviation is an evaluation of $u(y)$
  - Use them also to produce coverage intervals and to characterize Y's probability density

Monte Carlo Method

INGREDIENTS & ASSUMPTIONS

- Joint probability distribution of input quantities specified fully
- Specialized software to simulate draws from that distribution
- No need for linearizations (approximations) or for partial derivatives
- Results automatically interpretable probabilistically

Temperature Measurement

OBSERVATION EQUATION

- Replicated measurements of temperature τ: $t_1 = 99.49\,°C$, $t_2 = 100.42\,°C$, $t_3 = 99.56\,°C$, $t_4 = 100.95\,°C$
- *Observation equation* $t_i = \tau + \epsilon_i$
- If non-observable measurement errors $\{\epsilon_i\}$ have same Gaussian probability distribution with mean 0 and same standard deviation, then $\widehat{\tau} = \overline{t} = (99.49 + 100.42 + 99.56 + 100.95)/4 = 100.11\,°C$ estimates τ with minimum mean squared error

  Under different assumptions, best estimate need not be the average!

Temperature Measurement
UNCERTAINTY EVALUATION

- In the absence of any other recognized sources of uncertainty, and under same assumptions that make $\bar{t}$ best estimate:
  - $u(\bar{t}) = \text{SD}(99.49, 100.42, 99.56, 100.95)/\sqrt{4}$
    $= 0.35\,°\text{C}$
  - $100.11 \pm (4.30 \times 0.35) = (98.98\,°\text{C}, 101.23\,°\text{C})$ is approximate 95 % coverage interval for τ

    4.30 is the 97.5th percentile of Student's t distribution with 3 degrees of freedom

F-100 *Super Sabre* Lifetime
OBSERVATION EQUATION — EXAMPLE

- Measurand: component lifetime
- Times to failure (hour): 0.22, 0.50, 0.88, 1.00, 1.32, 1.33, 1.54, 1.76, 2.50, 3.00, 3+, 3+, 3+

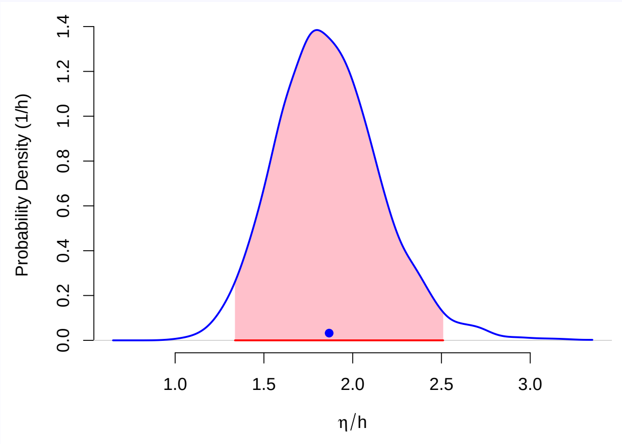
F-100 *Super Sabre* Lifetime
VALUE ASSIGNMENT, UNCERTAINTY EVALUATION

- Observation equation (*statistical model*)
  - Observed lifetimes are sample from Weibull distribution with shape α and scale β
- Value assignment
  - Maximum likelihood estimates of parameters are $\widehat{\alpha}$ and $\widehat{\beta}$ that maximize likelihood function, and $\widehat{\eta} = \widehat{\beta}\Gamma(1 + 1/\widehat{\alpha})$
- Uncertainty evaluation
  - Parametric statistical bootstrap
    - Draw samples from fitted distribution, of the same size and censored as experimental data
    - Estimate parameters from these samples and compute corresponding values of expected lifetime: their standard deviation is evaluation of $u(\eta)$

F-100 *Super Sabre* Lifetime
MAXIMUM LIKELIHOOD ESTIMATE, STATISTICAL BOOTSTRAP

- Measurement result: $\widehat{\eta} = 1.84\,\text{h}$, $u(\widehat{\eta}) = 0.30\,\text{h}$

  Lifetime shorter than 4.5 h with 99 % probability



Outline

Evaluation of Measurement Uncertainty

METHODS & TOOLS

METHODS

(1) Gauss's formula (*Measurement Equation*)

(2) Monte Carlo method (*Measurement Equation* and *Observation Equation*)

(3) Statistical method (*Observation Equation*)

TOOLS

(1) NIST Uncertainty Machine

(2) NIST Uncertainty Machine

(3) Collaboration between metrologist and statistician

NIST Uncertainty Machine

EXAMPLES

- Thermal expansion coefficient
- Falling ball viscometer
- End-Gauge calibration
- Resistance
- Stefan-Boltzmann constant
- Freezing point depression

NIST Uncertainty Machine — Example

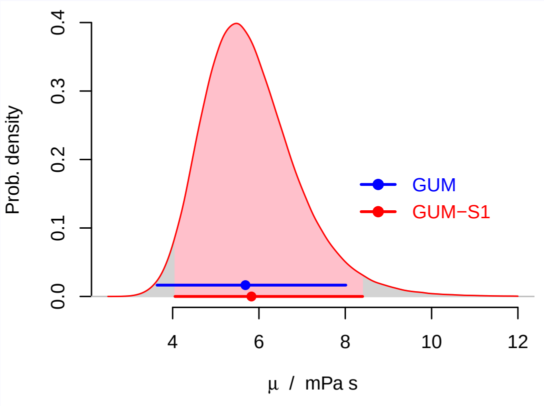
THERMAL EXPANSION COEFFICIENT

$$\alpha = \frac{L_1 - L_0}{L_0(T_1 - T_0)}$$

	x	$u(x)$	ν
T_0	288.15 K	0.02 K	3
L_0	1.4999 m	0.0001 m	3
T_1	373.10 K	0.05 K	3
L_1	1.5021 m	0.0002 m	3

Example
FALLING BALL VISCOMETER

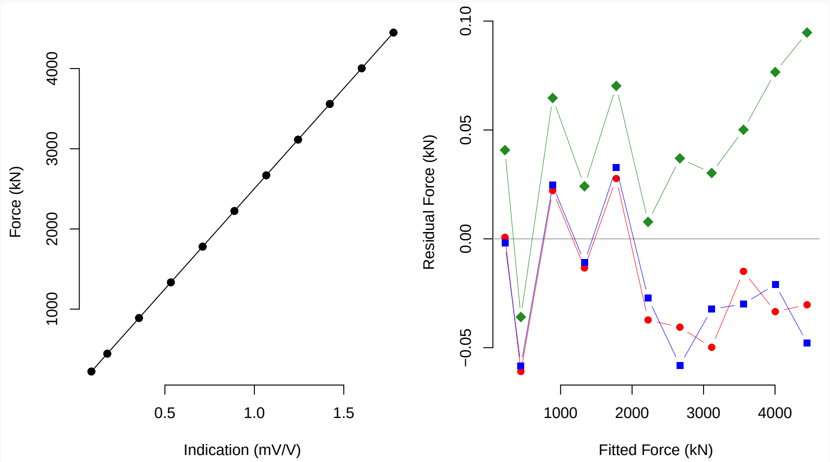
$$\mu_M = \mu_C \frac{\rho_B - \rho_M}{\rho_B - \rho_C} \frac{t_M}{t_C}$$



22% solution of sodium hydroxide in water at 20 °C
HAAKE boron silica glass ball no. 2

Outline

Load Cell Calibration
DATA & RESIDUALS



- Relative standard uncertainties for both forces and indications are 0.0005 %

Load Cell Calibration
PROBLEM & SOLUTION

- Calibration function produces values of force corresponding to response indications: $F = \varphi(R)$

PROBLEM
- Both variables are affected by uncertainty
- Dispersion of response indications exceeds uncertainty associated with each one individually

SOLUTION
Errors-in-variables regression, as in ISO 6143:2001(E) for calibration of gas mixtures

