

CHEM 2003

Analytical Chemistry

Dr Marc Humphries

marc.humphries@wits.ac.za



Course Outline 2018:

- | | |
|--|--------------|
| 1. Statistics and Errors: | Dr Humphries |
| 2. Calibration Methods: | Dr Humphries |
| 3. Volumetric analysis: | Dr Humphries |
| 4. Chemical equilibrium: | Dr Humphries |
| 5. Acids and bases: | Dr Pillay |
| 6. Complexation reactions: | Dr Pillay |
| 7. Redox reactions: | Dr Pillay |
| 8. Precipitation/gravimetric analysis: | Dr Pillay |

D.C. Harris, Quantitative Chemical Analysis

Course Information:

- **Tutorials**

Morning sessions: announced in class

Afternoon sessions: 15 – 18 October (lab)

- **Class test**

19 October 2018 at 10:15

- **Exam**

2 hours: 90 marks

Section A: 45 marks, Section B: 45 marks

Lab outline

Acid-Base Titrations

Determination of the concentration of acetic acid in vinegar

Complexometric Titrations

The determination of the hardness of water

Neutralisation Back Titrations

Analysis of Antacid tablets

Iodometric Titrations

Vitamin C in a commercial vitamin tablet

Spectrophotometry

Phosphate in Cola Drinks

Gravimetric Analysis

Determining the concentration of sulphate in Epsom salts

Method of Standard Addition

Analysis of manganese in steel

Labs

- Lab write ups submitted as **worksheets** or **full reports**.
- All **experimental data** must be entered online.
- All reports must be **typed** and correctly **formatted**.
- Your **signed datasheet** must be attached to the report/worksheet.
- **Late reports** are penalized 10% per day – after more than two days the report will not be marked.
- **Plagiarism** will not be tolerated.
- Marking issues – please consult your demonstrator.

Online data entry

[illegible]

When entering data:

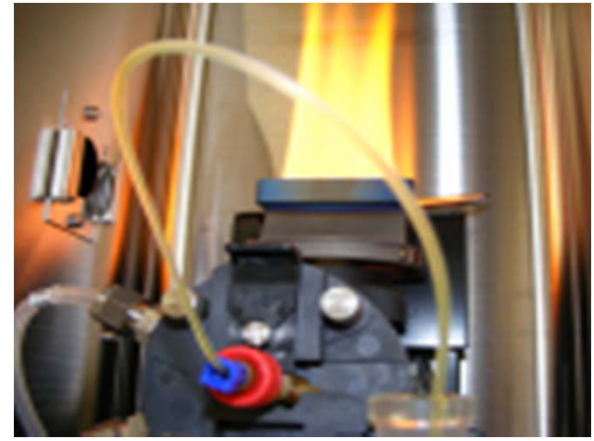
- Do NOT type in units (2.309 not 2.309 g)
- You MUST use a decimal point and NOT the decimal comma (2.30 not 2,30)
- Your data must be entered into the appropriate spreadsheet by Friday afternoon at 14:00 the week of your practical.
- If you fail to enter your data, you will not be awarded any calculation or accuracy marks!

What is chemical analysis

You might think that you are going to be just like CSI



You might end up more like this...



DEFINITION OF THE PROBLEM OR QUESTION

Information gathering

Select analytical technique or method

Implement analysis of known sample and unknowns

Reduce data, interpret and report results

**SOLUTION TO PROBLEM
OR QUESTION**

What is analysis?

- What is present?
- How much?
- What form?

Chemical analysis provides information about the composition of a sample:

- *Qualitative* (what is present in a sample)
- *Quantitative* (how much analyte is present in a sample)

Quantitative analysis

- So far, the only **quantitative** method you have come across has involved *titrations*!
- Not everything can be determined by volumetric titration.

WHY?

Two broad categories of chemical analysis:

1. *Classical methods*: Older methods of analysis based on volumetric analysis
2. *Instrumental methods*: Relatively modern and involve instrumentation (AAS, ICP, GC, XRD)

What is the difference?

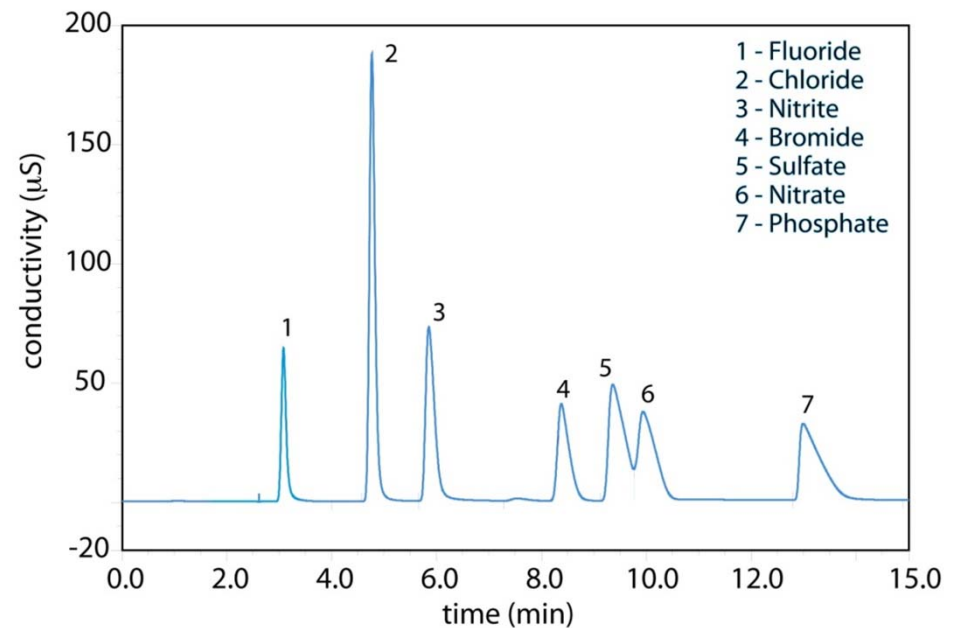
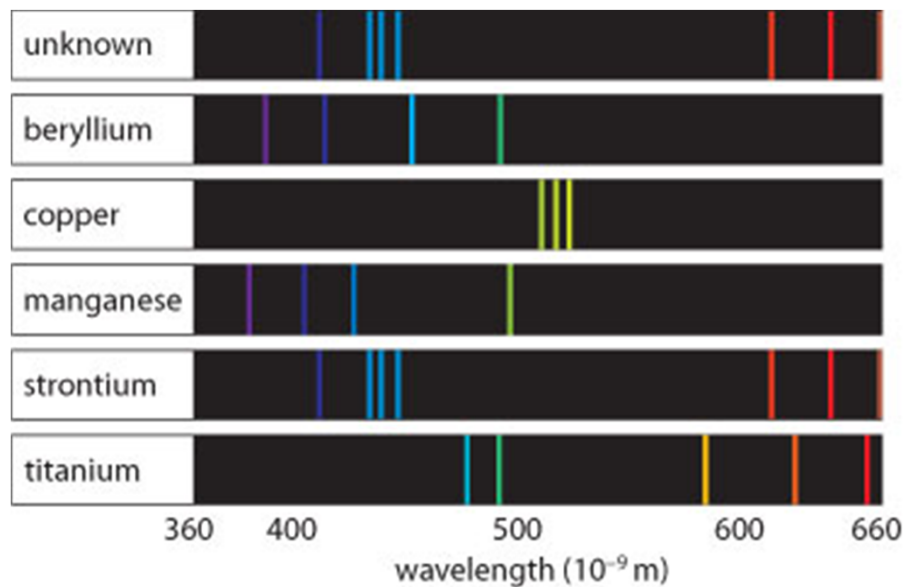
- All forms of chemical analyses result in an **analytical signal**
- *Classical analysis* uses indicators to signal the presence of an analyte or detect an endpoint (e.g. change in colour or formation of a precipitate)

