united nations educational, scientific and cultural organization
international atomic energy agency

the
abdus salam
international centre for theoretical physics

H4.SMR/1503 - 09

WORKSHOP ON NUCLEAR DATA FOR SCIENCE AND TECHNOLOGY: MATERIALS ANALYSIS

(19 - 30 May 2003)

STATISTICS

Dr. Piotr Robouch
Institute for Reference Materials and Measurements
European Commission - Joint Research Centre
Geel
BELGIUM

TrainMiC
Training in Metrology in Chemistry

STATISTICS

Statistics, why and when?

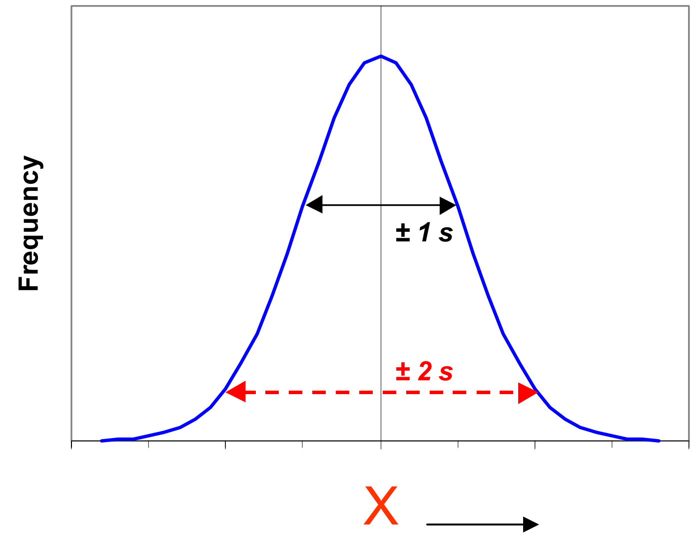
- Evaluation of uncertainty of results according to ISO-GUM
- Evaluation of Inter-Laboratory Comparison (ILC)
- Quality assurance:
  - method performance (accuracy; precision; ...)
- Optimisation of measurement procedures

TrainMiC
Training in Metrology in Chemistry

Statistics for evaluation of uncertainty

TrainMiC
Training in Metrology in Chemistry

Normal distribution



For a set of n values x_i

Mean Value (average) $$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}(x_i)$$

Standard Deviation $$s(x_i) = \sqrt{\frac{1}{n-1}\cdot\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Variance of the mean $$V(x_i) = s^2(x_i)$$

Relative Standard Deviation $$RSD = \frac{s(x_i)}{\bar{x}} \quad (absolute \quad or \quad \%)$$

Rectangular distribution

The Value is between the limits

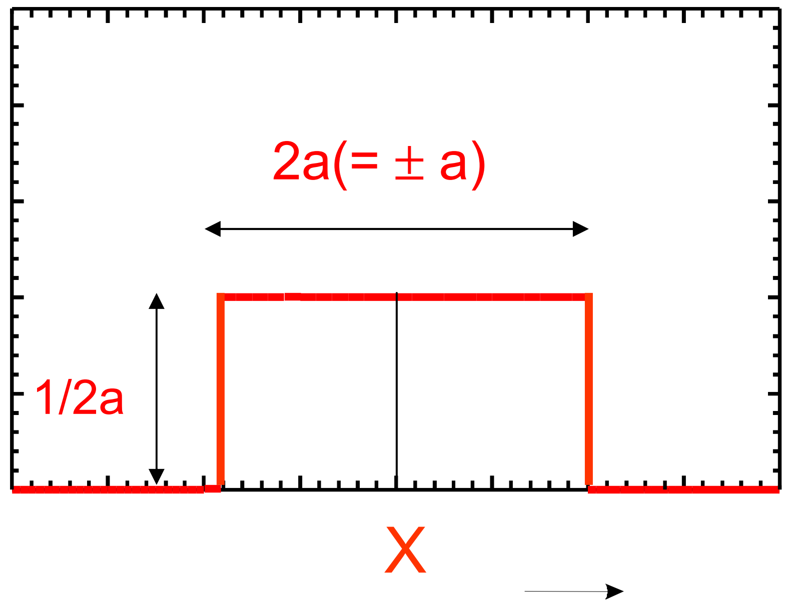
$$a_{-} \ldots a_{+}$$

The expectation

$$y = x \pm a$$

Assumed standard deviation:

$$s = a / \sqrt{3}$$



One can only assume that it is equally probable for the value to lie anywhere within the interval

TrainMiC
Training in Metrology in Chemistry

Example of Rectangular distribution

"*It is likely that the value is somewhere in that range*"

Rectangular distribution is usually described in terms of:
the average value and the range (±a)

Certificates or other specification give limits where the value could be, without specifying a level of confidence (or degree of freedom).

