Conference on Advanced Mathematical and Computational Tools in Metrology and Testing
23-25 June 2008 – Paris

Comparison of methods to measure uncertainty based on the ISO standard

Laura Deldossi - Diego Zappa

Department of Statistical Sciences

University of Milan Cattolica, Italy

Outline

2) ISO 5725 and GUM: a comparison

3) A conciliation between ISO 5725 and GUM:

- ✓ ISO 21749/2005: *Measurement uncertainty for metrological applications - Repeated measurements and nested experiments*
- ✓ ISO 21748/2004: *Guidance for the use of repeatibility, reproducibility and trueness estimates in measurement uncertainty estimation*

4) Conclusions, references.

In the last fifteen years the International Organization of Standardization (ISO) has produced several documents on uncertainty measurement.

The most important ones are:

- ISO 5725:1994 (Part 1-6): *Accuracy (trueness and precision) of measurement methods and results*
- ISO "*Guide to the expression of uncertainty in measurement*" (GUM) edited - in its revised form - in 1995

both proposing a method to calculate the uncertainty of a measure.

The principal characteristics of **ISO 5725** are:

1. y is directly measurable

2. it is based on a collaborative study, that is an interlaboratory experiment in which identical material is measured using the same standard method in <u>different</u> laboratory with <u>different</u> operators using <u>different</u> equipment.
   Data are obtained by an experimental design.

3. $y = \mu + B + e \quad B \sim N(0, \sigma_L^2), \quad e \sim N(0, \sigma_E^2)$

4. The uncertainty is measured as $\sigma_R = \sqrt{\sigma_L^2 + \sigma_E^2}$

5. There are no rules to obtain a confidence interval at a chosen level (usually 95%) around y (expanded uncertainties)

On the contrary the characteristics of the **GUM** method are:

1. it do not require to directly measure *y*

2. there is no experimental design to obtain Y

3. the measure y is determined from m input quantities x_1, …., x_m through a known equation (called measurement equation)

   The estimate *y* of the measurand Y is obtained as

$$y = f(x_1, x_2, \ldots, x_m)$$

4. the uncertainty is evaluated as $u_c^2(Y) = \sum_{i=1}^{m} c_i^2 u^2(X_i)$

5. it gives indication to obtain the expanded uncertainty U_p as

$$U_p = k_p u_c(Y)$$

$\mathbf{x_i}$ is an arithmetic mean of its independent repeated measurements if $\mathbf{X_i}$ is of type A

$\mathbf{x_i}$ is as the expected mean obtained from the assumed probability density function (PDF) of X_i if $\mathbf{X_i}$ is of type B

On the contrary the characteristics of the **GUM** method are:

1. it do not require to directly measure y

2. there is no experimental design to obtain Y

3. the measure y is determined from m input quantities x_1, …., x_m through a measurement equation

   The estimate y of the measurand Y is obtained as

$$y = f(x_1, x_2, \ldots, x_m)$$

4. the uncertainty is evaluated as $u_c^2(y) = \sum_{i=1}^{m} c_i^2 u^2(x_i)$

5. it gives indication to obtain the expanded uncertainty U_p as

$$U_p = k_p u_c(y)$$

COMPARISON

ISO 5725	GUM
1) y is measurable	1) y may be not directly measured
2) Experimental design	2) NO experimental design
3) $y = \mu + B + e$	3) $y = f(x_1, x_2, \ldots, x_m)$
Factors are only of type A	**Factors may be of type A and B**
4) $\sigma_R = \sqrt{\sigma_L^2 + \sigma_E^2}$	4) $u_c^2(y) = \sum_{i=1}^{m} c_i^2 u^2(x_i)$
5) NO expanded uncertainties	5) $U_p = k_p u_c(y)$

COMPARISON

ISO 5725	GUM
1) y is measurable	1) y may be not directly measured
2) Experimental design	2) NO experimental design
3) $y = \mu + B + e$	3) $y = f(x_1, x_2, \ldots, x_m)$
4) $\sigma_R = \sqrt{\sigma_L^2 + \sigma_E^2}$	4) $u_c^2(Y) = \sum_{i=1}^{m} c_i^2 u^2(X_i)$
5) NO expanded uncertainties	5) $U_p = k_p u_c(Y)$

From the beginning GUM was a **widely adopted standard approach**, probably because it was considered **a method** so **easy** that its application seemed to be achievable without the need of "at least one member of the panel with experience in the statistical design and analysis of experiments" as ISO 5725 did.

