Statistical Aspects of Measuring Uncertainties in Laboratories and Interlaboratory Tests

What are telling us the different characteristic values of the statistics of measurements in our laboratories and in interlaboratory tests (Round Robins)

„There a three kinds of lies:
lies, damned lies and statistics“

Benjamin Disraeli (1804 – 1881)

Determination of vertical ball behaviour - hockeyball

R

Sample A
EN 12235

Results submitted by participants

i.e., individual results x_{ik}
+ number of the test repetitions made by each lab (n_i)
+ within laboratory means (x_i) and standard deviations (s_i)
+ results of tests for outliers

Number of reporting laboratories p *: 6
Number of reported test results Σn_i: 12

Lab Code No.	Test results in % — Test replication No. (k) 1	2	3	4	5	6	Statistical evaluation of the submitted test results x_{ik} — n_i	x_i	s_i	Outliers — Cochran	Grubbs	$\lvert z \rvert > 2$
362	26,0	26,1					2	26,04	[illegible]		**	X
42	49,9	49,8					2	49,84	0,0283			
452	50,8	51,1					2	50,93	0,2404			
738	52,5	52,4					2	52,49	0,0707			
547	54,1	54,3					2	54,19	0,1273			
347	54,3	54,9					2	54,58	0,4243			
865			no results reported									X

** ... statistical outlier (99%)　　* ... straggler (95%)
X ... z > 2

Results of robust statistics

Convergence assumed at iteration number: 20

Robust average: $x^* = 51,3$	assigned value for the proficiency assessment		
Robust standard deviation for the proficiency assessment: $s^* = 3,75$			
Number of repeate measurements necessary due to s_r/s^*-ratio: $n' = 1$	OK	see page 4 for the meaning of NOT OK	
Standard uncertainty of the assigned value: $u_x = 1,91115$	NOT OK		

Additional check of the test method accuracy

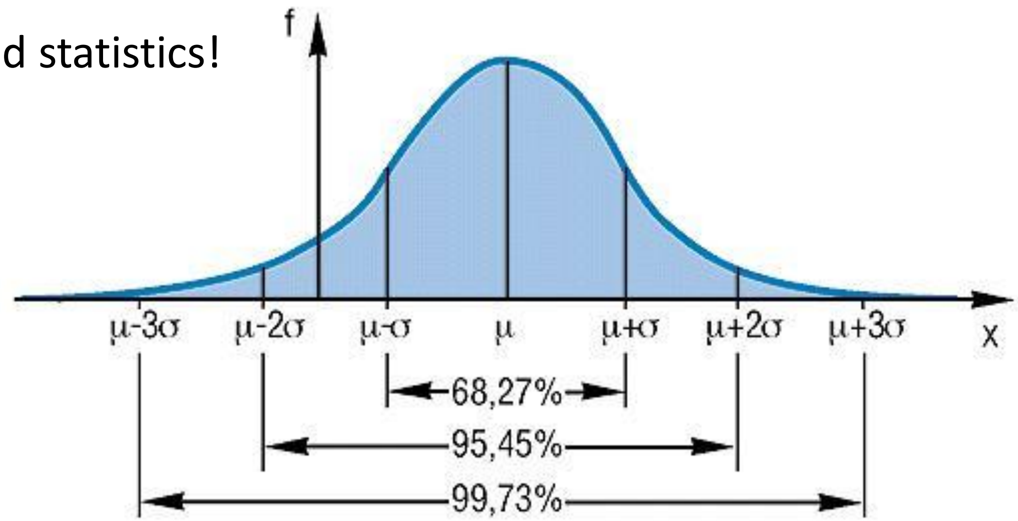
Do the input data come from a normal distribution ? (The results listed below shall be considered as really justified only if the input data come from a normal distribution)			YES
General mean $\Sigma n_i x_{ik} / \Sigma n_i$	m	52,4	%
Repeatability variance	s_r^2	0,0519600	
Repeatability standard deviation	s_r	0,22795	%
Repeatability coefficient of variation	$CV\%_r$	0,435	%
Between-laboratory variance	s_L^2	4,1437500	
Between-laboratory standard deviation	s_L	2,03562	%
Between-laboratory coefficient of variation	$CV\%_L$	3,884	%
Reproducibility variance s_R^2	$s_r^2+s_L^2$	4,1957100	
Reproducibility standard deviation	s_R	2,0483	%
Reproducibility coefficient of variation	$CV\%_R$	3,909	%
Repeatability limit	r	0,6	%
Relative repeatability limit	r_{rel}	1,22	%
Reproducibility limit	R	5,7	%
Relative reproducibility limit	R_{rel}	10,94	%
Number of participants included in the accuracy evaluation	p	5	
Number of tests included in the accuracy evaluation	Σn	10	