Examples:

Concentration of calibration standard is quoted as (1000 ± 2) mg/l
Assuming rectangular distribution the standard uncertainty is:

$$s = u(x) = a/\sqrt{3} = 2/\sqrt{3} = 1.16 \;\; mg/l$$

The purity of the cadmium is given on the certificate as (99.99 ± 0.01) %
Assuming rectangular distribution the standard uncertainty is:

$$s = u(x) = a/\sqrt{3} = 0.01/\sqrt{3} = 0.0058 \;\; \%$$

TrainMiC
Training in Metrology in Chemistry

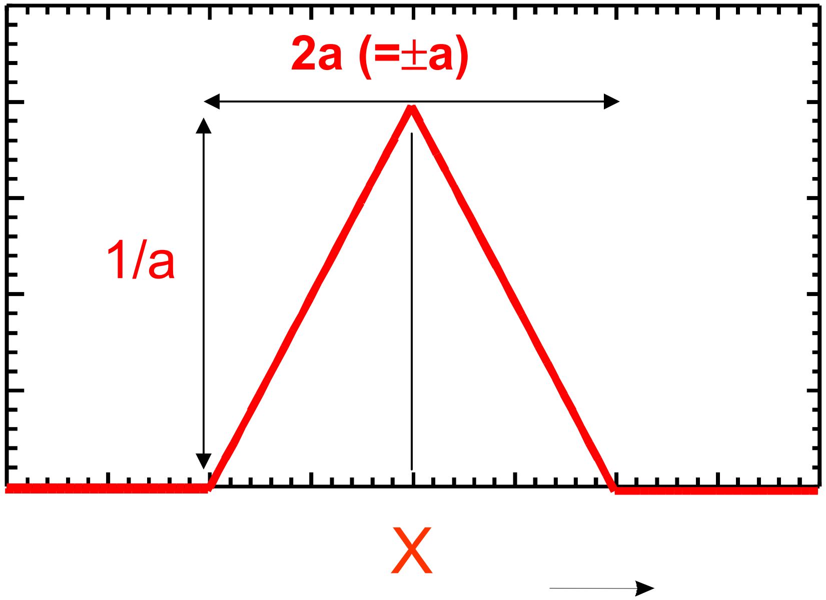
Triangular distribution

Distribution used when it is suggested that values near the centre of range are more likely than near to the extremes

$$y = x \pm a$$

Assumed standard deviation:

$$s = a \cdot 1 / \sqrt{6}$$



TrainMiC
Training in Metrology in Chemistry

Example of Triangular distribution

Values close to x are more likely than near the boundaries

The available information concerning the value is less limited than for rectangular distribution.

Example *(volumetric glassware)*

The manufacture quotes a volume for the flask of
(100 ± 0.1) ml at T = 20° C.
Nominal value most probable!
Assuming triangular distribution the standard uncertainty is:

$$s = u(x) = a \cdot 1/\sqrt{6} = 0.1/\sqrt{6} = 0.04\ ml$$

In case of doubt, use the rectangular distribution

Confidence Interval

The individual observations are distributed about the best estimate of the "True Value" with a spread, which depends on the precision