Load Cell Calibration
ERRORS-IN-VARIABLES REGRESSION

- Second-degree polynomial is an adequate model for calibration function: $\varphi(R) = \alpha + \beta R + \gamma R^2$
- Determine α, β, γ, and $\rho_1, \dots, \rho_m$ that minimize

$$\sum_{i=1}^{m} \left[\frac{(F_i - \varphi(\rho_i))^2}{u^2(F_i)} + \frac{(R_i - \rho_i)^2}{u^2(R_i) + \sigma^2} \right]$$

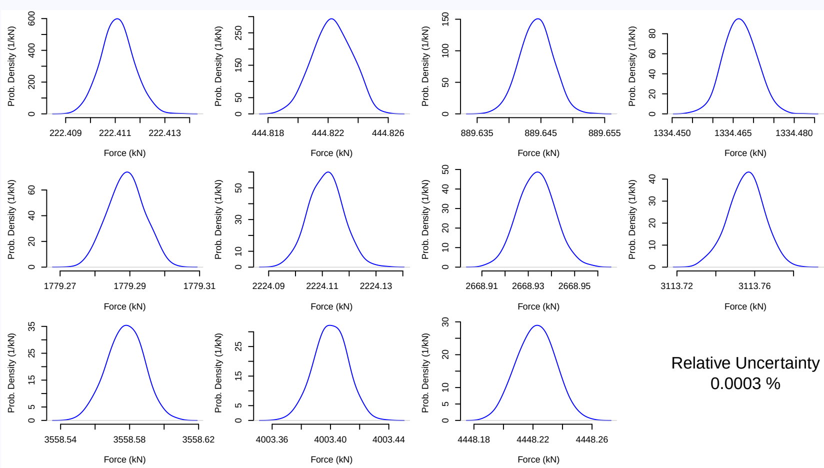
- σ describes the dispersion of replicated response indications at a fixed force setting

Load Cell Calibration
UNCERTAINTY EVALUATION — GUM SUPPLEMENT 1

- Repeat the following steps n times
  - Add Gaussian perturbations to the $\{\widehat{\rho}_i\}$ with mean 0 and variances $\{u^2(R_i) + \widehat{\sigma}^2\}$
  - Add Gaussian perturbations to the $\{\widehat{F}_i\}$ with mean 0 and variances $\{u^2(F_i)\}$
  - Fit the errors-in-variables model to the perturbed values

Load Cell Calibration
UNCERTAINTY EVALUATION — RESULTS

- Probability distributions describing uncertainties associated with forces



Uncertainty Machine — User's Manual

Thomas Lafarge Antonio Possolo

Statistical Engineering Division
Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, Maryland, USA

July 10, 2013

1 Purpose

NIST's **UncertaintyMachine** is a software application to evaluate the measurement uncertainty associated with an output quantity defined by a measurement model of the form $y = f(x_1, \dots, x_n)$, where the real-valued function f is specified fully and explicitly, and the input quantities are modeled as random variables whose joint probability distribution also is specified fully.

The **UncertaintyMachine** evaluates measurement uncertainty by application of two different methods:

- The method introduced by Gauss [1823] and described in the *Guide to the Evaluation of Uncertainty in Measurement* (GUM) [Joint Committee for Guides in Metrology, 2008a] and also by Taylor and Kuyatt [1994];
- The Monte Carlo method described by Morgan and Henrion [1992] and specified in the Supplement 1 to the GUM (GUM-S1) [Joint Committee for Guides in Metrology, 2008b].

2 Gauss's Formula *vs.* Monte Carlo Method

The method described in the GUM produces an approximation to the standard measurement uncertainty $u(y)$ of the output quantity, and it requires:

(a) Estimates $x_1, \dots, x_n$ of the input quantities;

(b) Standard measurement uncertainties $u(x_1), \dots, u(x_n)$;

(c) Correlations $\{r_{ij}\}$ between every pair of different input quantities (by default these are all assumed to be zero);

(d) Values of the partial derivatives of f evaluated at $x_1, \dots, x_n$.

When the probability distribution of the output quantity is approximately Gaussian, then the interval $y \pm 2u(y)$ may be interpreted as a coverage interval for the measurand with approximately 95 % coverage probability.

The GUM also considers the case where the distribution of the output quantity is approximately Student's t with a number of degrees of freedom that may be a function of the numbers of degrees of freedom that the $\{u(x_j)\}$ are based on, computed using the Welch-Satterthwaite formula [Satterthwaite, 1946, Welch, 1947].

In general, neither the Gaussian nor the Student's t distributions need model the dispersion of values of the output quantity accurately, even when all the input quantities are modeled as Gaussian random variables.

The GUM suggests that the Central Limit Theorem (CLT) lends support to the Gaussian approximation for the distribution of the output quantity. However, without a detailed examination of the measurement function f, and of the probability distribution of the input quantities (examinations that the GUM does not explain how to do), it is impossible to guarantee the adequacy of the Gaussian or Student's t approximations.

NOTE. The CLT states that, under some conditions, a sum of independent random variables has a probability distribution that is approximately Gaussian [Billingsley, 1979, Theorem 27.2]. The CLT is a *limit* theorem, in the sense that it concerns an infinite sequence of sums, and provides no indication about how close to Gaussian the distribution of a sum of a finite number of summands will be. Other results in probability theory provide such indications, but they involve more than just the means and variances that are required to apply Gauss's formula.

The Monte Carlo method provides an arbitrarily large sample from the probability distribution of the output quantity, and it requires that the joint probability distribution of the random variables modeling the input quantities be specified fully.

This sample alone suffices to compute the standard uncertainty associated with the output quantity, and to compute and to interpret coverage intervals probabilistically.

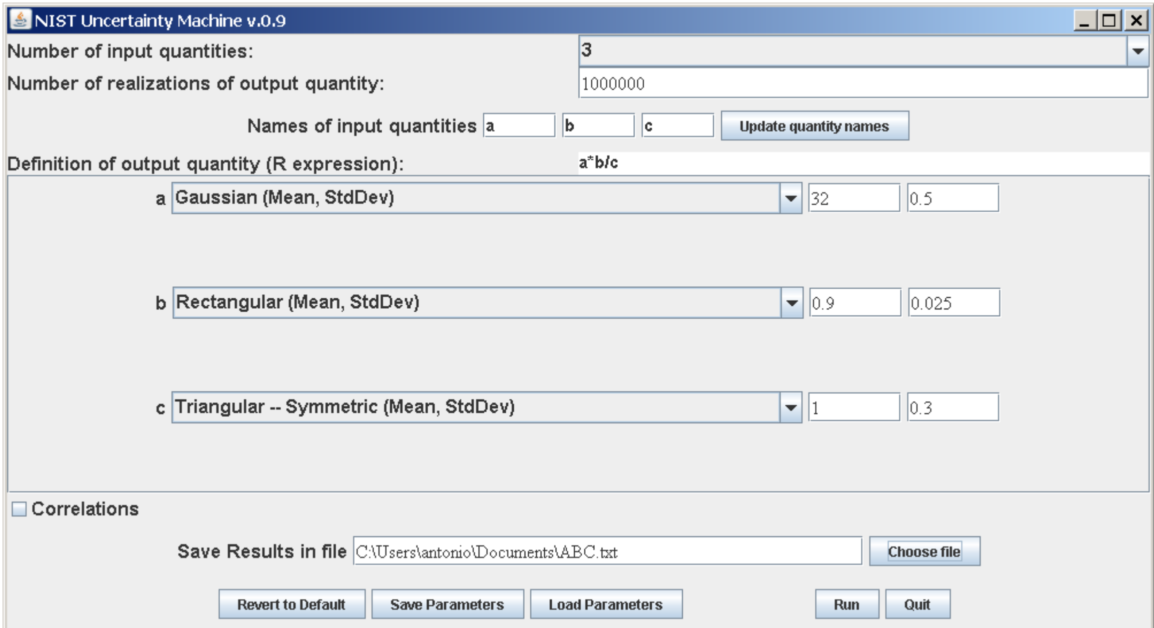
EXAMPLE. Suppose that the measurement model is $y = ab/c$, and that a, b, and c are modeled as independent random variables such that:

- a is Gaussian with mean 32 and standard deviation 0.5;
- b has a uniform (or, rectangular) distribution with mean 0.9 and standard deviation 0.025;
- c has a symmetrical triangular distribution with mean 1 and standard deviation 0.3.

