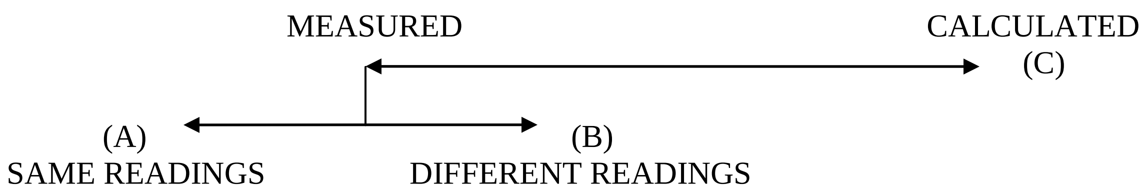
BRIEF SURVEY OF UNCERTAINITY IN PHYSICS LABS

THREE CASES OF UNCERTAINTY CALCULATION

There are two main situations when dealing with uncertainty calculation of a given parameter; or it is measured or it is calculated through some measured parameters by use of an expression.
Also, after several repeated measurements of a given measured parameter, there are two situations: one gets always the same reading or one gets different readings.



As a first step, one has to decide which case (A, B. C) one is dealing with.

CASE "A" - You measure several times a parameter and you get always the same numerical reading.

Example_1: You measure five times the length of a table with a meter stick which smallest unit is 1cm and read L= 85cm. Actually , you see that the reading is a little bit more or less than 85cm but you cannot read a precise number because you <u>cannot be precise</u> about what portion of 1cm is the quantity "a little bit *more or less". In such situations one uses "the half-scale rule" i.e.; the <u>absolute uncertainty</u> of the <u>measurement</u> is taken equal to the <u>half of the smallest unit available used for measurement</u>. In this case absolute uncertainty is $\Delta L= \pm 0.5cm$ and the result of measurement is reported as L= (85.0 ± 0.5)cm and the interval of uncertainty for the estimation is (84.5, 85.5)cm.*

-If one uses a meter stick with smallest unit available 1mm, one is going to have a more precise result but even in this case there is an uncertainty. Suppose that you get always the reading L= 853mm. Being aware that there is always a parallax error (eye position) on both sides reading, one gets easily[1] $\Delta L= \pm 1$, ±2, and even ±3mm) depending on the measurement circumstances. The result of measurement is reported as L= (853 ±1)mm or (853 ± 2)mm or (853 ± 3)mm . The <u>best estimation</u> for the table length is 853mm. If the <u>absolute uncertainty</u> of estimation is $\Delta L= \pm 1mm$, than the true length is inside the <u>uncertainty interval</u> (852, 854)mm.

-Assume that, using the same meter stick, you measure the length of a calculator and a room and find $\mathbf{L_{calc}= (14.0 \pm 0.5)cm}$ and $\mathbf{L_{room}= (525.0 \pm 0.5)cm}$. Both measurements have the ***same absolute uncertainty*** $\mathbf{\Delta L= \pm 0.5cm}$ but we are conscious that the length of **room** is ***measured more precisely. The precision of a measurement*** is estimated by the ***portion of uncertainty that belongs to the unit of measured length.***

Actually, it is ***estimated by the relative error***
$$\varepsilon = \frac{\Delta L}{\bar{L}} * 100\% \qquad \textbf{(1)}$$

[1] To get ΔL= ***±0.5mm one must apply special conditions (like a optic devicefor reading...)***

-Note that smaller relative error means higher precision of measurement. For these measurements, one gets $\varepsilon_{calc} = \frac{0.5}{14} * 100\% = 3.57\% \approx 4\%$ ***and*** $\varepsilon_{room} = \frac{0.5}{525} * 100\% = 0.095\% \approx 0.1\%$**.**

So, the room length is measured much more precisely (about 40 times).

Note: Don't mix the precision with accuracy. A measurement is accurate if uncertainty interval contains an expected (known by literature) value and is not accurate if it does not contain it.

CASE "B" - You measure several times a parameter and you get different numerical readings.

Example_2: You drop an object from a window and expect it to hit ground after 2sec. To verify your expectation, you *measure* this *time* several times and record the following results;

1.99s, 2.01s, 1.89s, 2.05s 1.96s, 1.99s, 2.68s, 1.97s, 2.03s, 1.95s

Note: ***5 measurements*** is a ***minimum acceptable number of repeated data readings***, i.e. repeat the same measurement 5 times. Due to time restrictions one may accept until 3 repeated measurements; the estimation based on 1 or 2 repeated measurements is not reliable.