What´s our goal when measuring a property?

- to find the true value of what we have to measure !!!
  (that´s what your customer expects from you because he pays for it)

Unfortunantely the true value will only be obtained by the average of an infinite numbers of single results performed by an infinte number of laboratories.

Thats why we need statistics!



***Gaussian Bell Curve* or Normal Distribution**

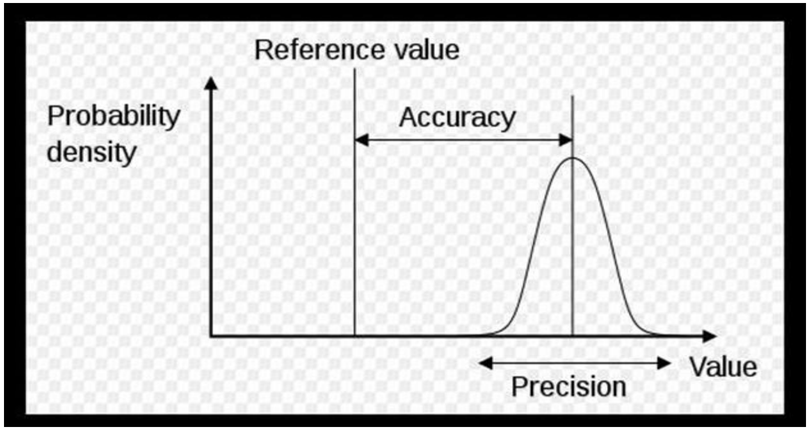
Laboratory statistics and interlaboratory statistics (Round Robins) give us the following idea about what we are measuring:

(1) How close is my measured value to the true value within certain (statistical) borders when measuring one, two or three times as it is ususal in our daily laboratory routine!

(2) When does a property of a measuring object do not meet the requirements, means failing the specification!!

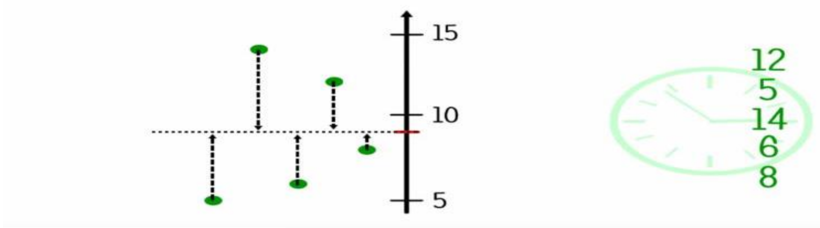
So what must an accurate measurement procedure **always** fullfill?

accuracy = trueness (systematical errors) and pecision (random errors)



Some necessary definitions to characterize measured data

(1) The Standard Deviation (SD) s (σ = true value in Normal Distribution)



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

n-1= estimates the SD of the universe
n = estimates the SD of the test

(!) The SD is a measure of the distribution of measured values

(2) The Variance (VAR)

$$\mathbf{s}^2 = \frac{\mathbf{1}}{(\mathbf{n}-\mathbf{1})} \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} (\mathbf{x}_\mathbf{i} - \bar{\mathbf{x}})^2$$