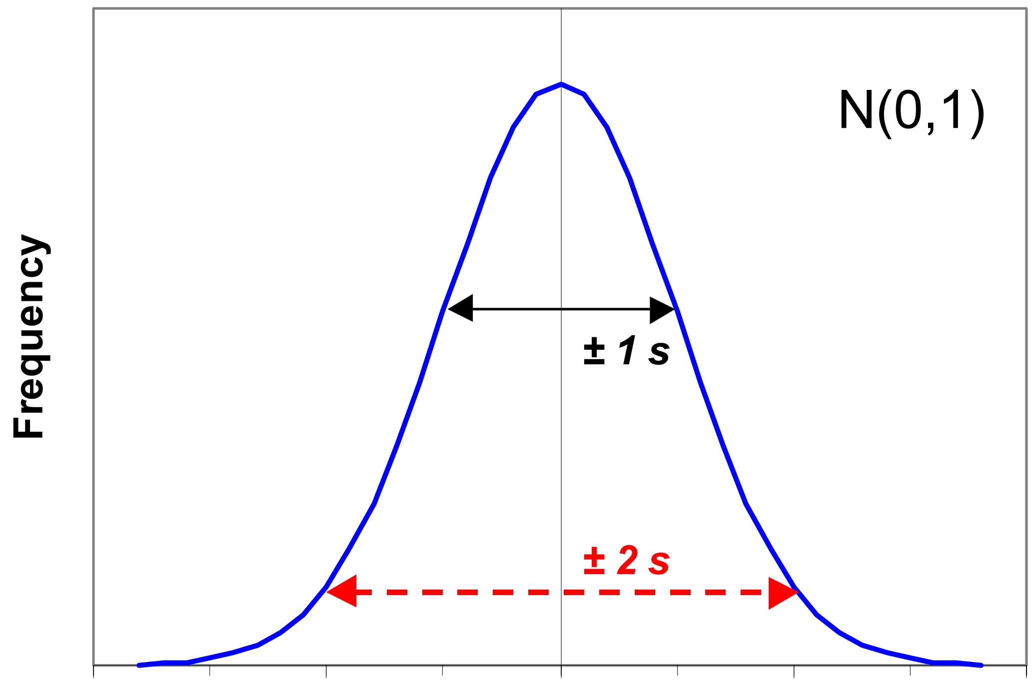
The estimate of the "True Value" *(μ)* lies within the confidence interval *(CI)*, with a probability of *($1-\alpha$)*, *having "n"* degrees of freedom:

$$\mu = \bar{x} \pm (1-\alpha)\%\ CI\ (n)$$

$$95\%\ CI = t(0.05, n-1) * s / \sqrt{n}$$

TrainMiC
Training in Metrology in Chemistry

Confidence Interval (2)



$\mu \pm 1s$	68 %
$\mu \pm 2s$	95 %
$\mu \pm 3s$	99.7%

TrainMiC
Training in Metrology in Chemistry

Law of "Uncertainty Propagation"

$$Y = f(X_1, X_2, \ldots\ldots, X_n)$$

$$u_c^2(Y) = \sum \left(\frac{\partial f}{\partial X_i} \right)^2 \cdot \left(u(X_i) \right)^2$$

$$C = (a + b)$$
$$C = (a - b)$$
⇒
$$u(C) = \sqrt{u(a)^2 + u(b)^2}$$

$$C = (a * b)$$
$$C = (a / b)$$
⇒
$$\frac{u(C)}{C} = \sqrt{\left(\frac{u(a)}{a} \right)^2 + \left(\frac{u(b)}{b} \right)^2}$$

Uncertainty "Type"

Type A evaluation of uncertainty:

statistical analysis of series of observations.

Type A standard uncertainty is measured from repeatability experiments and is quantified in terms of *the standard deviation* of the measured values

Type B evaluation of uncertainty:

by *other means* than statistical analysis

(previous experiments, literature data, manufacturer's information)

[GUM, 1993]

Training in Metrology in Chemistry (TrainMiC)

According to GUM...

Combined standard uncertainty, $u_c(y)$, is obtained by combining the standard uncertainties of the input quantities using the law of propagation

$$u_c^2(y) = \sum \left(\frac{\delta f}{\delta x_i} \right)^2 \cdot \left(u(x_i) \right)^2$$

Expanded Uncertainty, U, is obtained by multiplying the combined standard uncertainty by a coverage factor k:

$$U(y) = k * u_c(y)$$

often $k = 2$

TrainMiC
Training in Metrology in Chemistry

An uncertainty is given in the form of Standard Deviation *[s = u(x)]*

$$R = \overline{x} \pm \Delta x$$

But what is Δx ?