As a first step one proceeds by: ***VERIFYING THE VALIDITY OF RECORDED DATA***

The drawing of graphs during lab measurements is practical way to estimate quickly:

a) Whether the measurements confirm the expected behaviour predicted by the physics model.
b) If any of recorded data is measured in wrong way and must be excluded from further treatments.

To check out the data one includes them in a graph (fig.1). From this graph one can see that:

a) The fall time seems to be *constant* and very likely ~2s. So, in general, these are acceptable data.
b) Only the seventh measure is too far from the others results and this may be due to an abnormal circumstance during its measurement. To eliminate any doubt, one ***exclude*** this value from the following data analysis. The remaining nine date are enough for following calculations.
1.99s, 2.01s, 1.89s, 2.05s, 1.96s, 1.99s, 1.97s, 2.03s, 1.95s. .

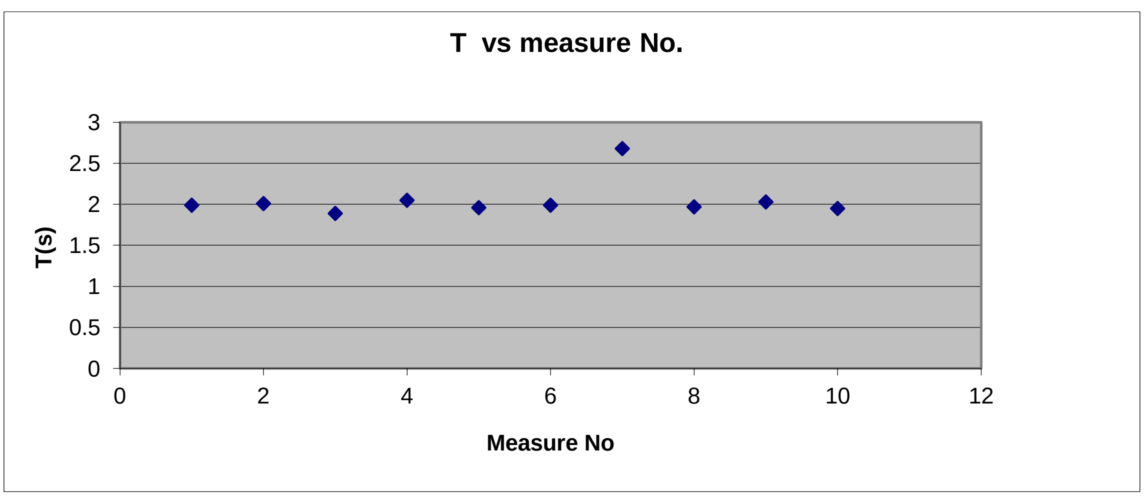


Fig.1

<u>*As a second step one proceeds by*</u> ***ORGANIZING RECORDED DATA IN A TABLE***

Include all data in a table organized in such a way that some cells be ready to include the results of uncertainty calculations. In this example, you are looking to estimate a single parameter "T", so you have to predict (at least) two cells; one for its average and one for its uncertainty.

Table_1

T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	Tav	ΔT
1.99s	2.01s	1.89s	2.05s	1.96s	1.99s	1.97s	2.03s	1.95s		

Then one follows with ***CALCULATIONS OF UNCERTAINTIES***

Keep in mind that the true value of parameter is unknown. So, one uses the recorded data to find an <u>estimation</u> of the true value and the <u>uncertainty</u> of this estimation.
In case ***"B"*** the ***best estimation*** of parameter is the ***average*** of recorded data and ***a good estimation*** of ***absolute uncertainty*** is the ***mean deviation*** of recorded data***.***

So, for the data collected in experiment_2

b.1) The best estimation for falling time is the **<u>average</u>**.

$$\bar{T} = \frac{1}{9}\sum_{i=1}^{9} T_i = \frac{1}{9}[1.99 + 2.01 + 1.89 + 2.05 + 1.96 + 1.99 + 1.97 + 2.03 + 1.95] = 1.982 \cong 1.98s \quad (2)$$

b.2) One may get a good **estimation** for **absolute uncertainty by using** the ***spread of measured data***. and more precisely their ***<u>mean deviation</u>*** ***from*** their ***average value***.