Figure 1 on Page 3 shows the graphical user interface of the **`UncertaintyMachine`** filled in to reflect these modeling choices, and the results that are printed on-screen. Figure 2 on Page 4 shows a probability density estimate of the distribution of the output quantity.

The method described in the GUM produces $y = 28.8$ and $u(y) = 8.7$. According to the conventional interpretation, the interval $(11.4, 46.2)$ may be a coverage interval with approximately 95 % coverage probability.

A sample of size 1×10^7 produced by application of the Monte Carlo method has average 32.20 and standard deviation 12.53. Since only 88 % of the sample values lie within $(11.4, 46.2)$, the coverage probability of this coverage interval is much lower than the conventional interpretation would have led one to believe.




```
Monte Carlo Method

Summary statistics for sample of size 1e+06

ave     = 32.2
sd      = 12.5
median  = 28.8
mad     = 8.9

Coverage intervals

99% ( 17.12 ,  85 )     k = 2.7
95% ( 18 ,     67 )     k = 2
90% ( 19.11 ,  58 )     k = 1.6
68% ( 21.8 ,   42 )     k = 0.81
50% ( 23.7 ,   37 )     k = 0.53

--------------------------------------------

Gauss's Formula (GUM's Linear Approximation)

      y    = 28.8
      u(y) = 8.69

             SensitivityCoeffs Percent.u2
a                          0.9       0.27
b                         32.0       0.85
c                        -29.0      99.00
Correlations                NA       0.00
============================================
```

Figure 1: **ABC.** Entries in the GUI correspond to the example discussed in §2. In each numerical result, only the digits that the **UncertaintyMachine** deems to be significant are printed.

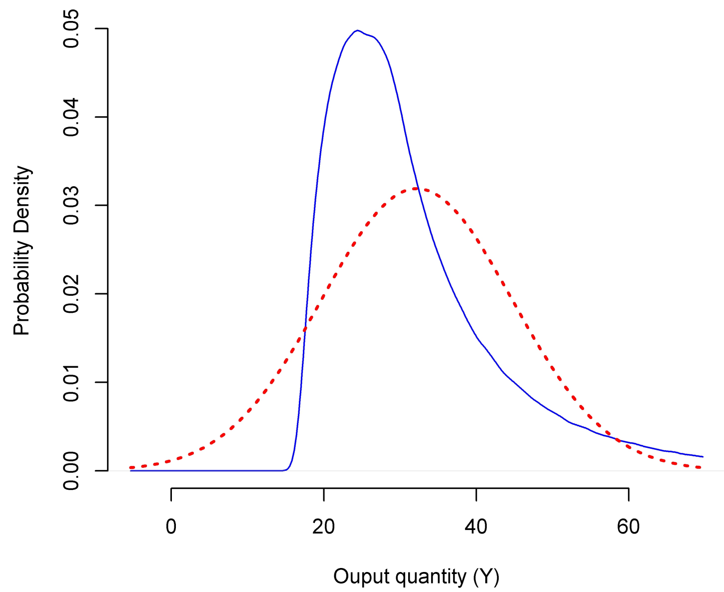


Figure 2: **ABC — Densities.** Estimate of the probability density of the output quantity (solid blue line), and probability density (dotted red line) of a Gaussian distribution with the same mean and standard deviation as the output quantity, corresponding to the results listed in Figure 1 on Page 3. In this case, the Gaussian approximation is very inaccurate.

3 Software

NIST's **`UncertaintyMachine`** should run on any computer where Oracle's Java (`www.java.com`) is installed, irrespective of the operating system. Since the computations are done using facilities of the R environment for statistical computing and graphics [R Development Core Team, 2012], this too, must be installed. The software is installed as described in §10.

> NOTE. Some commercial products, including software, are identified in this manual in order to specify the means whereby the **`UncertaintyMachine`** may be employed. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the software identified is necessarily the only or best available for the purpose.

4 Usage

The following instructions are for using the **`UncertaintyMachine`** under the Microsoft Windows operating system: under other operating systems, the steps are similar.

(U-1) Either by double-clicking the **`UncertaintyMachine`** icon where it will have been installed, or by selecting the appropriate item in the `Start` menu, launch the application's graphical user interface (GUI), which is displayed in a resizable window.

(U-2) If one wishes to use a saved configuration, click **`Load Parameters`**, select the file where parameters will have been saved previously, and continue.

(U-3) Choose the number of input quantities from the drop-down menu corresponding to the entry **`Number of input quantities`**. In response to this, the GUI will update itself and show as many boxes as there are input quantities, and assign default names to them (which may be changed as explained below).

(U-4) Enter the size of the sample to be drawn from the probability distribution of the output quantity, into the box labeled **`Number of realizations of output quantity`**: the default value, 1×10^6, is the minimum recommended sample size.

(U-5) Enter the names of the input quantities into the boxes following **`Names of input quantities`**.

(U-6) Click the button labeled **`Update quantity names`**: this will update the labels of the boxes that appear farther down in the GUI that are used to assign probability distributions to the input quantities.

(U-7) Enter a valid R expression into the box labeled **`Value of output quantity (R expression)`** that defines the value of the output quantity. This expression should involve only the input quantities, and functions and numerical constants that R knows how to evaluate. (Remember that R is case sensitive.)

Alternatively, the definition may comprise several R expressions, possibly in different lines within this box (pressing `Enter` on the keyboard, with the cursor in this box, creates a new line), but the last expression must evaluate the output quantity (without assigning this value to any variable).

> EXAMPLE. If the measurement model is $A = (L_1 - L_0)/\big(L_0(T_1 - T_0)\big)$, then the R expression that should then be entered into this box is `(L1-L0)/(L0*(T1-T0))`.
> Alternatively, the box may comprise these three lines:
> ```
> N = L1-L0
> D = L0*(T1-T0)
> N / D
> ```
> Note that the last expression produces the value $f(x_1, \ldots, x_n)$ that the measurement function takes at the estimates of the input quantities.

(U-8) Assign a probability distribution to each of the input quantities, using the drop-down menus in front of them. Once a choice is made, one or more additional input boxes will appear, where values of parameters must be entered fully to specify the probability distribution that was selected. Table 1 on Page 7 lists the distributions implemented currently, and their parametrizations. Note that some distributions have more than one parametrization: in such cases, only one of the parametrizations needs to be specified.

(U-9) If there are correlations between input quantities that need to be taken into account, then check the box marked **`Correlations`**, and enter the values of non-zero correlations into the appropriate boxes in the upper triangle of the correlation matrix that the GUI will display.

(U-10) If the box marked **`Correlations`** has been checked, then besides having specified correlations in (U-9), also select a copula (currently, either Gaussian or Student's t) to manufacture a joint probability distribution for the input quantities. If the copula chosen is (multivariate) Student's t, then another box will appear nearby to receive the number of degrees of freedom.

NOTE. The resulting joint distribution reproduces the correlation structure that has been specified, and has the distributions specified for the input quantities as margins. Possolo [2010] explains and illustrates the role that copulas play in uncertainty analysis.

(U-11) Optionally, if you wish to save the results of the calculations (numerical and graphical), then click the button labeled **`Choose file`**, and use the file selection dialog that is displayed to select the location, and the prefix for the output files.