- Standard deviation ?
- Rectangular distribution uncertainty ?
- Triangular distribution uncertainty ?
- Confidence interval w/o specified degree of freedom ?
- Confidence Interval with specified degree of freedom ?
- Combined Uncertainty *?*
- Expanded uncertainty *? Is “k” specified ?*

Standard deviation of a single measurement

0. Experimental Measurement ➔ Type A uncertainty !

1. Single measurement with several instrumental replicates:

$$R = \bar{x} \pm s$$

s

- provided by the instrument
- calculated from (instrumental) replicates

TrainMiC
Training in Metrology in Chemistry

Standard deviation of n independent measurements

2. Several (**n**) independent measurements with several instrumental replicates

$$R_i = \bar{x}_i \pm s_i$$

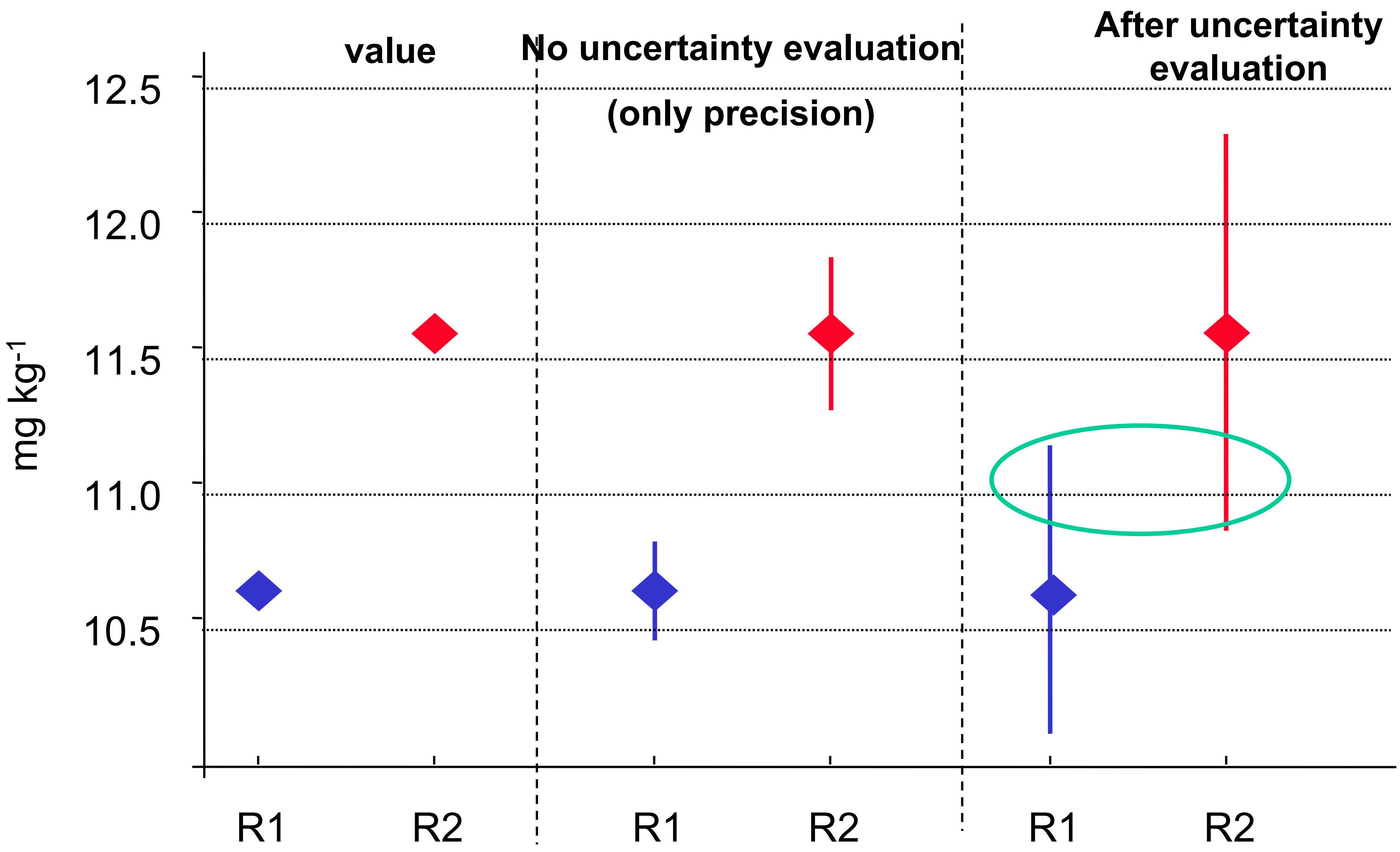
assuming that ALL s_i are similar *(= s)*

$$R_i = \bar{x}_i \pm s$$

$$R = (\bar{R}_i) \pm s_{mean} = (\bar{R}_i) \pm \frac{s}{\sqrt{n}}$$

TrainMiC
Training in Metrology in Chemistry

Are these results different?