$$\Delta T = \frac{1}{n}\sum_{i=1}^{n}\left|T_i - \bar{T}\right| = \frac{1}{9}\sum_{i=1}^{9}\left|T_i - 1.982\right| = 0.0353\text{s} \cong 0.04\text{s} \quad (3)$$

This way, one can say that the **true value** *of fall time* is inside the ***<u>uncertainty interval</u>*** (1.94, *2.02*)***sec*** or between T_{max} = 2.02***s*** and T_{min} =1.94***s*** with **best estimation** 1.98***s***. (after *rounding off*)

The result is reported as ***T = (1.98 +/- 0.04)sec*** (4)

b.3) <u>*Another (**statistically better**) estimation of spread*</u> is the "**standard[2] deviation**" of data.

Based on our example data we get

$$\sigma T = \sqrt{\frac{\sum_{i=1}^{n}\left(T_i - \bar{T}\right)^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{9}\left(T_i - 1.982\right)^2}{8}} = 0.047s. \quad (5)$$

The result is reported as ***T = (1.98 +/- 0.05)sec*** (6)

For *spread estimation*, a ***larger interval of uncertainty*** means a more "***conservative estimation***"

[2] The standard deviation can be calculated direct in Excel and in many calculators.

but in the same time a ***more reliable estimation***. That's why the ***standard deviation*** is a better estimation for the ***absolute uncertainty***. Note that we get $\Delta T = \pm 0.05s$ when using the ***standard deviation*** and $\Delta T = \pm 0.04s$ when using the ***mean deviation.*** *Also, the* ***relative error (or relative uncertainty)*** *calculated from the* ***standard deviation*** is bigger. In our example the ***relative uncertainty*** of measurements is

$$\varepsilon = \frac{\sigma T}{\bar{T}} * 100\% = \frac{0.047}{1.982} * 100\% = 2.4\%$$ when using the ***standard deviation***

and $$\varepsilon = \frac{\Delta T}{\bar{T}} * 100\% = \frac{0.035}{1.982} * 100\% = 1.8\%$$ when using the ***mean deviation***

The relative uncertainty 2.4% estimates less precision of measurements but it is more reliable.

Important: The absolute and relative uncertainty can never be zero.

-Assume that you **repeat** 5 times a given measurement and you read all times the **<u>same value</u>** ***X***. So, by applying the rules of case "B" you would report $X_{best} = X_{Av} = 5X/5 = X$ ***and*** $\Delta X_b = 0$. ***But here you are actually*** *dealing with a* ***case "A"*** and this means that there is a $\Delta X_a (\neq 0) = \frac{1}{2}$***(smallest unit of measurement scale).*** This example shows that, in general, the absolute uncertainty should be calculated as

$$\Delta X = \Delta X_a + \Delta X_b \qquad (7)$$

Note that in those cases where $\Delta X_b >> \Delta X_a$ ***one may simply disregards*** ΔX_a.

In exemple_2 the time is measured with 2 decimals. This means that $\Delta X_a = \Delta T_a = 1/2(0.01) = 0.005s$
As $\Delta X_b = \Delta T_b = 0.05s$ *is* ***ten times bigger one <u>may</u>*** *neglect* $\Delta T_a = 0.005s$ and get $\Delta T = 0.005s$. ***But if*** ΔT_b *were* $= 0.02s$ *and* $\Delta T_a = 0.005s$ *one should not neglect* $\Delta T_a = 0.005s$ *because it is 25% of* ΔT_b. *So, one would have to use the expression (7) and calculate the absolute uncertainty as* $\Delta T = 0.02 + 0.005 = 0.025s$

Note: In college labs one considers that a measurement has a *good precision* if the ***relative uncertainty*** $\varepsilon < 10\%$. If the relative uncertainty is $\varepsilon > 10\%$, you may proceed by:

a) *Cancelling* any particular data "***shifted too much from the average value***" ;
b) *Increasing* the number of recorded data by repeating more times the measurement;
c) *Improving* the measurement procedure.

CASE "C" Estimation of Uncertainties for Calculated Quantities (Uncertainty propagation)

Very often, one uses the experimental data recorded for some parameters and a mathematical expression to estimate the value of a given parameter of interests (POI). As the measured parameters are estimated with a certain uncertainty, it is clear that the estimation of POI with have some uncertainty, too. Actually, the calculation of best estimation for POI is based on the best estimations of measured parameters and the formula that relates POI with measured parameters. Meanwhile, the uncertainty of POI estimation is calculated by using the Max_Min method. This method calculates the limits of uncertainty interval, POI_{min} and POI_{max} by using the formula relating POI with other parameters and the combination of their limiting values in such a way that the result be the smallest or the largest possible.