(!) The VAR is a measure of how the measured values are distributed around the mean value

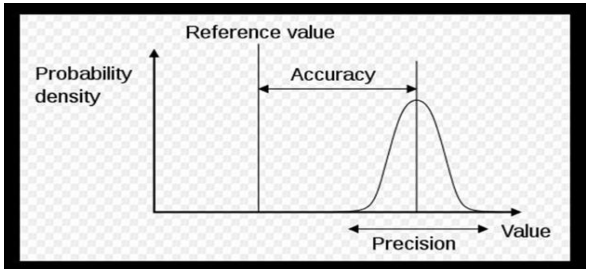
We always have to deal with two kinds of statistics:

(a) the laboratory statistics in your own lab

(b) the interlaboratory statistics between different laboratories (Round Robins)

Please keep in mind that both statistics about your measurement uncertainties are based on same principles!

trueness (systematical errors) and pecision (random errors)



(a) In your own lab: GUM = $u = \sqrt{u_{R_W}^2 + u_{bias}^2}$

(b) In Interlaboratory Tests (Round Robins) = $s_R = \sqrt{s_r^2 + s_L^2}$

Let us now talk about the interpretation of interlaboratory tests (Round Robins)

key figure	calculation	meaning
u_x	$u_{C_{ref},i} = 1{,}25 \cdot \frac{s_{R,i}}{\sqrt{n_{p,i}}}$	Uncertainty of the assigned value x^* 1,25 x 3,75/√6 = 1,911 -> 0,3 x s^* = 1,12
x^*	ISO 13528 DIN ISO 5725-5	robust mean regarding all individual means (assigned value!)
s^*	ISO 13528 DIN ISO 5725-5	robust mean standard deviation regarding all individual SD's (target value for analytical performance called z-score)
s_r	$s_r^2 = \frac{\sum s_i^2}{p}$,	SD of all individual SD's (repeatability)
s_L	$s_L^2 = \frac{1}{\bar{n}}\left[\frac{1}{p-1}\sum_{i=1}^{p} n_i(\bar{x}_i - \bar{\bar{x}})^2 - s_r^2\right]$	SD of all differences from the individual mean to the assigned value x^*
s_R	$s_R = \sqrt{s_r^2 + s_L^2}$	the overall SD taking into account the precision and trueness

Why are these chracteristic values s_r and s_R so important?

r = 2,8 * s_r (repeatability limit 95 % confidence interval)
R = 2,8 * s_R (reproducibility limit 95 % confidence interval)

$$CD = 2{,}8 \cdot s_r \cdot \sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}$$

How do your employees work?

$$CD = \sqrt{(2{,}8\, s_R)^2 - (2{,}8\, s_r)^2\left(1 - \frac{1}{2n_1} - \frac{1}{2n_2}\right)}$$

Do two labs have the same performance?

$$CD = \frac{1}{\sqrt{2}} \cdot \sqrt{(2{,}8\, s_R)^2 - (2{,}8\, s_r)^2\left(\frac{n-1}{n}\right)}$$

y‘ - m_0 : Has the test object failed in your lab?

$$CD = \frac{1}{\sqrt{2p}} \cdot \sqrt{(2{,}8\, s_R)^2 - (2{,}8\, s_r)^2\left(1 - \frac{1}{p}\sum_{i=1}^{p}\frac{1}{n_i}\right)}$$

y‘‘- m_0 : Has the test object failed in different labs?

2 employees making n_i determinations / 2 labs making n_i determinations
p labs making n_i determinations to confirm a reference value m_0

Now, what to do when the critical differences (CD) are exceeded??

Don‘t worry be happy and ask the DIN ISO 5725 or the EN ISO 4259 Annex 1, because they will tell you what to do to avoid a lawsuit with your customer or to fire one of your employees.

Thank you for having so much patience with statistics and the guy who was talking about it.

For further Informations:

Darell Huff, Irving Geis, ***How to Lie with Statistics,*** W.W. Norton & Company Ltd, New York, 1954