TrainMiC
Training in Metrology in Chemistry

(Traditional) Statistical Approach

Measurement Cd content in plant
3 digested samples

1st Digestion : 22 mg/kg		mean [Cd] = 21.0 mg/kg
2nd Digestion : 21 mg/kg	⇨	
3rd Digestion : 20 mg/kg		(stdev) s = 0.21 mg/kg

(mean ± stdev) C_{Cd} = (21.0 ± 0.2) mg/kg
(mean ± 95% CI) C_{Cd} = (21.0 ± 0.5) mg/kg, with n = 3

t(0.05,2)= 4.3

TrainMiC
Training in Metrology in Chemistry

GUM Approach

Measurement Cd content in plant
3 digested samples

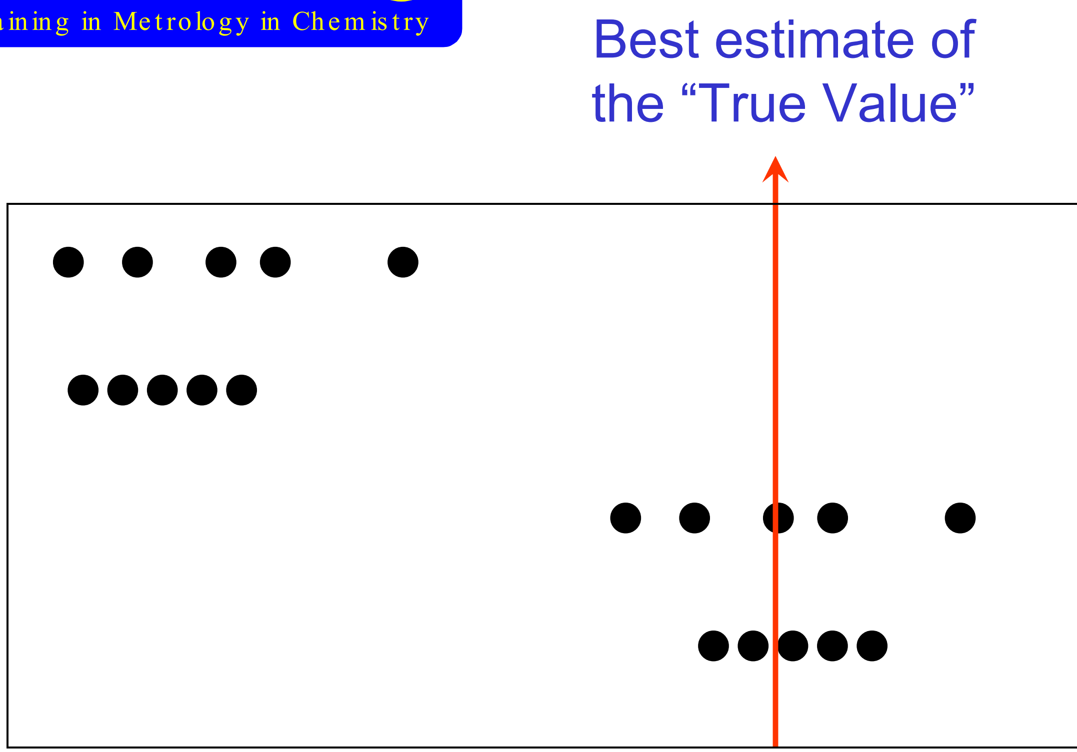
1st Digestion : 22 mg/kg	mean [Cd] = 21.0 mg/kg
2nd Digestion : 21 mg/kg ⇒	Combined unc. u_c = 2.1 mg/kg
3rd Digestion : 20 mg/kg	

Uncertainty Budget calculation ➔ Combined Uncertainty (including contribution from all parameters)

mean ± Expanded uncertainty

C_{Cd} = (21.0 ± 4.2) mg/kg, with k = 2

Statistics for method performance studies



Accurate?	Precise?
no	no
no	yes
yes	no
yes	yes
(close)	*(scatter)*

Precision: The closeness of agreement between independent test results obtained under stipulated conditions [ISO 5725]