Furthermore GUM gave the possibility to take into consideration **the uncertainty of type B**, that is uncertainties specified by judgment based on available information not obtained through collaborative trials, and it can be applied even when **y is not directly measurable.**

WARNINGS on GUM - 1

In the computation of the uncertainty $u_c(Y)$ also the covariances terms should be inserted if the input variables are not independent as it is often mentioned in the GUM, but nothing is said on how to compute the covariances terms or, on the other side, how to prove that the input variables are independent.

It is evident that sampling from X_i marginally with respect to all the other variables does not allow to evaluate the dependence between the variables.

The consequence is that the uncertainty $u_c(Y)$ can be overestimated or underestimated depending on the sign of the covariances, if they are not equal to 0.

WARNINGS on GUM - 2

The choice of the coverage factor k_p on the basis of the normal or *t*-Student distribution to obtain the expanded uncertainty U_p

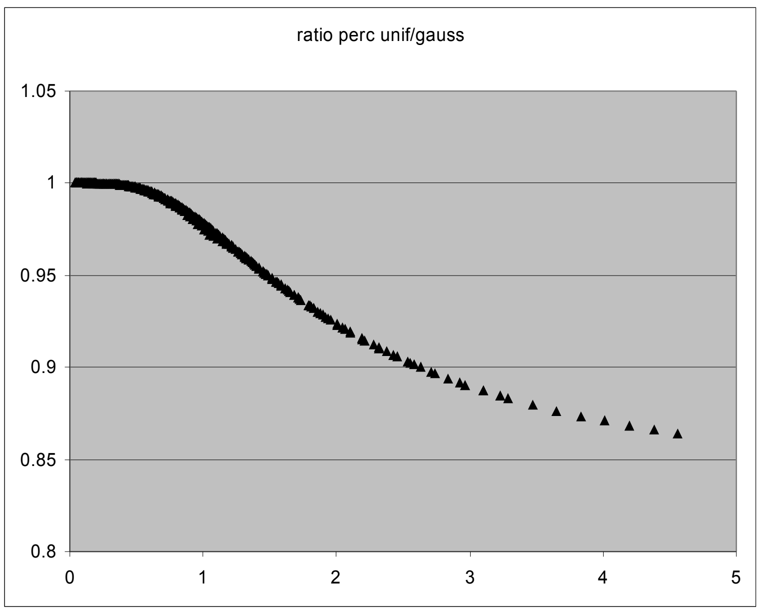
Many criticism appeared in the metrological literature and we know of the new proposal for obtaining the complete probability density function (pdf) of the measurand *y* by Monte Carlo simulation

WARNINGS on GUM – 2 follow

As the variance of the input of type B increases wrt to the correspondent of type A the coverage factor (k_p) always overestimate the true percentile of the convolution.

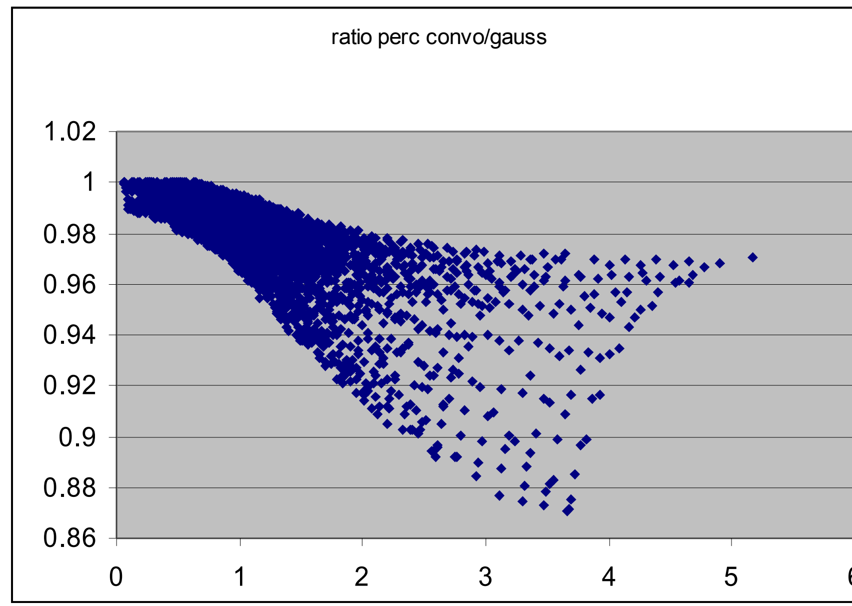
$X_1 \sim N(0,\sigma^2)$ (a type A variable) $X_2 \sim U(-a,a)$ and $X_3 \sim U(-b,b)$ (a type B variable), we computed the cumulative density function of the convolution of