Example_3. To find the volume of a rectangular pool with constant depth, one measures its length L, its width W and its depth D by a meter stick. Then, one calculates the volume by using the formula V=L*W*D. Assume that the measurement results are L = (25.5 ± 0.5)m, W = 12.0 ±0.5m, D = 3.5 ±0.5m

In this case the **best estimation** for the volume is $\mathbf{V_{best}}$ = 25.5*12.0*3.5=**1071.0 m^3**. This estimation of volume is associated by an uncertainty calculated by *Max-Min methods* as follows

$V_{min}=L_{min}*W_{min}*D_{min}$= 25*11.5*3 = 862.5$m^3$ and $V_{max}=L_{max}*W_{max}*D_{max}$= 26*12.5*4 = 1300.0$m^3$

So, the ***uncertainty interval*** for volume is (862.5, 1300.0) and the ***absolute uncertainty** is*

ΔV= (V_{max}-V_{min})/2 = (1300.0 - 862.5)/2= 218.7m^3 while the **relative error** is $\varepsilon_V = \frac{218.7}{1071.0} * 100\% = 20.42\%$

When applying the sig. figure rules V_{best} =1100 m^3; ΔV= 220m^3; ε_V=20.%

Note_1: When applying the Max-Min method to calculate the uncertainty, one must pay attention to the mathematical expression that relates POI to measured parameters.

Examples_4 - You ***measure*** the ***period*** of an oscillation and you use it to ***calculate*** the ***frequency*** (POI). As $f = 1/T$, $f_{av} = 1/T_{av}$ *the* ***max-min method*** *gives* $f_{min}=1/T_{max}$ ***and*** $f_{max}=1/T_{min}$

Examples_5 - *If* $z = x - y$, $z_{av} = x_{av} - y_{av}$ *and* $z_{MAX} = x_{MAX} - y_{MIN}$ *and* $z_{MIN} = x_{MIN} - y_{MAX}$.

Note_2. Use the *best estimations of parameters in the expression* to calculate the *best estimation for POI. If they are missing* one may use POI_{middle} as the best estimation for POI

$$POI_{middle} = \frac{POI_{MAX} + POI_{MIN}}{2} \qquad (8)$$

Be aware though, that POI_{middle} is not always equal to **POI *best estimation.***
So, for the pool volume $\mathbf{V_{middle}}$= (1300+862.5)/2 =**1081.25m3** which is different from $\mathbf{V_{best}}$ **=1071.0 m^3**

<u>HOW TO PRESENT THE RESULT OF UNCERTAINTY CALCULATIONS</u>?

You must provide the **best estimation,** the **absolute uncertainty** and the **relative uncertainty.** So, for the last example, the result of uncertainty calculations should be presented as follows: **V= (1071.0 ± 218.7) m^3 , ε =(218/1071)*100%= 20.42%**

Note: Absolute u*ncertainties* must be ***quoted*** to the ***<u>same number of decimals</u> as*** the **<u>best estimation</u>**. The use of *scientific notation* helps to prevent confusion about the number of significant figures.

Example_6: If calculations generate, say A = (0.03456789 ± 0.00245678)m
This should be presented after being rounded off (leave 1,2 or 3 digits after decimal point):
A = (3.5 ± 0.2) * 10^{-2}m or A = (3.46 ± 0.25) * 10^{-2}m

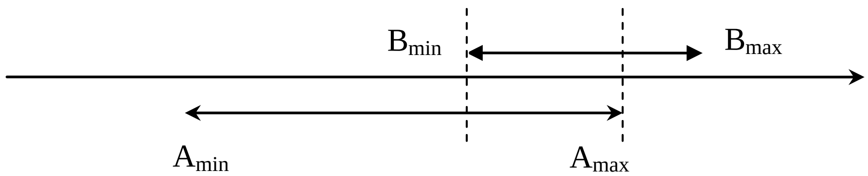
<u>HOW TO CHECK WHETHER TWO QUANTITIES ARE EQUAL?</u>

This question appears essentially in two situations:
1. We measure the ***same parameter*** by two ***different methods*** and want to verify if the results are equal.
2. We use measurements to ***verify*** if ***a theoretical expression*** is right.