Precision ↗ ⇒ Scatter ↘ ⇒ uncertainty ↘

Accuracy

Closeness of agreement between a test result of a measurement and the accepted reference value (ISO 3534-1)

Accuracy is not given by the spread of a normal distribution, but by the deviation of the arithmetic mean of a series of results from accepted reference value

Accuracy ↗ ⇒ Bias ↘ (zero)

Repeatability

Precision recorded under repeatability conditions:

– same laboratory, analyst, equipment, time (short interval)

Typically used for studying variation within a batch or between replicated measurements.

Within-run precision = Repeatability

TrainMiC
Training in Metrology in Chemistry

Reproducibility

Precision recorded under reproducibility conditions:

- different laboratory, analyst, equipment, time (short interval)

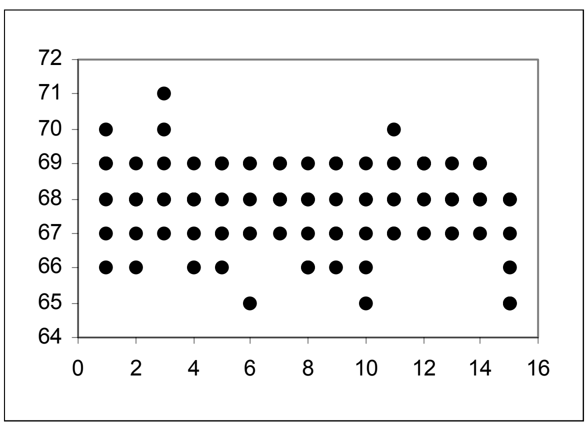
Typically used for studying variation

on measurements made between laboratories.

Between-run precision = Reproducibility

Anova Single factor

Replicates		1	2	3	4	5	6
Vials	1	66	68	67	69	70	69
	2	66	67	68	68	68	69
	3	71	67	68	69	68	70
	4	66	68	67	68	68	69
	5	67	67	66	69	69	68
	6	65	67	67	69	68	69
	7	67	68	68	68	69	69
	8	67	66	66	68	68	69
	9	67	67	66	69	68	69
	10	66	65	67	68	69	68
	11	67	67	69	68	68	70
	12	67	68	69	69	68	69
	13	67	67	68	69	68	68
	14	67	68	68	69	68	69
	15	65	66	65	68	68	67



$$R = 2^* \sqrt{2} * s_R$$
$$r = 2^* \sqrt{2} * s_r$$

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
1	6	409	68.2	2.2
2	6	406	67.7	1.1
3	6	413	68.8	2.2
4	6	406	67.7	1.1
5	6	406	67.7	1.5
6	6	405	67.5	2.3
7	6	409	68.2	0.6
8	6	404	67.3	1.5
9	6	406	67.7	1.5
10	6	403	67.2	2.2
11	6	409	68.2	1.4
12	6	410	68.3	0.7
13	6	407	67.8	0.6
14	6	409	68.2	0.6
15	6	399	66.5	1.9

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	26.2	14	1.87	1.34	0.207	1.83
Within Groups	104.8	75	1.40			
Total	131.0	89				

repeatability stdev	s_r	1.18	=sqrt(MSW)
reproducibility stdev	s_R	1.21	=sqrt(MSW+(MSB-MSW)/N)
			(n replicates)

Statistics for Inter-Laboratory Comparison (ILC), Proficiency Testing (PT)

TrainMiC
Training in Metrology in Chemistry

(Traditional) Z-score

$$Z = \frac{x_{lab} - x_{ref}}{"s"}$$

Difference ➔ distance ➔ accuracy

"Normalized" versus ...