X_1 and X_2



$$\sqrt{\frac{\mathrm{Var}(X_2)}{\mathrm{Var}(X_1)}}$$

X_1, X_2 and X_3



$$\sqrt{\frac{\mathrm{Var}(X_2 * X_3)}{\mathrm{Var}(X_1)}}$$

r

Ratio between the convolution percentile and the normal percentile

It follows that the confidence intervals are conservative and wider than necessary.

The major drawback is the strong increase of the II type probability error, that is, <u>it increases the probability to consider a measure within the uncertainty limits when it is not.</u>

The normal approximation may be supposed to be a good solution only if the variances due to type B variables do not dominate the variance of the type A variables (generally supposed to be gaussian)

What about ISO 5725?

y is obtained by an experimental design:

1) the randomization of the trials in the design guarantees the supposed normal distribution of the random effects;

2) the experimental design gives the possibility to estimate, if they exist, interactions among factors and to take them into account;

3) numerous statistical results are available to obtain approximated confidence intervals of variances components as

- Modified Large Sample (MLS)
- Generalized Confidence Intervals (GCI)
- Simulated Confidence Intervals (SCI)

An example

Consider an R&R study for monitoring a Scanning Electronic Microscope to evaluate line widths and hole diameters stamped over a "wafer" with a precision of a few nm.

The random sample is taken from
- a single wafer randomly chosen from a process
- 13 points (parts) on it
- 3 operators (represented by the loading/unloading of the wafer)
- 3 replications

MLS, GCI, SCI

$$Y_{ijk} = \mu + P_i + O_j + (PO)_{ij} + E_{ijk}$$

GUM

$$Y = f(P, O, E)$$

3 operators

13 Parts

	1	1	1	2	2	2	3	3	3
	r1	r2	r3	r1	r2	r3	r1	r2	r3
1	y_{111}	y_{112}	y_{113}	y_{121}	y_{122}	y_{123}	y_{131}	y_{132}	y_{133}
2	y_{211}	y_{212}	y_{213}	y_{221}	y_{222}	y_{223}	y_{231}	y_{232}	y_{233}
3	y_{311}	y_{312}	y_{313}	y_{321}	y_{322}	y_{323}	y_{331}	y_{332}	y_{333}
4	…	…	…	…	…	…	…	…	…
5	…	…	…	…	…	…	…	…	…
6	…	…	…	…	…	…	…	…	…
7	…	…	…	…	…	…	…	…	…
8	…	…	…	…	…	…	…	…	…
9	…	…	…	…	…	…	…	…	…
10	…	…	…	…	…	…	…	…	…
11	y_{1111}	y_{1112}	y_{1113}	y_{1121}	y_{1122}	y_{1123}	y_{1131}	y_{1132}	y_{1133}
12	y_{1211}	y_{1212}	y_{1213}	y_{1221}	y_{1222}	y_{1223}	y_{1231}	y_{1232}	y_{1233}
13	y_{1311}	y_{1312}	y_{1313}	y_{1321}	y_{1322}	y_{1323}	y_{1331}	y_{1332}	y_{1333}

3 replications

Expanded uncertainties

	Parts	Gauge	Total
Point Estimate	0,000040	0,000001	0,000042
GUM with normal	0,0374	0,0072	0,0122
MLS	0,0067	0,0023	0,0107
GCI	0,0067	0,0021	0,0068
SCI	0,0056	0,0008	0,0056

COMPARISON - conclusion

ISO 5725 approach is preferable to measure uncertainty if Y can be measured directly.