In the first case, we have to compare the estimations $A \pm \Delta A$ and $B \pm \Delta B$ of the "two parameters". The second case can be transformed easily to the first case by noting the left side of expression ***A*** and the right side of expression ***B***. Then, the procedure is the same. ***Example***: We want to verify if the thins lens equation ***1/p + 1/q = 1/f*** is right. For this we note ***1/p + 1/q = A*** and ***1/f = B***

Rule: We will consider that the quantities ***A*** and ***B*** are ***equal***[3] if their ***uncertainty intervals overlap***.

Fig.2



WORK WITH GRAPHS

One uses the graphs to check the theoretical expressions or to find the values of physical quantities.

Example_7; Theoretical study of simple pendulum oscillation shows that its period is $T = 2\pi\sqrt{L/g}$ and one wants to verify this experimentally. To make the work easier, one prefers to get a linear relationship between two quantities one can measure; in this case period T and length L. So, one squares both sides of the relation $T^2 = \frac{4\pi^2}{g} * L$, notes $\mathbf{T^2 = y, L = x}$ and gets the <u>linear expression</u> **y = a*x** where $\mathbf{a = 4\pi^2/g}$.
So, one has to verify experimentally if there is such a relation between T^2 and L. Note that if this is verified one can use the experimental value of "**a**" to calculate the free fall constant value "$\mathbf{g = 4\pi^2/a}$".

- Assume that after measuring the period for a given pendulum length several times, calculated the average values and uncertainties for y(=T^2) and repeated this for a set of different values of length x(L=1,...,6m), one gets the data shown in table No 1. At first, one graphs the average data (X_{av},Y_{av}). You see that they are aligned on a straight line, <u>*as expected*</u>. Then, one uses Excel to find the best linear fitting for the data and forces this line to pass from (x = 0, y = 0) because this is predicted from the theoretical formula. One gets a straight line with $a_{av} = 4.065$. Using the theoretical formula one calculates the estimation for $g_{av} = 4\pi^2/a_{av} = 4\pi^2/4.065 = 9.70$ which is not far from ***expected value 9.8***. Next, one adds the uncertainties in the graph(that appear as **error bars**) and draws the best linear fitting with maximum /minimum slope that pass by origin. From the graphs one gets $a_{min} = 3.635$/ $a_{max} = 4.202$.
So, $g_{min} = 4\pi^2/a_{max} = 4\pi^2/4.202 = \mathbf{9.38}$ and $g_{max} = 4\pi^2/a_{min} = 4\pi^2/3.635 = \mathbf{10.85}$

Table_2

X	Y(av.)	ΔY (+/-)	Ymin	Ymax	Max. Slope	Min. Slope
1	4	1.5	2.5	5.5	4.202	3.635
2	8.3	1.8	6.5	10.1		
3	11.8	1.3	10.5	13.1	P1 (1; 1.5)	P1 (1; 5.5)
4	17	1.6	15.4	18.6	P2 (6; 25.5)	P2 (6; 21.5)
5	21	1.1	19.9	22.1		
6	23.5	2	21.5	25.5		

This way, by using the graphs one:

1- has ***proved experimentally*** that the ***relation*** between ***T*** and ***L*** is right.

2- found out that the measurements are ***accurate*** because the ***uncertainty interval*** (9.38, 10.85) for "**g**" does *include the officially accepted value* $\mathbf{g = 9.8 m/s^2}$

3-found the ***absolute error*** $\Delta g = (10.85-9.38)/2 = 0.735 m/s^2$
The relative error is $\varepsilon = (0.735/9.70)*100\% = 7.6\%$ which means a acceptable (< 10%) ***precision of measurement***.

[3] They should be expressed in the same unit, for sure.