The prefix will be used to define the names of the three output files that will be created: (i) a plain ASCII text file where the sample of values of the output quantity will be written to, one per line; (ii) a JPEG file with a plot; and (iii) a plain ASCII text file with summary statistics of the Monte Carlo sample drawn from the distribution of the output quantity, and with the estimate of the measurand and the standard uncertainty evaluated as specified in the GUM.

EXAMPLE. If the specified prefix is `ABC.txt or ABC`, then the three output files that will be created and named automatically will be called `ABC-values.txt`, `ABC-density.jpg` and `ABC-results.txt`.

(U-12) Optionally, save the parameters specified in the GUI by clicking **`Save Parameters`** and choosing a file to save the GUI's current configuration to. This configuration comprises the definition of the measurement model and the parameter settings.

(U-13) Click the button labeled **`Run`**. In response to this, a window will open where numerical results will be printed, and a graphics window will also open to display graphical results. If a file name will have been specified in (U-11), then both numerical and graphical output are saved to files.

The **`UncertaintyMachine`** estimates the number of significant digits in the results, and reports only these. To increase the number of significant digits, another run will have to be done with a larger sample size than what was specified in (U-4).

(U-14) To quit, press the button labeled **`Quit`** on the GUI, and close the window that the **`UncertaintyMachine`** created in (U-13), which will induce the graphics window also to close.

5 Results

The **`UncertaintyMachine`** produces output in two windows on the screen, and optionally writes three files to disk, described next.

- One of the outputs shown on the computer screen is an R graphics window that shows a kernel estimate [Silverman, 1986] of the probability density of the output quantity (drawn in a solid blue line), and the probability density of the Gaussian distribution with the same mean and standard deviation as the Monte Carlo sample of values of the output quantity (drawn as a red dotted line).

NAME	PARAMETERS	CONSTRAINTS
Bernoulli	Prob. of success	$0 <$ Prob. of success < 1
Beta	Mean, StdDev	$0 <$ Mean < 1, $0 <$ StdDev $< ½$
	Shape1, Shape2	Shape1 > 0, Shape2 > 0
Chi-Squared	DF	DF > 0
Exponential	Mean	Mean > 0
Gamma	Mean, StdDev	Mean > 0, StdDev > 0
	Shape, Scale	Shape > 0, Scale > 0
Gaussian	Mean, StdDev	StdDev > 0
Gaussian – Truncated	Mean, StdDev, Left, Right	StdDev > 0, Left $<$ Right
Rectangular	Mean, StdDev	StdDev > 0
	Left, Right	Left $<$ Right
Student's t	Mean, StdDev, DF	StdDev > 0, DF > 2
	Center, Scale, DF	Scale > 0, DF > 0
Triangular – Symmetric	Mean, StdDev	StdDev > 0
	Left, Right	Left $<$ Right
Triangular – Asymmetric	Left, Right, Mode	Left $\leqslant$ Mode $\leqslant$ Right; Left $\neq$ Right
Uniform	Mean, StdDev	StdDev > 0
	Left, Right	Left $<$ Right
Weibull	Mean, StdDev	Mean > 0, StdDev > 0
	Shape, Scale	Shape > 0, Scale > 0

Table 1: **Distributions.** Several distributions are available with alternative parametrizations: for these, it suffices to select and specify one of them. The rectangular distribution is the same as the uniform distribution. DF stands for number of degrees of freedom. Left and Right denote the left and right endpoints of the interval to which a distribution assigns probability 1. For the truncated Gaussian distribution, Mean and StdDev denote the mean and standard deviation without truncation: the actual mean and standard deviation depend also on the truncation points, and it is the actual mean and standard deviation that the GUM and Monte Carlo methods use in their calculations. The mode of a distribution is where its probability density reaches its maximum. A Student's *t* distribution will have infinite standard deviation unless DF > 2, and its mean will be undefined unless DF > 1. The values assigned to the parameters must satisfy the constraints listed.

- Numerical output is written in another window, in the form of a table with summary statistics for the sample that was drawn from the probability distribution of the output quantity: average, standard deviation, median, MAD.

  NOTE. "MAD" denotes the median absolute deviation from the median, multiplied by a factor (1.4826) that makes the result comparable to the standard deviation when applied to samples from Gaussian distributions.

  Also listed are coverage intervals with coverage probabilities 99 %, 95 %, 90 %, 68 %, and 50 %. The interval with 68 % coverage probability is often called a "1-sigma interval", and the interval with 95 % coverage probability is often called a "2-sigma interval": however, these designations are appropriate only when the distribution of the output quantity is approximately Gaussian. Next to each interval is listed the value of the corresponding *coverage factor* k (*cf.* GUM §3.3.7, §6.2).

  Below these, and in the same window, are listed the value of the output quantity corresponding to the estimates of the input quantities, and the value of $u(y)$ computed using formula (13) in the GUM [Joint Committee for Guides in Metrology, 2008a, Page 21].

  Finally, a table shows the sensitivity coefficients that are defined in the GUM §5.1.3: the values of the partial derivatives of the measurement function f evaluated at the estimates of the input quantities.

  The same table also shows the percentage contributions that the different input quantities make to the squared standard uncertainty of the output quantity. If the input quantities are uncorrelated, then these contributions add up to 100 % approximately. If they are correlated, then the contributions may add up to more or less than 100 %: in this case, the line labeled `Correlations` will indicate the percentage of $u^2(y)$ that is attributable to those correlations (this percentage is positive if $u^2(y)$ is larger than it would have been in the absence of correlations).

- If the user has specified a file name prefix in (U-11), following **`Save results in file`** in the GUI, then the first output file name ends in `-values.txt` and is a plain ASCII text file with one value per line of the sample that was drawn from the probability distribution of the output quantity. This file may be read into R or into any other computer program for statistical analysis, to produce additional numerical and graphical summaries.

- The second output file has the same prefix as the file just mentioned, but its name ends in `-results.txt`, and contains the same summary statistics and GUM uncertainty evaluation that were already displayed on the screen.

- The third output file has the same prefix as the file just mentioned, but its name ends in `-density.jpg`, and it is a JPEG file with the same graphical output that was displayed in the graphics window on the screen.

6 Example — Thermal Expansion Coefficient

To measure the coefficient of linear thermal expansion of a cylindrical copper bar, the length $L_0 = 1.4999$ m of the bar was measured with the bar at temperature $T_0 = 288.15$ K, and then again at temperature $T_1 = 373.10\,\text{K}$, yielding $L_1 = 1.5021\,\text{m}$. The measurement model is $A = (L_1 - L_0)/\big(L_0(T_1 - T_0)\big)$ (this "A" denotes uppercase Greek alpha).

For the purpose of this illustration we will assume that the input quantities are like (scaled and shifted) Student's t random variables with 3 degrees of freedom, with means equal to the measured values given, and standard deviations $u(L_0) = 0.0001\,\text{m}$, $u(L_1) = 0.0002\,\text{m}$, $u(T_0) = 0.02\,\text{K}$, and $u(T_1) = 0.05\,\text{K}$.

> NOTE. The assignment of distributions to the four input quantities would be appropriate if their estimates were averages of four replicated readings each, and these were outcomes of independent Gaussian random variables with unknown common mean and standard deviation.

The GUM's approach yields $\alpha = 1.727 \times 10^{-5}\,\text{K}^{-1}$ and $u(\alpha) = 1.8 \times 10^{-6}\,\text{K}^{-1}$, and the Monte Carlo method reproduces these results. Figure 3 on Page 10 reflects these facts, and lists the numerical results. The graphical results are displayed in Figure 4 on Page 11.

7 Example — End-Gauge Calibration

In Example H.1 of the GUM (which is reconsidered by Guthrie et al. [2009]), the measurement model is $l = l_S + d - l_S(\delta\alpha \cdot \theta + \alpha_S \cdot \delta\theta)$. The estimates and standard measurement uncertainties of the input quantities are listed in Table 2. For the Monte Carlo method, we model the input quantities as independent Gaussian random variables with means and standard deviations equal to these estimates and standard measurement uncertainties.