If it is not possible, GUM approach can be adopted but with some cautions.

Only in some particular circumstances (linearity of the function f, normality assumption, independence of the input variables) the two approaches give the same results.

A conciliation between GUM and ISO 5725

ISO 21749/2005

Measurement uncertainty for metrological applications - Repeated measurements and nested experiments

ISO 21748/2004

Guidance for the use of repeatibility, reproducibility and trueness estimates in measurement uncertainty estimation

ISO 21749

ISO 21749 follows the approach taken in GUM but it proposes to use the analysis of variance for estimating uncertainties of components classified as Type A.

The novelty of this approach consists on

1) the introduction of nested designs for estimating the uncertainties due to:

- time-dependent sources of variability
- measurement configuration (differences due to different instruments, operators, geometries)
- material inhomogeneity

2) the use of GUM approach to compute the expanded uncertainty.

First consideration

Two-Stage Nested Design

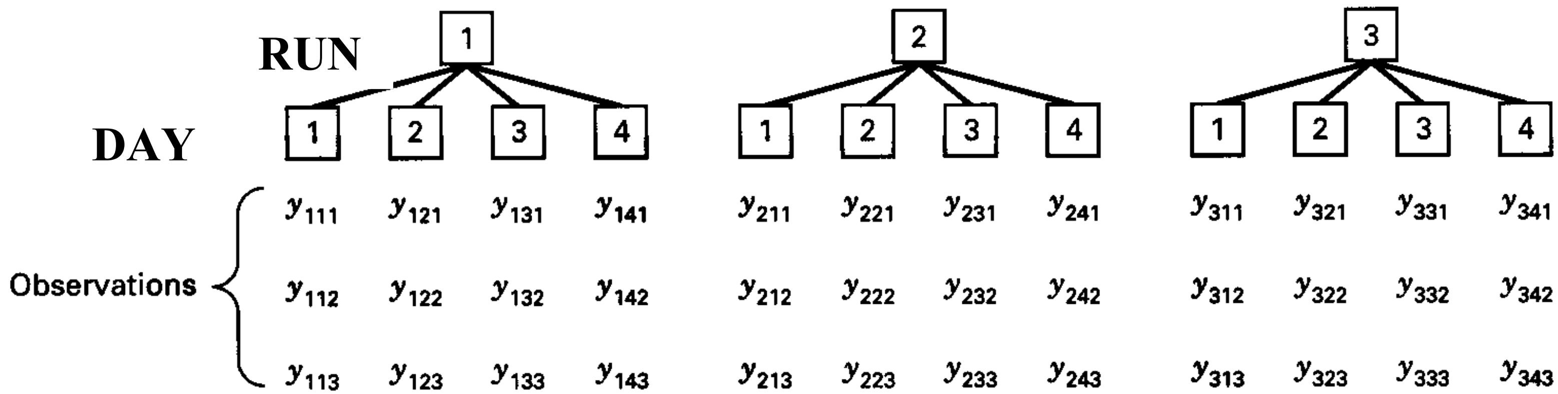


Figure 13-1 A two-stage nested design.

It is a right design for estimating the uncertainties due to time-dependent sources of variability but not always for those due to measurement configuration and material inhomogeneity