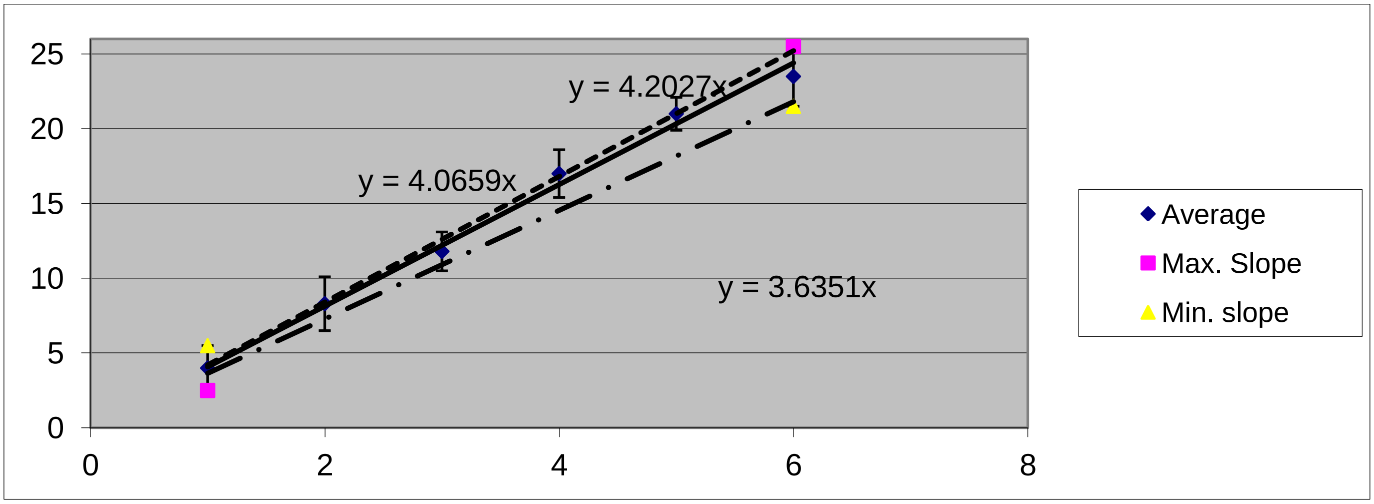
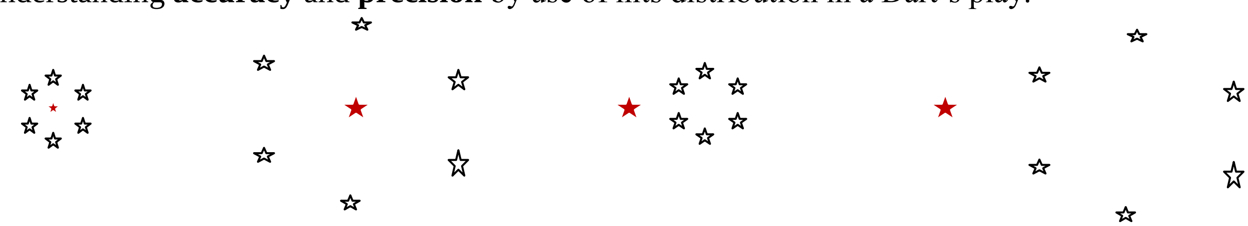


Fig.3

ABOUT THE ACCURACY AND PRECISION

- Understanding **accuracy** and **precision** by use of hits distribution in a Dart's play.



Accurate Good precision | ***Accurate Low precision*** | ***Inaccurate Good precision*** | ***Inaccurate Low precision***

- As a rule, before using a method (or device) for measurements, one should make sure that the method (or device) produces accurate results in the range of expected values for the parameter under study. This is an obligatory step in research and industry and it is widely known as the calibration procedure. During a calibration procedure one records a set of data and makes sure that the result is accurate.

In principle, the *result of an experiment is accurate* if the "***best estimation***" fits to the" *officially accepted value*". You will consider that the experiment is "*enough accurate*" if the" *officially accepted value*" falls ***inside*** *the interval of uncertainty* of measured parameter; *otherwise the result is inaccurate.*

The quantity $\varepsilon_{accu} = \frac{|C_{Av} - C_{off}|}{C_{off}} x100$ ***% (often ambiguously named as error) gives the relative shift of average from the officially accepted value*** $C_{official}$***. It is clear that the accuracy is higher when*** ε_{accu} ***is smaller. But, the measurement is <u>inaccurate if</u>*** $\underline{\varepsilon_{accu} > \varepsilon}$ ***(relative uncertainty of measurement).***

Remember that relative uncertainty $\varepsilon = \frac{\Delta C}{C_{Av}} x100\,\%$ ***is different from*** ε_{accu}.

Note: For an <u>a big number of measured data</u> and <u>accurate measurement</u>, the average should fit to the expected value of parameter and ε_{accu} ***should be practically zero. Meanwhile the relative error ε tents to a fixed value different from zero. Actually, ε can never be equal to zero.***