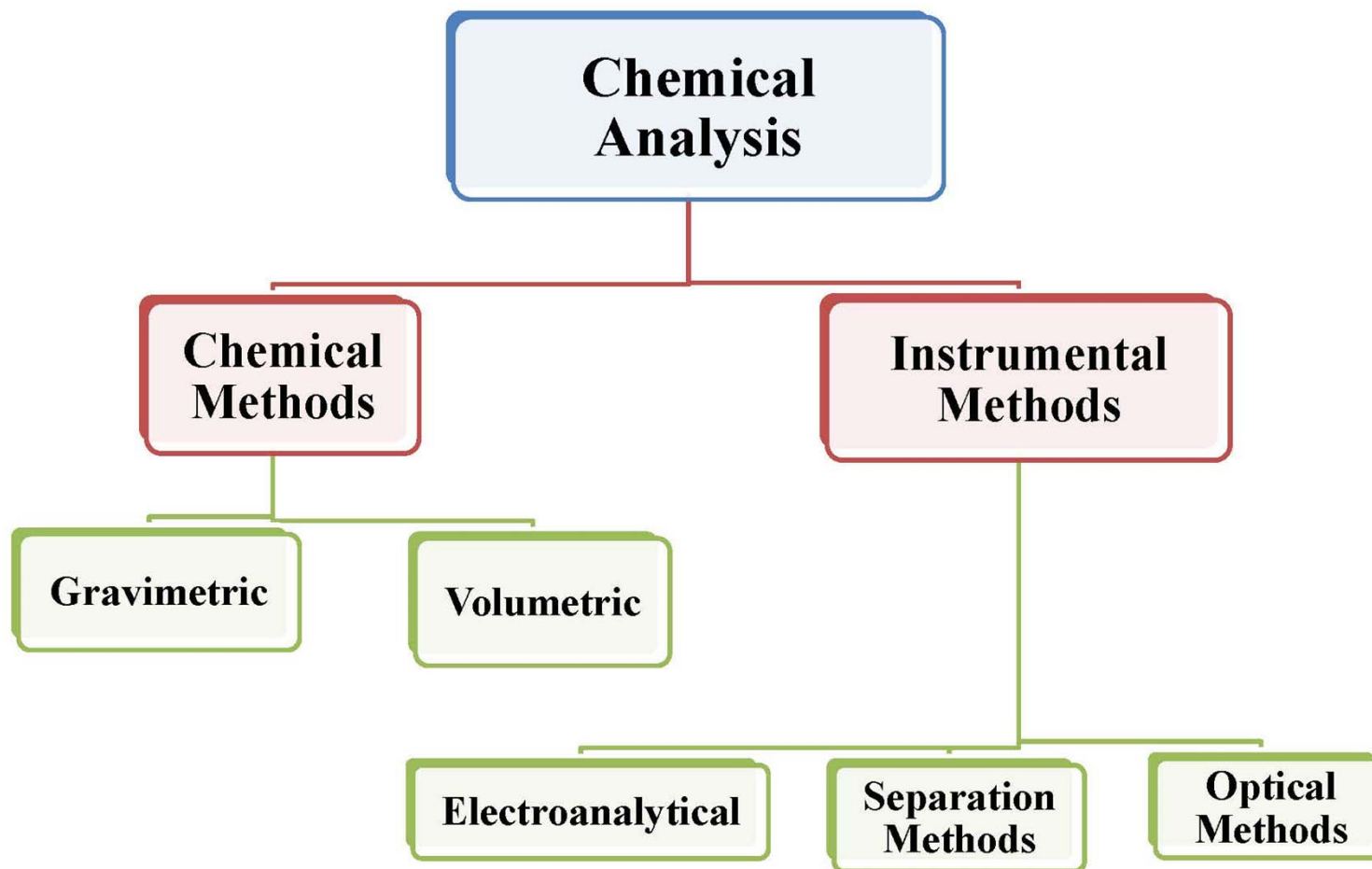


Visual endpoint

Instrumental analysis uses an instrument to convert a signal into a form that can be measured.