- Target performance *(i.e. 5%)*
- Reference uncertainty (nominal value)
- Inter-Laboratory Comparison reproducibility

Performance evaluation:

$0 < |Z| < 2$: good

$2 < |Z| < 3$: warning ➔ preventive action

$|Z| > 3$: unsatisfactory ➔ corrective action

TrainMiC
Training in Metrology in Chemistry

En-score according to GUM

$$En = \frac{x_{lab} - x_{ref}}{\sqrt{(u_{lab}^2 + u_{ref}^2)}}$$

"Normalized" versus ...

propagated combined uncertainties

Performance evaluation:

```
0 <|En|< 2 : good
2 <|En|< 3 : warning → preventive action
   |En|> 3 : unsatisfactory → corrective action
```

TrainMiC
Training in Metrology in Chemistry

The Uncertainty Budget
Step-by-step Tutorial

TrainMiC
Training in Metrology in Chemistry

❶ Model: $Y = X_1 * X_2 / (X_3 * X_4)$ *part 1*

❷

RSD	stdev	value	description
??	0,02	2,46	**X1**
3,0%	??	4,32	**X2**
??	0,11	6,38	**X3**
2,3%	??	2,99	**X4**

❸

RSD	stdev	value	description
0,8%	0,02	2,46	**X1**
3,0%	0,13	4,32	**X2**
1,7%	0,11	6,38	**X3**
2,3%	0,07	2,99	**X4**

❹

RSD	stdev	value	description
0,8%	0,02	2,46	**X1**
3,0%	0,13	4,32	**X2**
1,7%	0,11	6,38	**X3**
2,3%	0,07	2,99	**X4**
??	**??**	**0,557**	Result

TrainMiC
Training in Metrology in Chemistry

❶ Model: $Y = X_1 * X_2 / (X_3 * X_4)$ *part 2*

❺

RSD	stdev	value	description	X1	X2	X3	X4
0,8%	0,02	2,46	**X1**		2,46	2,46	2,46
3,0%	0,13	4,32	**X2**	4,32		4,32	4,32
1,7%	0,11	6,38	**X3**	6,38	6,38		6,38
2,3%	0,07	2,99	**X4**	2,99	2,99	2,99	
??	**??**	**0,557**	Result				

❻

RSD	stdev	value	description	X1	X2	X3	X4
0,8%	0,02	2,46	**X1**	2,48	2,46	2,46	2,46
3,0%	0,13	4,32	**X2**	4,32	4,45	4,32	4,32
1,7%	0,11	6,38	**X3**	6,38	6,38	6,49	6,38
2,3%	0,07	2,99	**X4**	2,99	2,99	2,99	3,06
4,2%	**0,024**	**0,557**	Result	0,562	0,574	0,548	0,544

x+Δx

❼

diff	0,005	0,017	-0,009	-0,013	0,001

sumsq($diff_i$)

$$u_c = \sqrt{\sum_i (y_i - y)^2}$$

TrainMiC
Training in Metrology in Chemistry

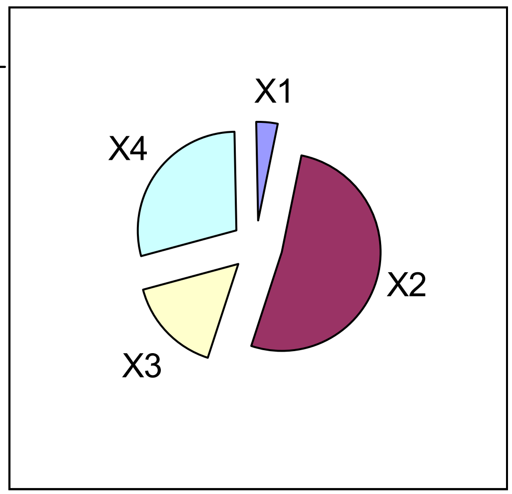
❶ Model: $Y = X_1 * X_2 / (X_3 * X_4)$ *part 3*

RSD	stdev	value	description	X1	X2	X3	X4	
0,8%	0,02	2,46	**X1**	2,48	2,46	2,46	2,46	
3,0%	0,13	4,32	**X2**	4,32	4,45	4,32	4,32	
1,7%	0,11	6,38	**X3**	6,38	6,38	6,49	6,38	
2,3%	0,07	2,99	**X4**	2,99	2,99	2,99	3,06	
4,2%	**0,024**	**0,557**	Result	0,562	0,574	0,548	0,544	
			diff	0,005	0,017	-0,009	-0,013	0,001

❽

index	3,7%	**50,8%**	16,1%	**29,4%**	100,0%
					sum

$$index = \frac{(y_i - y)^2}{\sum_i (y_i - y)^2}$$



Major Contributor :

- Type B? ☹
- Type A? ☺
- Replicates?
- Much work?
- Control Charts?