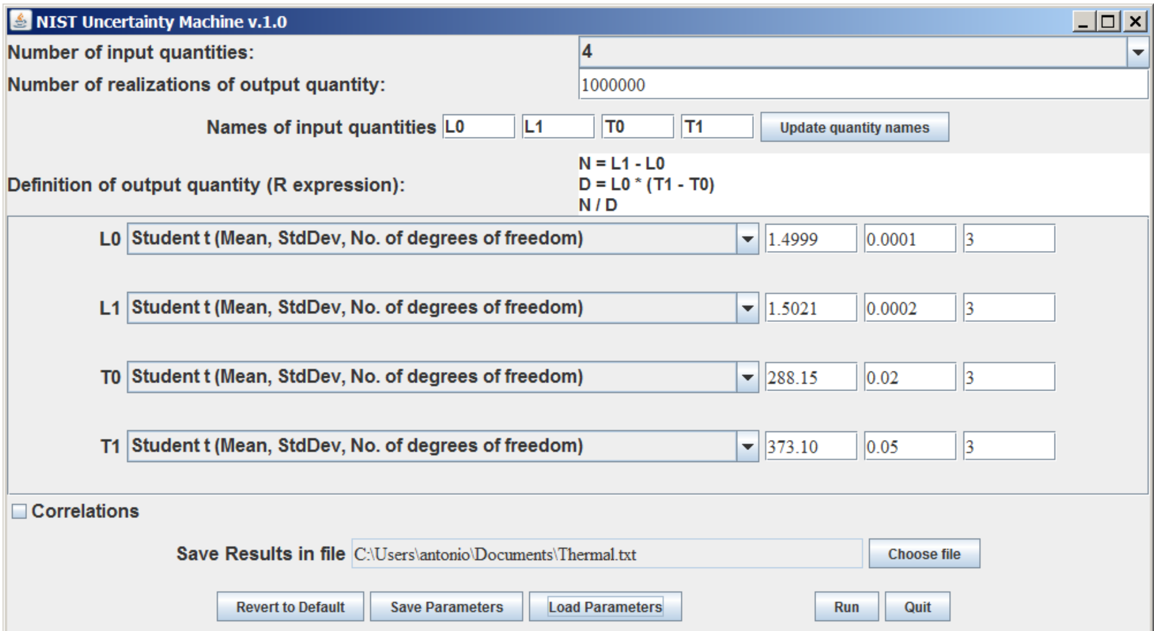
QUANTITY	x	$u(x)$
l_S	50 000 623 nm	25 nm
d	215 nm	9.7 nm
$\delta\alpha$	0 °C^{-1}	0.58×10^{-6} °C^{-1}
θ	−0.1 °C	0.41 °C
α_S	11.5×10^{-6} °C^{-1}	1.2×10^{-6} °C^{-1}
$\delta\theta$	0 °C	0.029 °C

Table 2: **End-Gauge Calibration.** Estimates and standard measurement uncertainties for the input quantities in the measurement model of Example H.1 in the GUM.

The GUM's approach yields $l = 50\,000\,838\,\text{nm}$ and $u(l) = 32\,\text{nm}$, while the Monte Carlo method reproduces the value for l but evaluates $u(l) = 34\,\text{nm}$.

The GUM (Page 84) gives (50 000 745 nm, 50 000 931 nm) as an approximate 99 % coverage interval for l, and the results of the Monte Carlo method confirm this coverage probability. If one chooses a coverage interval that is probabilistically symmetric (meaning that it leaves 0.5 % of the Monte Carlo sample uncovered on both sides), then the Monte Carlo method produces (50 000 749 nm, 50 000 927 nm) as 99 % coverage interval (and this is not quite centered at the estimate of y).

Figure 5 on Page12 reflects these facts and lists the numerical results. The graphical results are displayed in Figure 6 on Page 13.




```
Monte Carlo Method

Summary statistics for sample of size 1e+06

ave     = 1.727e-05
sd      = 1.7e-06
median  = 1.727e-05
mad     = 1e-06

Coverage intervals

99% ( 1.15e-05 ,  2.3e-05 )     k = 3.3
95% ( 1.4e-05 ,   2.05e-05 )    k = 1.9
90% ( 1.48e-05 ,  1.973e-05 )   k = 1.4
68% ( 1.6e-05 ,   1.857e-05 )   k = 0.74
50% ( 1.643e-05 , 1.811e-05 )   k = 0.48

--------------------------------------------

Gauss's Formula (GUM's Linear Approximation)

      y    = 1.727e-05
      u(y) = 1.8e-06

             SensitivityCoeffs Percent.u2
L0                    -7.9e-03    2.0e+01
L1                     7.8e-03    8.0e+01
T0                     2.0e-07    5.4e-04
T1                    -2.0e-07    3.4e-03
Correlations                NA    0.0e+00
## ===============================================
```

Figure 3: **Thermal Expansion Coefficient.** Entries in the GUI correspond to the example discussed in §6. In each numerical result, only the digits that the **`UncertaintyMachine`** deems to be significant are printed.

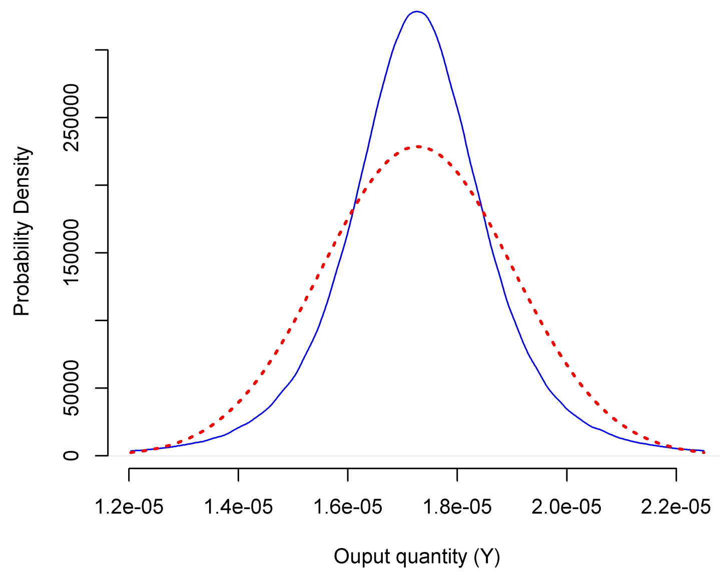


Figure 4: **Thermal Expansion Coefficient — Densities.** Estimate of the probability density of the output quantity (solid blue line), and probability density (dotted red line) of a Gaussian distribution with the same mean and standard deviation as the output quantity, corresponding to the results listed in Figure 3 on Page 10.