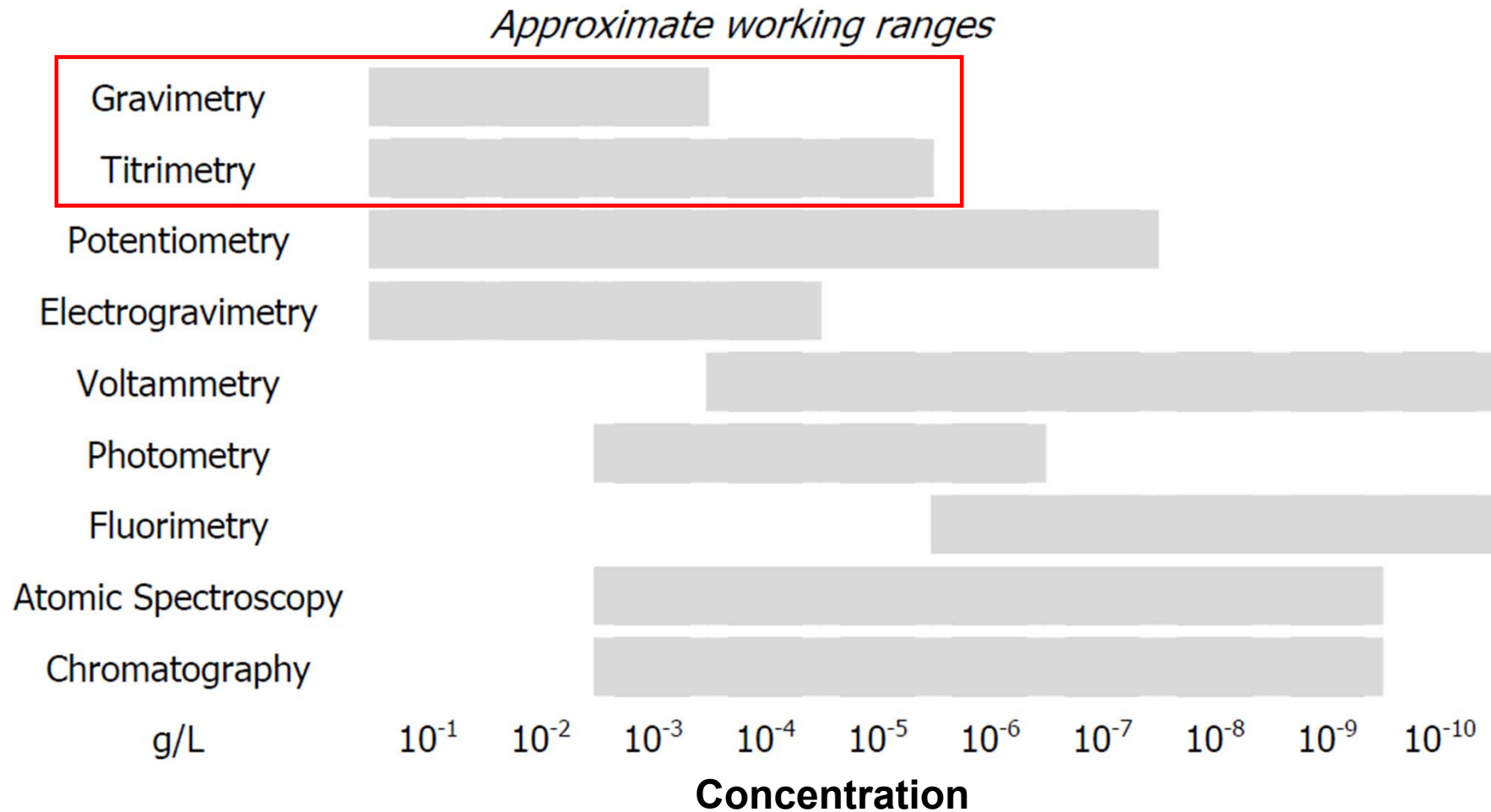
E.g. Emission of light, diffraction of radiation, electric potential





How do we decide on what method to use?

Comparison of analytical methods:



How accurate do you need to be?

Analysis of signal

- Seawater contains: Ca^{2+} , Mg^{2+} , Na^{+} , K^{+} , Sr^{2+}
- These contribute to its electrical conductivity
- Conductivity is NOT selective to an element
- We get conductivity of the SOLUTION



If we wanted to measure the concentration of each cation, then we would need to **separate** each cation signal and THEN quantify the signal.

Examples:

Spectroscopy: measure radiation emitted

Chromatography: separate ions based on size or polarity

Errors, Uncertainties & Statistics

Statistics

We need to understand what are we trying to measure and how **precise** our results need to be.

A medical example:

Your blood $[K^+]$ should be between 3.5 - 5.5 mmol/L.

This might seem a big variation, but is it?

Expressed as g/mL = 0.00014 – 0.00021 g/mL

If you titrated 10 mL of blood plasma with 0.1 M Ag^+ , the above range = **0.2 mL of titrant**.

How good are you at titrating? If you are wrong, the patient dies!

In reality you would not accept this risk and either choose a different method of analysis or perhaps change the concentration of the titrant.

Statistics are now helping us decide on the concentration of reagents.



A legal example:

The legal blood alcohol limit is 80 mg ethanol per 100 mL of blood.

The problem is that the precision of the measuring devices is around $\pm 10\%$.

79 mg/100 mL could read between 71 and 87 mg/100 mL.

In reality, the State won't prosecute until **90 mg/100 mL** is reached.

Otherwise people within the limit could be prosecuted due to the lack of precision of the equipment.

At the end of this section you *should* be able to:

- Look at a method of analysis and say whether it is appropriate or not.
- Decide on what level of precision is acceptable.

Measurement Uncertainty

Whenever we make a measurement of a physical quantity there will be some *uncertainty* in the result.

- Suppose you are weighing an object:
15 g **15.22 g** **15.1353 g**
- There is uncertainty associated with each of these measurements.
- The more decimal places we have, the **more** information we have and the ***less uncertainty*** there is.
- Do we always ***need*** as many decimal places as possible?

Significant Figures

- The number of significant values in a figure reflects the **certainty** we have in the value.
- You should only use a large number of significant figures when you have a high level of certainty about the value.

Revise the rules for significant figures!

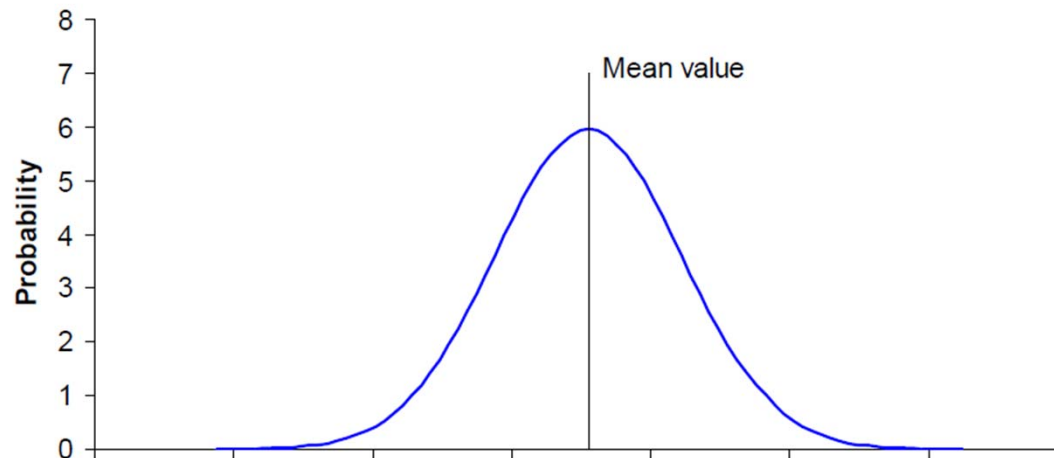
Types of experimental error: Accuracy and Precision

One thing is certain...you can never be certain

- Some forms of experimental error are unavoidable
- Some errors are avoidable
- Some errors are **determinate** (systematic) and affect the **accuracy** of a measurement.
- Some errors are **indeterminate** (random) and affect the **precision** of a measurement.