operators

	1	1	1	2	2	2	3	3	3
	r1	r2	r3	r1	r2	r3	r1	r2	r3
1	y_{111}	y_{112}	y_{113}	y_{121}	y_{122}	y_{123}	y_{131}	y_{132}	y_{133}
2	y_{211}	y_{212}	y_{213}	y_{221}	y_{222}	y_{223}	y_{231}	y_{232}	y_{233}
3	y_{311}	y_{312}	y_{313}	y_{321}	y_{322}	y_{323}	y_{331}	y_{332}	y_{333}
4	...	...	...	...	...	...	...	...	...
5	...	...	...	...	...	...	...	...	...
6	...	...	...	...	...	...	...	...	...
7	...	...	...	...	...	...	...	...	...
8	...	...	...	...	...	...	...	...	...
9	...	...	...	...	...	...	...	...	...
10	...	...	...	...	...	...	...	...	...
11	y_{1111}	y_{1112}	y_{1113}	y_{1121}	y_{1122}	y_{1123}	y_{1131}	y_{1132}	y_{1133}
12	y_{1211}	y_{1212}	y_{1213}	y_{1221}	y_{1222}	y_{1223}	y_{1231}	y_{1232}	y_{1233}
13	y_{1311}	y_{1312}	y_{1313}	y_{1321}	y_{1322}	y_{1323}	y_{1331}	y_{1332}	y_{1333}

Second consideration

Once the various variance components are estimated, the expanded uncertainty can be obtained using:

- MLS, GCI and SCI method

instead of

- GUM procedure

with the coverage factor based on the t-distribution with degrees of freedom approximated by the Welch-Satterwhaite formula.

ISO 21748

This guidance is based on the model

$$y = \mu + \delta + B + \sum_{i=1}^{m} c_i x_i + e$$

ISO5725

(s_r, s_R)

GUM

$x_i \sim N(0, u^2(x_i))$

It is based on the idea that effects not observed within the context of the collaborative study (s_R obtained by ISO 5725) shall be negligible or explicitly allowed for. In general are considered negligible those effects whose uncertainties are less than 0.2 times the correspondent expected one.

What are the main factors that are likely to change between collaborative study and routine testing ?

- sampling process
- inhomogeneity
- changes in test-item type

Their uncertainties have to be added to increment the reproducibility variance of the collaborative studies.

$s_{R'}$ instead of s_R

There would be other effects that may operate in particular instances:

- factors which are held constant
- factors which vary insufficiently during collaborative studies

The guidance proposes to estimate these effects separately, either from experimental variation or by prediction from established theory and, if not negligible, to add components in the summation.

In our opinion there is a risk in this procedure.

Including uncertainties components in subsequent steps do not allow to take into account the possible **interaction among factors**.

Then, we think, it would be better, if possible, to **improve the native experimental design** of the collaborative trial.

For example, if factors are quantitative, one can refer to the so called combined array approach in which the controllable factors as well as the noise factors, which typically arise in the real production/utilization of some goods, are unified in a single experimental design.

This approach is different from the one proposed in the ISO/TS 21748 as it allows to take into account also all the interactions between controllable and noise factors.

Conclusion

- If the response variable *Y* is measurable, the design of a suitable experiment and the use of the correspondent ANOVA table provide the best approach to estimate the different sources of variation derived from a measurement system.

1) Even if in none ISO standards the uncercentainty due to the manufacturing process is taken into account in the model as “measurement of check standards are used”, often the factor “part” should be considered together with its interaction with the other factors.

1) The use of mixed effect model is also suggested to consider fixed and random factors in a single factorial (crossed or nested) experiment for estimating all uncertainties components (systematic and random) with a single ANOVA table.

- When y is not measurable or there are input variable of type B, GUM approach can be applied, with some caution in obtaining the expanded uncertainty as it is relevant the right choice of the coverage factor and the interaction among factors.

References

. R.K. Burdick, C.M. Borror, D.C. Montgomery, Journal of Quality Technology, 35, 342-354 (2003)

. F.A. Graybill, C. Wang C., JASA 75, 372, 869-873 (1980).

. R.K. Burdick , G.A. Larsen , Journal of Quality Technology, 29, 3, 261-273 (1997).

. R.K. Burdick, C.M. Borror, D.C. Montgomery , ASA-SIAM Series on Statistics and Applied Probability (2005)

. L. Deldossi, D. Zappa, in Proceedings of ENBIS-DEINDE 2007 105-116 (2007)

. ISO, *Guide to the Expression of Uncertainty in Measurement*, (1995).

. ISO/TS 21748:2004 (2004).

. ISO/TS 21749:2005, (2005).

. ISO 5725:1994 (1-6), (1994).

. D.C. Montgomery Wiley, New-York (2001).

Thanks for your attention

Laura Deldossi, Diego Zappa