8 Example — Resistance

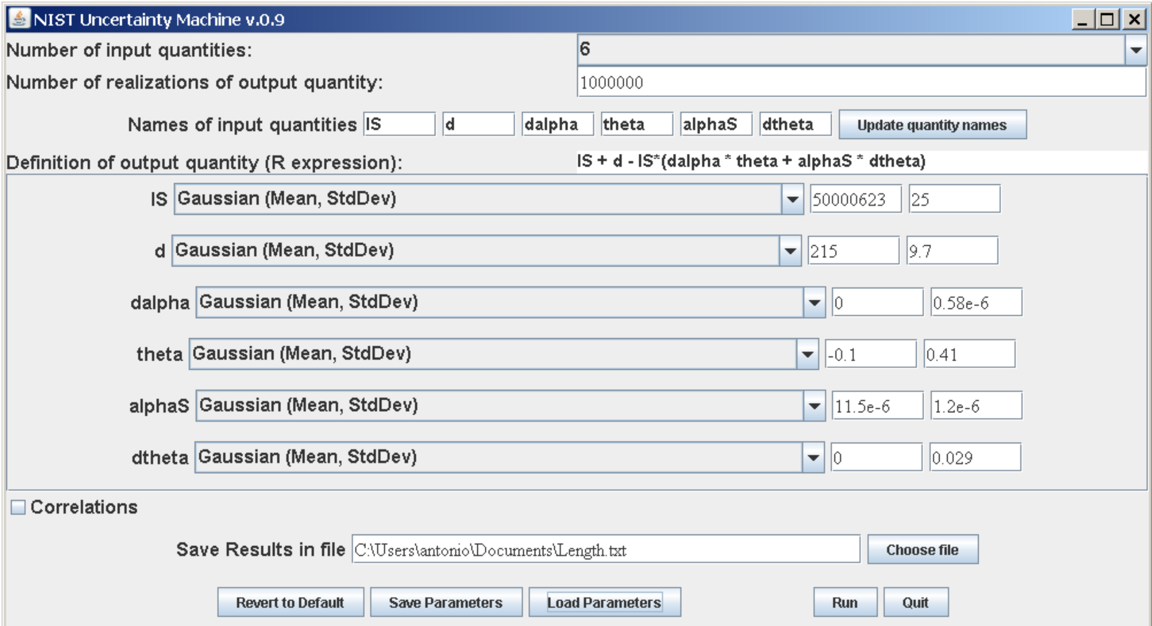
In Example H.2 of the GUM, the measurement model for the resistance of an element of an electrical circuit is $R = (V/I)\cos(\phi)$. The estimates and standard uncertainties of the input quantities, and the correlations between them, are listed in Table 3 on Page 11.

For the Monte Carlo method, we model the input quantities as correlated Gaussian random variables with means and standard deviations equal to the estimates and standard uncertainties listed in Table 3, and with correlations identical to those given in the same table. We also adopt a Gaussian copula to manufacture a joint probability distribution consistent with the assumptions already listed.

QUANTITY	x	$u(x)$
V	4.9990 V	0.0032 V
I	19.6610×10^{-3} A	0.0095×10^{-3} A
ϕ	1.044 46 rad	0.000 75 rad
$r(V,I) = -0.36$	$r(V,\phi) = 0.86$	$r(I,\phi) = -0.65$

Table 3: **Resistance.** Estimates and standard measurement uncertainties for the input quantities in the measurement model of Example H.2 in the GUM, and correlations between them, all as listed in Table H.2 of the GUM.

The GUM's approach and the Monte Carlo method produce the same values of the output quantity $R = 127.732\,\Omega$ and of the standard uncertainty $u(R) = 0.07\,\Omega$. The Monte Carlo method yields $(127.595\,\Omega, 127.869\,\Omega)$ as approximate 95 % coverage interval for the resistance without invoking any additional assumptions about R. Figure 7 on Page 14 reflects




```
Monte Carlo Method

Summary statistics for sample of size 1e+06

ave     = 50000838
sd      = 33.9
median  = 50000838
mad     = 34

Coverage intervals

99% ( 50000750 ,  50000927 )     k = 2.6
95% ( 50000770 ,  50000900 )     k = 1.9
90% ( 50000780 ,  50000894 )     k = 1.7
68% ( 50000804 ,  50000872 )     k = 1
50% ( 50000815 ,  50000861 )     k = 0.68

------------------------------------------

Gauss's Formula (GUM's Linear Approximation)

      y    = 50000838
      u(y) = 31.7

             SensitivityCoeffs Percent.u2
lS                           1      62.00
d                            1       9.40
dalpha                 5000000       0.84
theta                        0       0.00
alphaS                       0       0.00
dtheta                    -580      28.00
Correlations                NA       0.00
==========================================
```

Figure 5: **End-Gauge Calibration.**

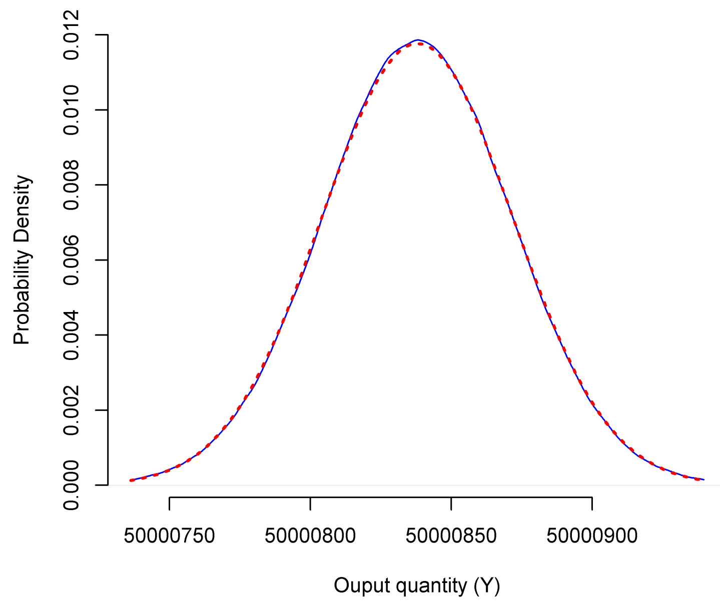


Figure 6: **End-Gauge Calibration — Densities.** Estimate of the probability density of the output quantity (solid blue line), and probability density (dotted red line) of a Gaussian distribution with the same mean and standard deviation as the output quantity, corresponding to the results listed in Figure 5 on Page 12.

these facts, and lists the numerical results. The graphical results are displayed in Figure 8 on Page 15.

9 Example — Stefan-Boltzmann Constant

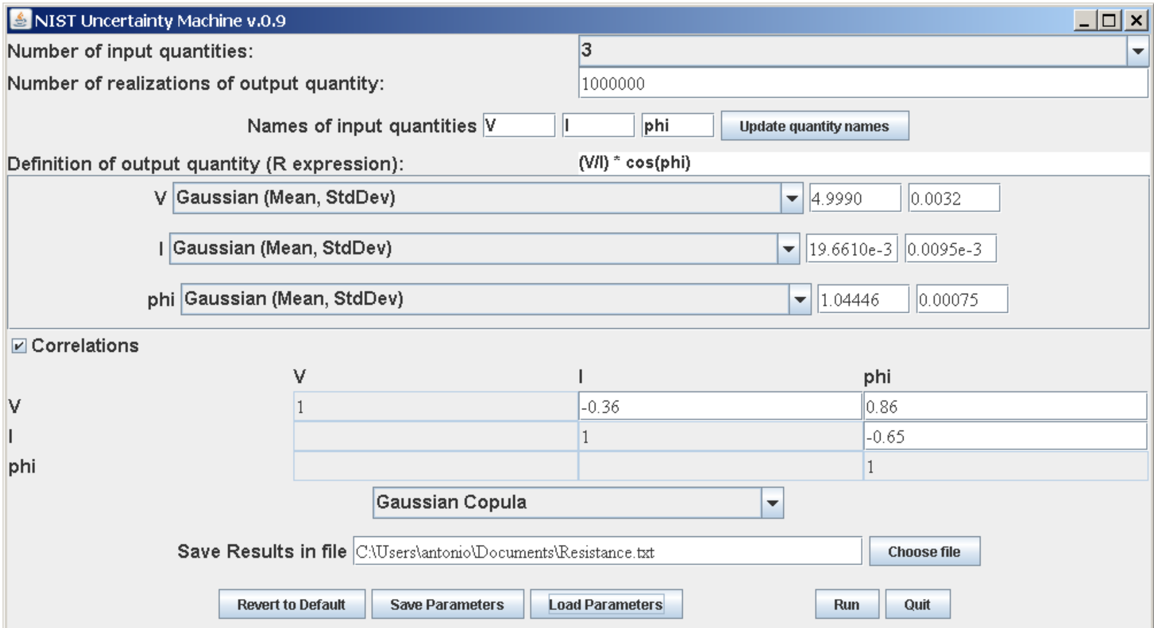
The functional relation used to define the Stefan-Boltzmann constant σ involves the Planck constant h, the molar gas constant R, Rydberg's constant R_∞, the relative atomic mass of the electron $A_r(e)$, the molar mass constant M_u, the speed of light in vacuum c, and the fine-structure constant α:

$$\sigma = \frac{32\pi^5 h R^4 R_\infty^4}{15 A_r(e)^4 M_u^4 c^6 \alpha^8}. \quad (1)$$

Table 4 lists the 2010 CODATA [Mohr et al., 2012] recommended values of the quantities that determine the value of the Stefan-Boltzmann constant, and the measurement uncertainties associated with them.