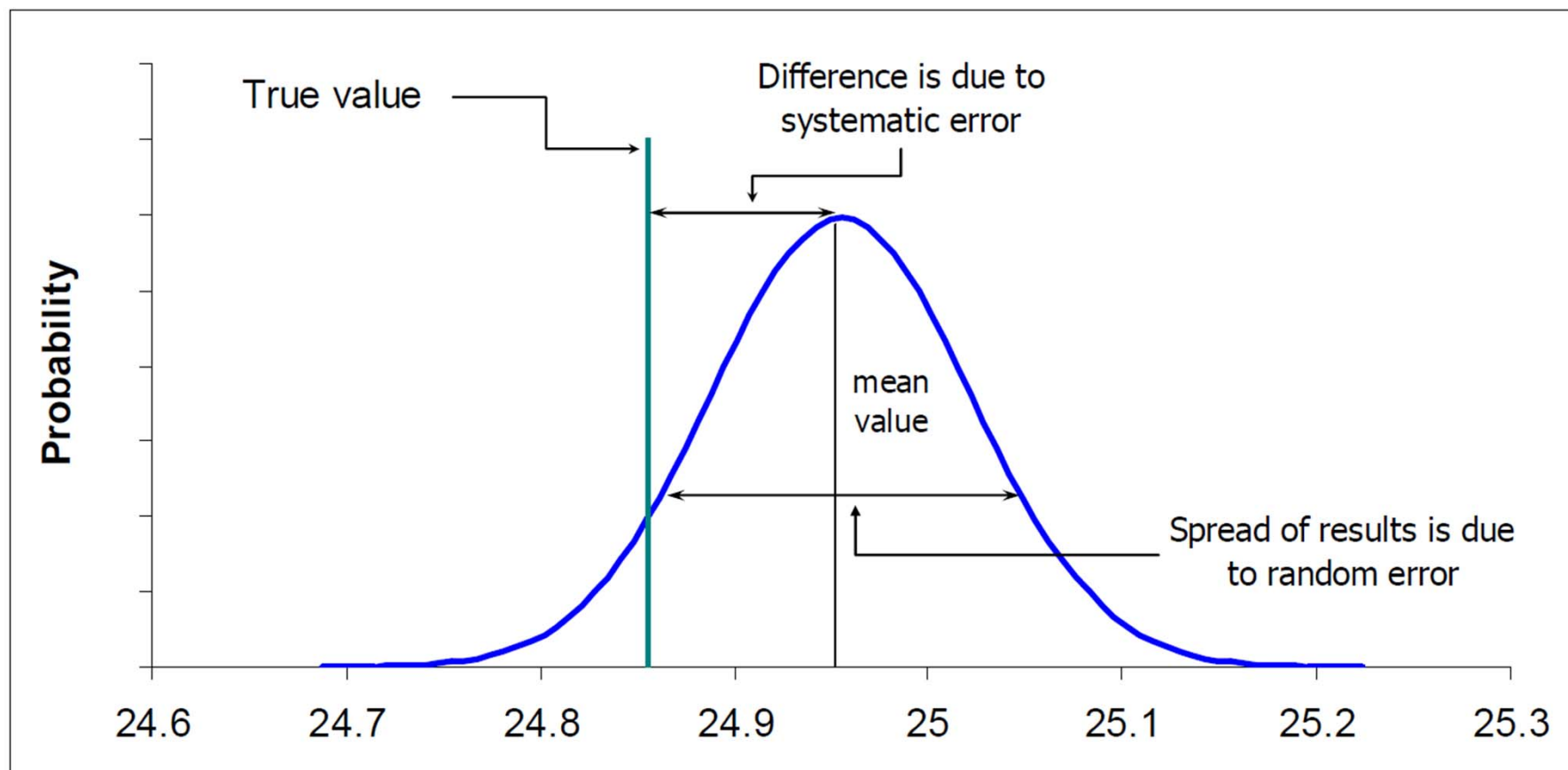
Indeterminate (Random) Errors

- These are caused by many **uncontrollable** variables and cannot be avoided.
- Can be positive or negative.
- Cannot be corrected for.
- If we repeat the measurement enough times the random errors will tend to cancel out when we calculate the *mean* or *average* value.



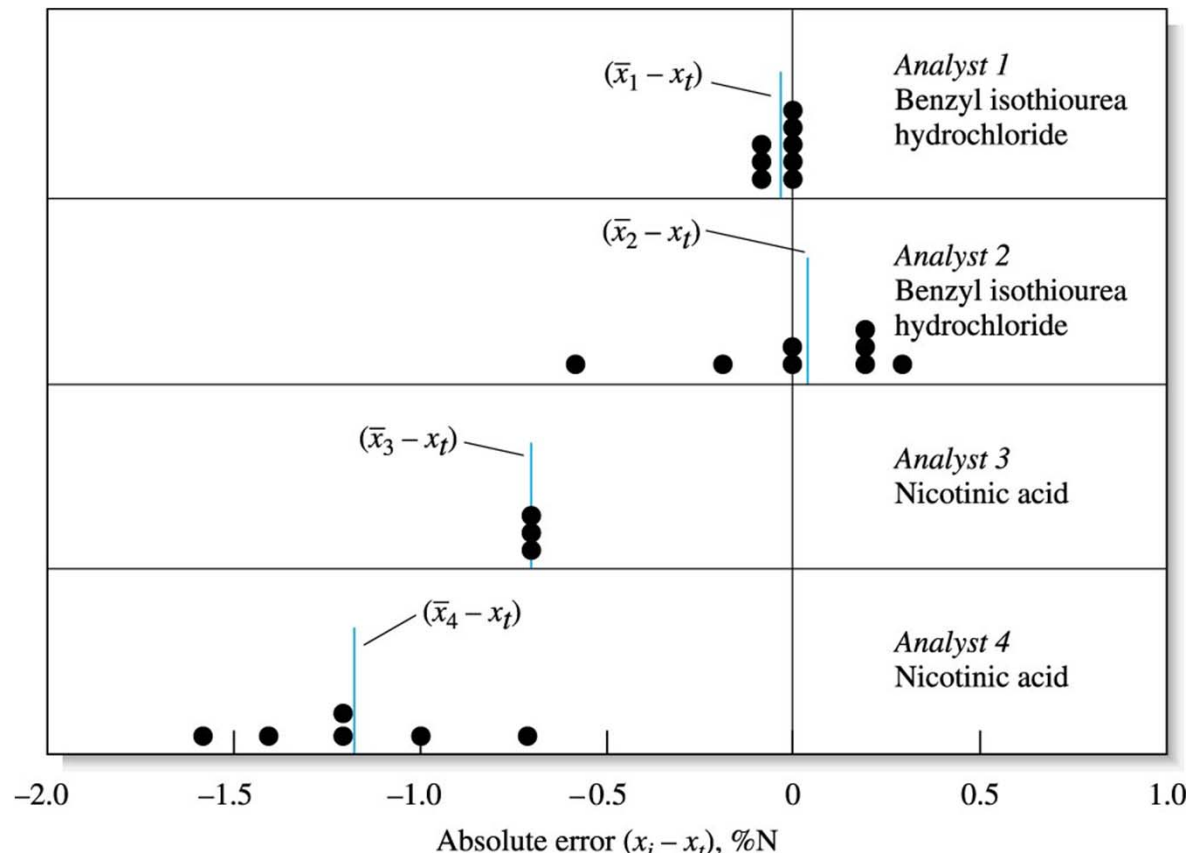
Determinate or Systematic Errors

- Reproducible under the same conditions in the same experiment.
- Can be detected and corrected for.
- Are always positive or always negative.
- Results in the mean value for the data differing from the true value.
- *Instrumental errors*: incorrect functioning of an instrument (faulty calibration, incorrect conditions)
- *Method errors*: non-ideal chemical or physical behaviour of reagents (incorrect choice of indicator)
- *Personal errors*: associated with the experimenter's own judgement.



Accuracy and Precision

- **Accuracy** – a measure of how close the experimental result is to the true value.
- **Precision** – a measure of the agreement within a set of replicate results



Errors and Statistics

Analysing a sample to get results is only the first step. We need to analyse our experimental results to obtain meaningful information.

Some basics:

Mean or average (\bar{X}): $\bar{x} = \frac{\sum x}{n}$

Standard deviation (s): $s = \sqrt{\frac{\sum_i \left(x_i - \bar{x} \right)^2}{n - 1}}$

This gives an indication of how good the measured precision is

Other ways of expressing precision...

- Relative standard deviation (RSD):

$$\text{RSD} = \frac{s}{\bar{x}}$$

- Percent RSD / coefficient of variation (%RSD):

$$\% \text{RSD} = \frac{s}{\bar{x}} \times 100$$

- In a scientific study, we usually infer information about a **population** from observations made on a **sample**.
- In reality therefore, we have a limited number of data.

The **sample mean** is given the symbol: \bar{x}

The **population mean** is given the symbol: μ

The **sample standard deviation**: s

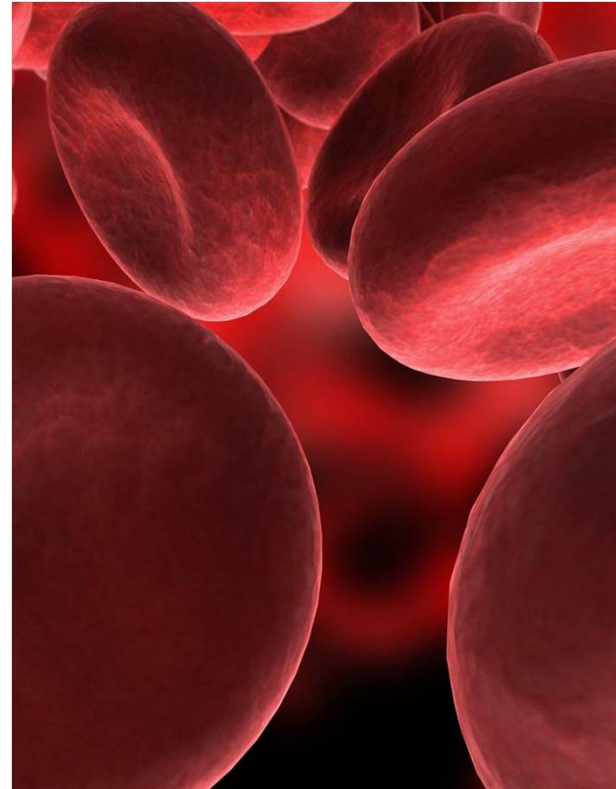
The **population standard deviation**: σ

$$\text{As } N \rightarrow \infty ; \quad \bar{x} \rightarrow \mu ; \quad s \rightarrow \sigma$$

Example:

Replicate results were obtained for the analysis of lead in blood. Calculate the **mean** and the **standard deviation** of this set of data.

Replicate	[Pb] / ppb
1	752
2	756
3	752
4	751
5	760

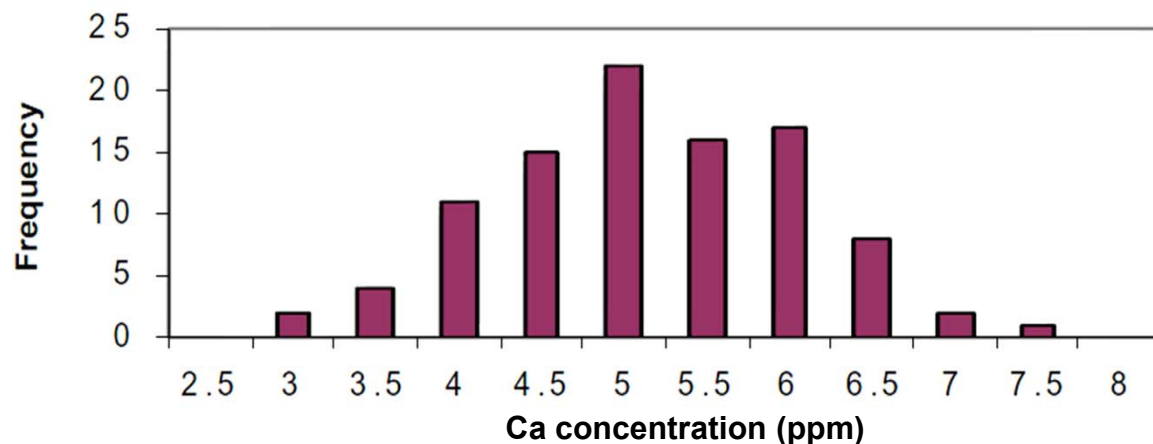


Another example:

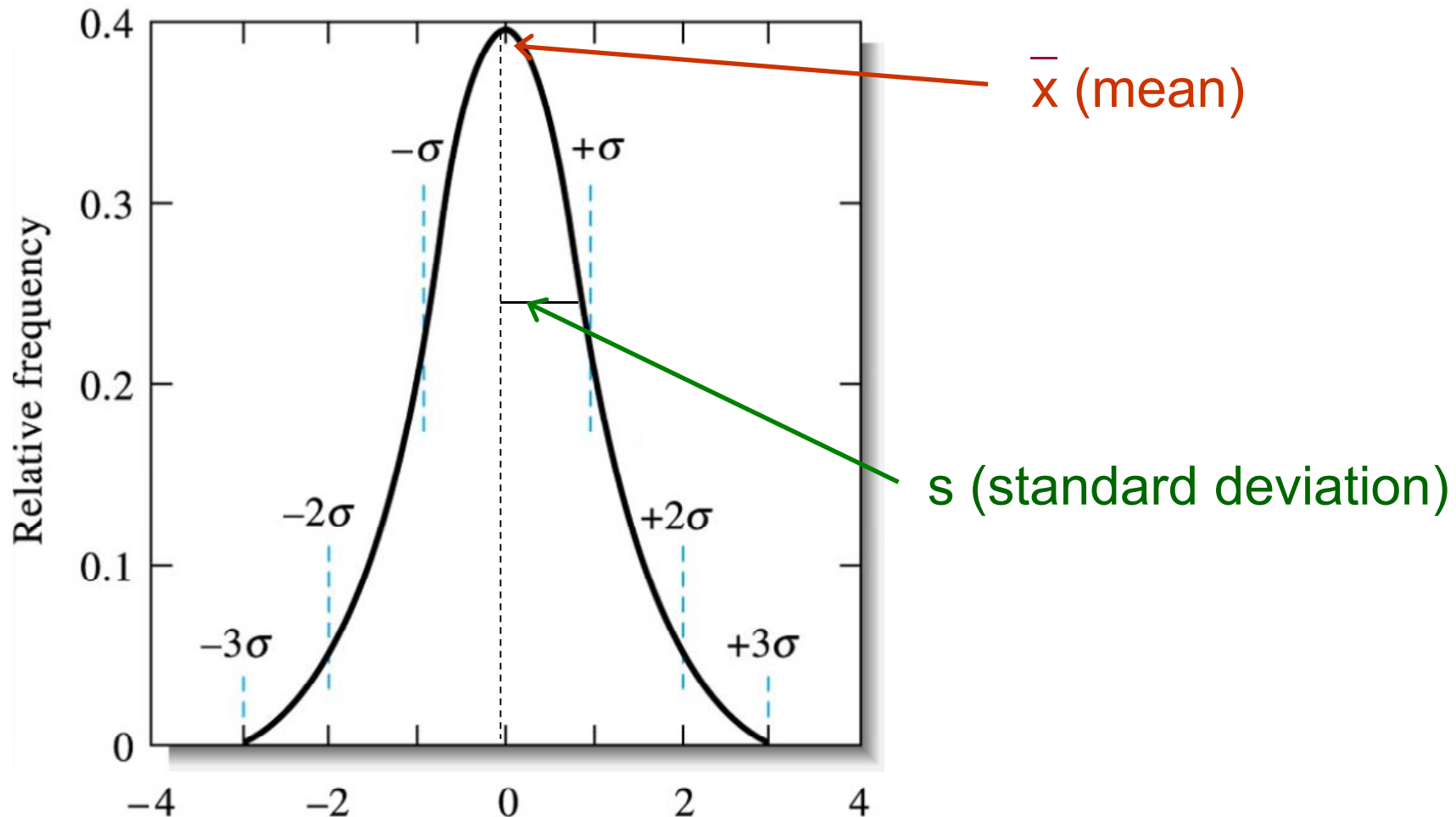
Let's look at the results obtained for the determination of Ca in mineral water (actual concentration = 5.00 ppm)

No. of readings	1	2	5	10	20	50	100
Mean conc (ppm)	3.0	3.8	4.5	4.6	4.7	5.0	5.0

If we use enough readings, the random errors cancel out and the average tends towards the “true mean”.

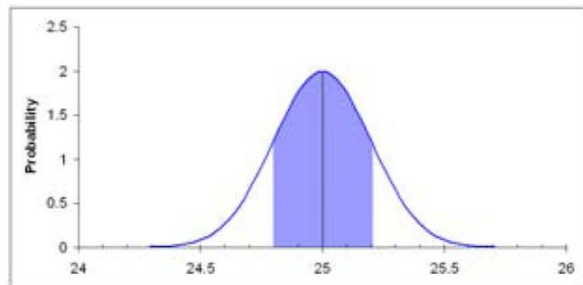


For a large number of experimental replicates the results approach an ideal smooth curve called the **Gaussian** or **Normal distribution** curve.

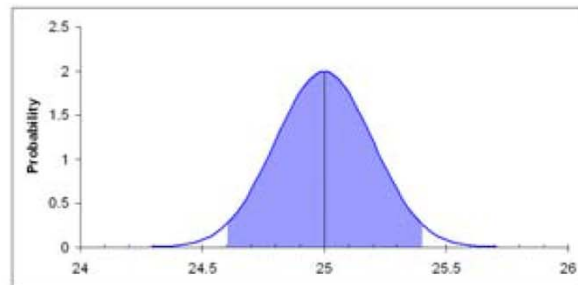


Gaussian Distributions

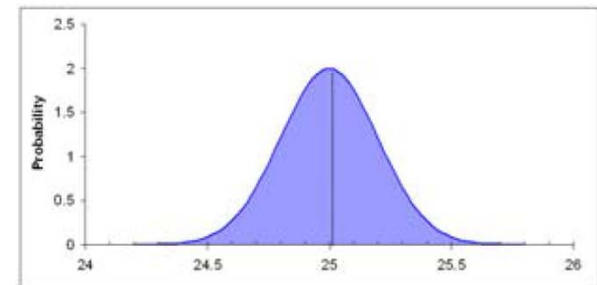
- We often assume that experimental results follow a Gaussian distribution.
- This allows us to make predictions about our data.
- E.g. we usually expect **95% of our readings** to be within a range **two standard deviations** either side of the true mean.
- We can estimate the range within which the true value has a 95% chance of being found.