According to the GUM, the estimate of the measurand equals the value of the measurement function evaluated at the estimates of the input quantities, as $\sigma = 5.670\,37 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$. Both the GUM's approximation and the Monte Carlo method produce the same evaluation of $u(\sigma) = 2 \times 10^{-13}\,\mathrm{W\,m^{-2}\,K^{-4}}$.

These evaluations disregard the correlations between the input quantities that result from the adjustment process used by CODATA. However, once these correlations are taken into account via Equation (13) in the GUM, the same value still obtains for $u(\sigma)$ to within the single significant digit reported above.




```
Monte Carlo Method

Summary statistics for sample of size 1e+06

ave     = 127.732
sd      = 0.07
median  = 127.732
mad     = 0.07

Coverage intervals

99% ( 127.55 ,   127.912 )     k = 2.6
95% ( 127.595 ,  127.8689 )    k = 2
90% ( 127.617 ,  127.847 )     k = 1.6
68% ( 127.662 ,  127.802 )     k = 1
50% ( 127.685 ,  127.779 )     k = 0.67

------------------------------------------

Gauss's Formula (GUM's Linear Approximation)

      y    = 127.732
      u(y) = 0.07

             SensitivityCoeffs Percent.u2
V                           26        140
I                        -6500         78
phi                       -220        560
Correlations                NA       -670
==========================================
```

Figure 7: **Resistance.** Entries in the GUI correspond to the example discussed in §8. Note that, in this case, the **UncertaintyMachine** reconfigured its graphical user automatically to accommodate the correlations that had to be specified. In each numerical result, only the digits that the **UncertaintyMachine** deems to be significant are printed.

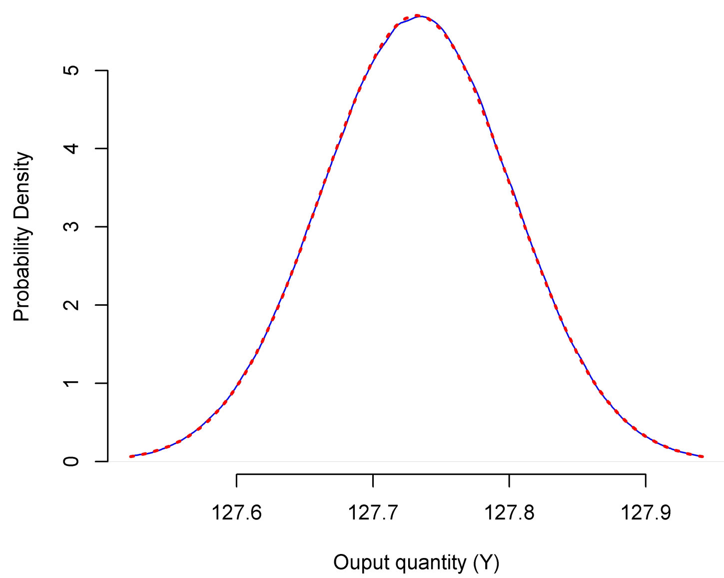


Figure 8: **Resistance — Densities.** Estimate of the probability density of the output quantity (solid blue line), and probability density (dotted red line) of a Gaussian distribution with the same mean and standard deviation as the output quantity, corresponding to the results listed in Figure 7 on Page 14.

	VALUE	STD. MEAS. UNC.	UNIT
h	$6.626\,069\,57 \times 10^{-34}$	$0.000\,000\,29 \times 10^{-34}$	J s
R	$8.314\,462\,1$	$0.000\,007\,5$	$\mathrm{J\,mol^{-1}\,K^{-1}}$
R_∞	$10\,973\,731.568\,539$	$0.000\,055$	$\mathrm{m^{-1}}$
$A_r(e)$	$5.485\,799\,094\,6 \times 10^{-4}$	$0.000\,000\,002\,2 \times 10^{-4}$	u
M_u	1×10^{-3}	0	kg/mol
c	$299\,792\,458$	0	m/s
α	$7.297\,352\,569\,8 \times 10^{-3}$	$0.000\,000\,002\,4 \times 10^{-3}$	1

Table 4: **Stefan-Boltzmann.** 2010 CODATA recommended values and standard measurement uncertainties for the quantities used to define the value of the Stefan-Boltzmann constant.

Without additional assumptions, it is impossible to interpret an expression like $\sigma \pm u(\sigma)$ probabilistically. The assumptions that are needed to apply the Monte Carlo method of the GUM Supplement 1 deliver not only an evaluation of uncertainty, but also enable a probabilistic interpretation.

If the measurement uncertainties associated with h, R, R_∞, $A_\mathrm{r}(e)$, and α are expressed by modeling these quantities as independent Gaussian random variables with means and standard deviations set equal to the values and standard measurement uncertainties listed in Table 4, then the distribution that the Monte Carlo method of the GUM Supplement 1 assigns to the measurand happens to be approximately Gaussian as gauged by the Anderson-Darling test of Gaussian shape [Anderson and Darling, 1952].

Figure 9 on Page 17 reflects these facts and lists the numerical results, which imply that the interval from $5.670\,332 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$ to $5.670\,412\,6 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$ is a coverage interval for σ with approximate 95 % coverage probability. The probability density of σ, and the corresponding Gaussian approximation are displayed in Figure 10 on Page 18.

10 Software Installation

If R has not been previously installed in the target machine, then it will have to be installed first. R is free and open-source, with versions available for all major operating systems: it may be downloaded from `www.r-project.org`. A Java Runtime Environment (JRE) is also necessary: it may be downloaded from `www.java.com`.

10.1 Microsoft Windows

Open the distribution ZIP archive, which contains the installer and the user's manual, and execute the installer. During installation, a dialog box will prompt the user to select whether (the default is to do them all):

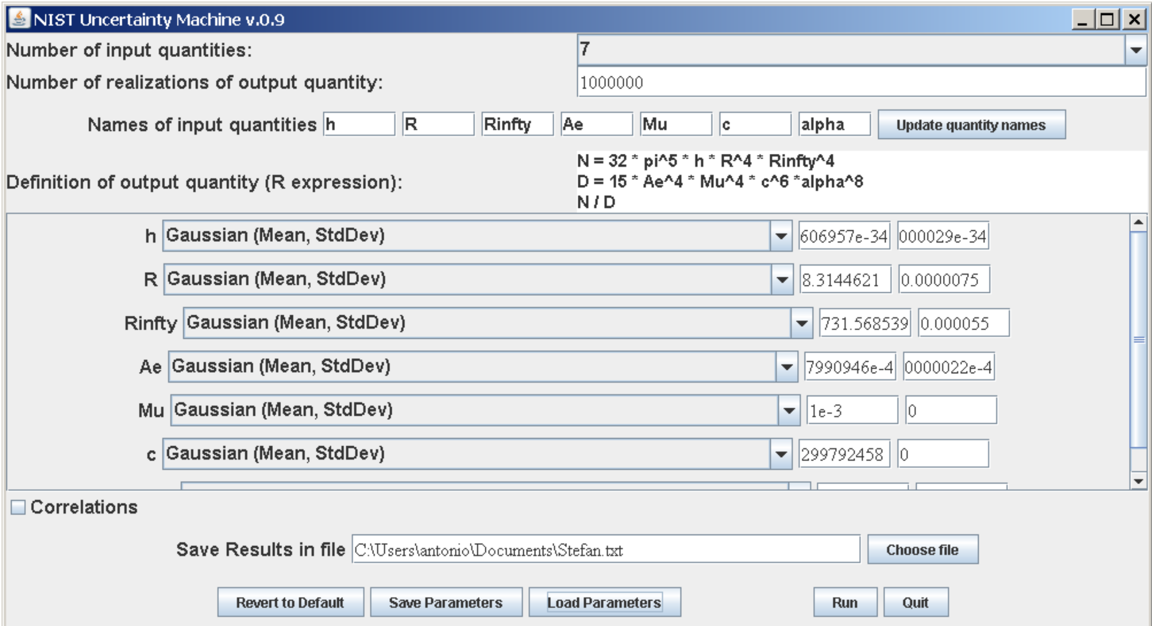
- `Rscript.exe` should be added to the search path for executables;
- Missing R packages should be installed automatically;
- A desktop icon should be created.