$\pm 1 \sigma = 68.3\%$
of area under curve



$\pm 2 \sigma = 95.4\%$
of area under curve



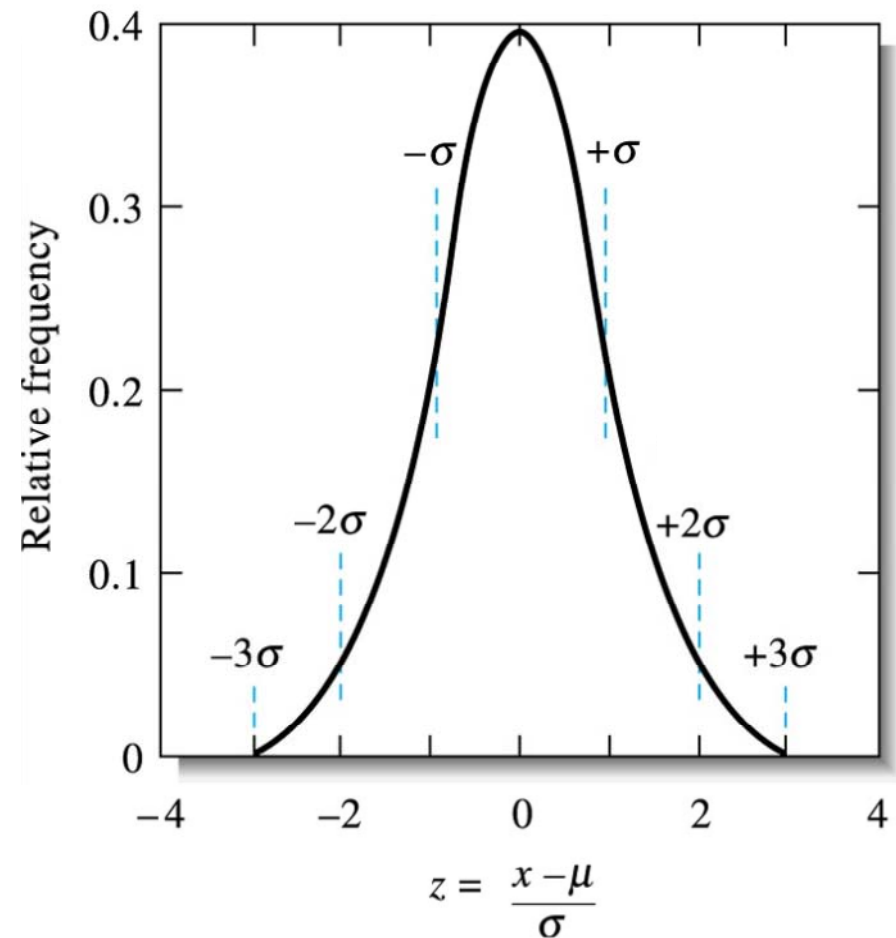
$\pm 3 \sigma = 99.7\%$
of area under curve

Confidence Intervals

- Confidence intervals define an interval around \bar{x} that probably contains μ

$$CI \text{ for } \mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

The value of z is
obtained from tables



Example: A sample of drinking water was analyzed for total chloro-hydrocarbons. From the data below, calculate the 95% confidence interval.

Number of Replicates with same value	ppb
5	17
7	20
6	16
8	14
5	19
6	22

Using Student's t test to find confidence intervals

When s is not a good estimate of σ
(small *no.* of replicates) then we use Student's t .

$$CI \text{ for } \mu = \bar{x} \pm \frac{ts}{\sqrt{n}}$$

The value of t depends on the desired confidence interval and the degrees of freedom.

Degrees of freedom = $n-1$

The number of values that are free to vary. If you have to take ten different courses to graduate, and only ten different courses are offered, then you have nine degrees of freedom. For nine semesters you will be able to choose which class to take; in the tenth semester, there will only be one class left to take - there is no choice, if you want to graduate.

If you want the t value for **6** replicates at the 95% CI

Values of t at Various Levels of Probability			
Degrees of Freedom	90%	95%	99%
1	6.31	12.7	63.7
2	2.92	4.30	9.92
3	2.35	3.18	5.84
4	2.13	2.78	4.60
5	2.02	2.57	4.03
6	1.94	2.45	3.71
7	1.90	2.36	3.50
8	1.86	2.31	3.36
9	1.83	2.26	3.25
10	1.81	2.23	3.17
15	1.75	2.13	2.95
20	1.73	2.09	2.84
40	1.68	2.02	2.70
60	1.67	2.00	2.62
∞	1.64	1.96	2.58