10.2 Linux

Assuming that R (version 2.14 or newer), a JRE, and `xterm` are installed on the system and are in the user's PATH for executables, then installation amounts to extracting the contents of the distribution archive, and placing them in the desired folder, then executing the command `Rscript configure-R.R` to install the required R packages. The software can be started by executing either the command `java -jar UncertaintyMachine.jar` or the shell script `run.sh`.

10.3 Apple OS X

The installation under Apple OS X is similar to the installation under Linux, including the requirement that `xterm` be installed: it is different from the Terminal application, and it




```
Monte Carlo Method

Summary statistics for sample of size 1e+06

ave     = 5.67037e-08
sd      = 2.05e-13
median  = 5.6703725e-08
mad     = 2.05e-13

Coverage intervals

99% ( 5.67032e-08 ,    5.670425e-08 )     k = 2.6
95% ( 5.670332e-08 ,   5.6704126e-08 )    k = 2
90% ( 5.670339e-08 ,   5.6704062e-08 )    k = 1.6
68% ( 5.6703521e-08 ,  5.670393e-08 )     k = 1
50% ( 5.670359e-08 ,   5.670386e-08 )     k = 0.66

--------------------------------------------

Gauss's Formula (GUM's Linear Approximation)

      y    = 5.67037e-08
      u(y) = 2.05e-13

             SensitivityCoeffs Percent.u2
h                      8.6e+25    1.5e-02
R                      2.7e-08    1.0e+02
Rinfty                 2.1e-14    3.1e-09
Ae                    -4.1e-04    2.0e-05
Mu                    -2.3e-04    0.0e+00
c                     -1.1e-15    0.0e+00
alpha                 -6.2e-05    5.3e-05
Correlations                NA    0.0e+00
============================================
```

Figure 9: **Stefan Boltzmann constant.**

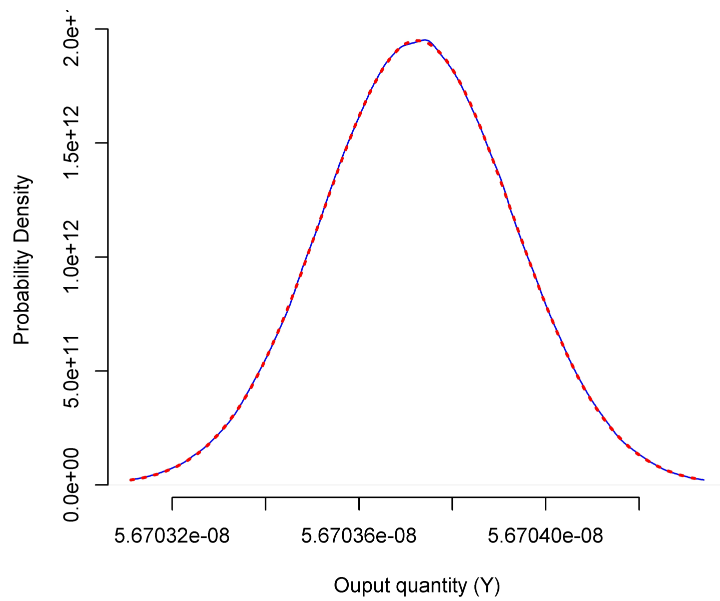


Figure 10: **Stefan-Boltzmann — Densities.** Estimate of the probability density of the output quantity (solid blue line), and probability density (dotted red line) of a Gaussian distribution with the same mean and standard deviation as the output quantity, corresponding to the results listed in Figure 9 on Page 17.

is part of Apple's X11 package. Under *Mountain Lion*, X11 installs on demand: when an application is first launched that requires X11 libraries, the user is directed to a download location for the most up-to-date version of X11 for the Mac.

Acknowledgments

The authors are grateful to their colleagues Will Guthrie, Alan Heckert, Narain Krishnamurthy, Adam Pintar, Jolene Splett, and Jack Wang, for the many suggestions that they offered for improvement of the **UncertaintyMachine** and of this user's manual.

The authors will be very grateful to users of the software, and to readers of this manual, for information about errors and other deficiencies, and for suggestions for improvement, which may be sent via eMail to `antonio.possolo@nist.gov`.

References

T. W. Anderson and D. A. Darling. Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23:193–212, 1952.

P. Billingsley. *Probability and Measure*. John Wiley & Sons, New York, NY, 1979.

C. Gauss. Theoria combinationis observationum erroribus minimis obnoxiae. In *Werke, Band IV, Wahrscheinlichkeitsrechnung und Geometrie*. Könighlichen Gesellschaft der Wissenschaften, Göttingen, 1823. `http://gdz.sub.uni-goettingen.de/`.

W. F. Guthrie, H.-k. Liu, A. L. Rukhin, B. Toman, J. C.M. Wang, and N.-f. Zhang. Three statistical paradigms for the assessment and interpretation of measurement uncertainty. In F. Pavese and A. B. Forbes, editors, *Data Modeling for Metrology and Testing in Measurement Science*, Modeling and Simulation in Science, Engineering and Technology, pages 1–45. Birkhäuser Boston, 2009. doi: 10.1007/978-0-8176-4804-6_3.

Joint Committee for Guides in Metrology. *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. International Bureau of Weights and Measures (BIPM), Sèvres, France, September 2008a. URL http://www.bipm.org/en/publications/guides/gum.html. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM 100:2008, GUM 1995 with minor corrections.

Joint Committee for Guides in Metrology. *Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method*. International Bureau of Weights and Measures (BIPM), Sèvres, France, 2008b. URL http://www.bipm.org/en/publications/guides/gum.html. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM 101:2008.

P. J. Mohr, B. N. Taylor, and D. B. Newell. CODATA recommended values of the fundamental physical constants: 2010. *Reviews of Modern Physics*, 84(4):1527–1605, October-December 2012.

M. G. Morgan and M. Henrion. *Uncertainty — A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, New York, NY, first paperback edition, 1992. 10th printing, 2007.

A. Possolo. Copulas for uncertainty analysis. *Metrologia*, 47:262–271, 2010.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL http://www.R-project.org. ISBN 3-900051-07-0.

F. E. Satterthwaite. An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 2(6):110–114, December 1946.

B. W. Silverman. *Density Estimation*. Chapman and Hall, London, 1986.

B. N. Taylor and C. E. Kuyatt. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. National Institute of Standards and Technology, Gaithersburg, MD, 1994. URL http://physics.nist.gov/Pubs/guidelines/TN1297/tn1297s.pdf. NIST Technical Note 1297.

B. L. Welch. The generalization of 'Student's' problem when several different population variances are involved. *Biometrika*, 34:28–35, January 1947.