$$CI \text{ for } \mu = \bar{x} \pm \frac{ts}{\sqrt{n}}$$

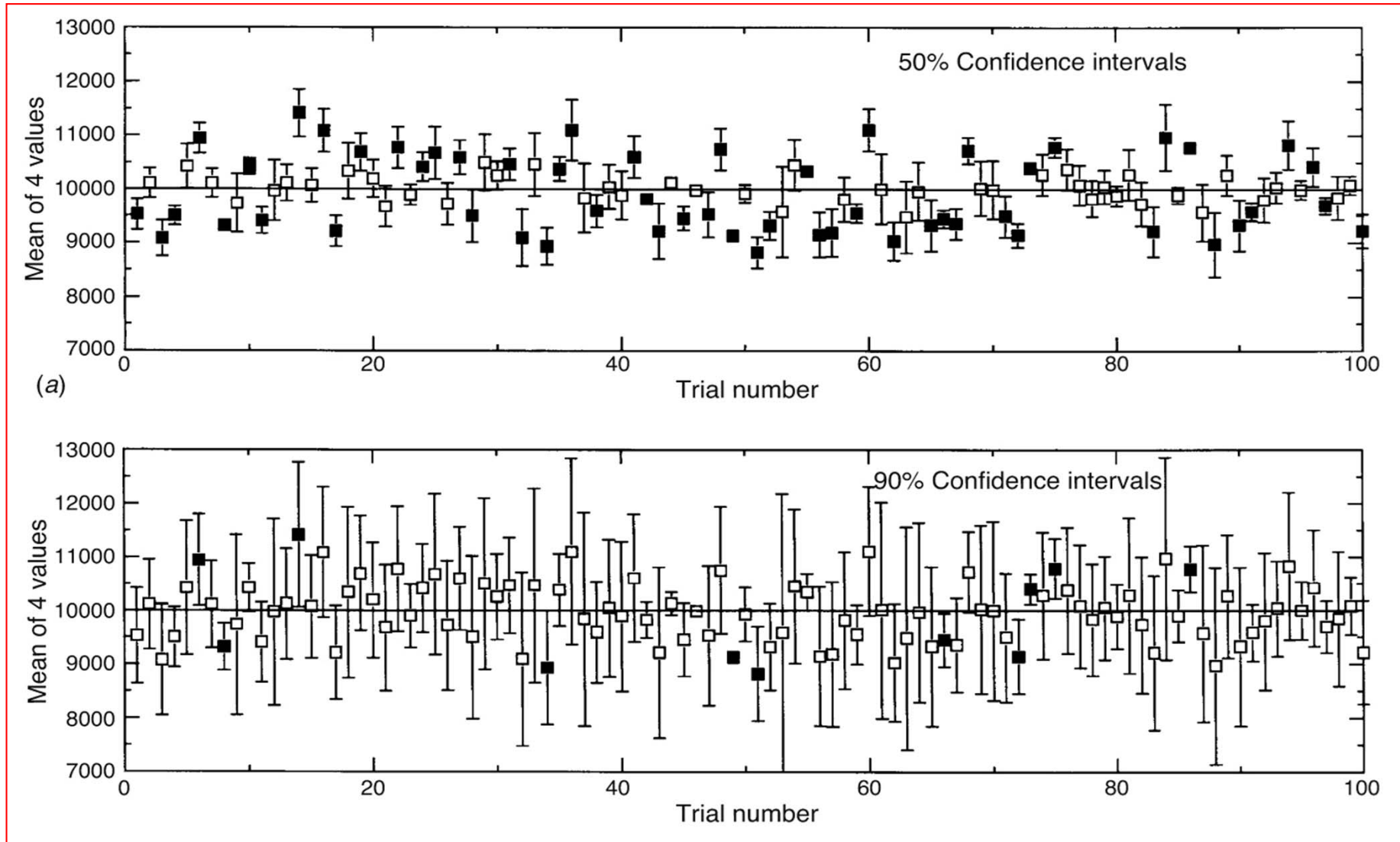


$$CI \text{ for } \mu = \bar{x} \pm \frac{2.57 \times s}{\sqrt{6}}$$

Example:

The mercury content in fish samples was determined as follows: 1.80, 1.58, 1.64, 1.49 ppm Hg. Calculate the 50% and 90% confidence intervals for the mercury content.

Confidence intervals: Experimental uncertainty



Comparing Experimental and True Mean Values using Student's t test

We can **assess the accuracy** of our data by comparing the sample mean to the population mean.

$$t_{calc} = \frac{|\bar{x} - \mu|}{s} \sqrt{n}$$

If $t_{calc} < t_{table} \rightarrow$ **NO** significant difference between the sample and the true value.

If $t_{calc} > t_{table} \rightarrow$ experimental results **ARE** significantly different from the true value.

Example:

A new procedure for the rapid determination of the %S in kerosenes was tested on a sample known to contain 0.123% S. The results are tabulated below.

0.112% S	0.118% S	0.115% S	0.119% S
----------	----------	----------	----------

Do the results agree with the accepted value at the 95% confidence level?

Comparison of two experimental means

- Comparing two means that each have an uncertainty.
- Do the two confidence intervals overlap?

For 2 sets of data with number of measurements n_1 , n_2 and means \bar{x}_1 , \bar{x}_2 :

$$t_{calc} = \frac{|\bar{x}_1 - \bar{x}_2|}{S_{pooled}} \sqrt{\frac{n_1 \times n_2}{n_1 + n_2}}$$

$$\text{d.f.} = n_1 + n_2 - 2$$

Where S_{pooled} = pooled std dev. from both sets of data

$$S_{pooled} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$$

If $t_{calc} > t_{table}$, the difference is significant

Comparing individual differences

We have two methods and a restricted number of replicates but several different samples.

Do the two methods produce different results?

We need to take account of variability between samples.

$$t_{\text{calc}} = \frac{\bar{d}}{s_d} \sqrt{n}$$

d = the mean difference

s_d = standard deviation of the differences

N = number of pairs of data

Example: Comparison of two methods for measuring cholesterol.

Plasma sample (d_i)	Cholesterol content (g/L)		Difference
	Method A	Method B	
1	1.46	1.42	0.04
2	2.22	2.38	-0.16
3	2.84	2.67	0.17
4	1.97	1.80	0.17
5	1.13	1.09	0.04
6	2.35	2.25	0.10
			$\bar{d} = 0.06$

Comparison of two standard deviations

What if we want to compare the **precision** obtained by two different labs or methods?

The F-test:

$$F_{\text{calc}} = \frac{s_1^2}{s_2^2}$$

- Always put the larger standard deviation in the numerator.
- We usually use the 95% confidence interval.
- If $F_{\text{calc}} < F_{\text{tab}}$, then there is no significant difference.

Table 4-5 Critical values of $F = s_1^2/s_2^2$ at 95% confidence level

Degrees of freedom for s_2	Degrees of freedom for s_1													
	2	3	4	5	6	7	8	9	10	12	15	20	30	∞
2	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5
3	9.55	9.28	9.12	9.01	8.94	8.89	8.84	8.81	8.79	8.74	8.70	8.66	8.62	8.53
4	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.63
5	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.36
6	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.67
7	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.58	3.51	3.44	3.38	3.23
8	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	2.93
9	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.71
10	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.84	2.77	2.70	2.54
11	3.98	3.59	3.36	3.20	3.10	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.57	2.40
12	3.88	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.47	2.30
13	3.81	3.41	3.18	3.02	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.38	2.21
14	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.31	2.13
15	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.07
16	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.19	2.01
17	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.15	1.96
18	3.56	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.11	1.92
19	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.07	1.88
20	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.84
30	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.62
∞	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.00

Example: As part of a inter-laboratory study, a sample of river water was sent to two laboratories for phosphate analysis (in ppb).

Lab A	20.7	27.5	30.4	23.9	21.7	24.1	24.8	28.9
Lab B	20.9	21.4	24.9	20.5	19.7	26.3	22.4	20.2

Is there a significant difference in precision between the two labs?

Detection of Gross Errors (outliers)

Let's consider the following dataset:

20.5	20.0	20.4	20.5	0.25	20.6	20.3
------	------	------	------	------	------	------

It is not always so easy to recognise an outlier.

10.3%	10.5	10.3%	9.8%	10.6%	10.4%	10.4%
-------	------	-------	------	-------	-------	-------

Is 9.8% an outlier? We need an objective test.

The Q-test:

$$Q_{calc} = \frac{| \text{suspected value} - \text{nearest neighbour} |}{\text{range of all values}}$$

In the above case, $Q_c = 0.5/0.8 = 0.625$

For Q_{table} you look at number of observations **NOT** degrees of freedom!

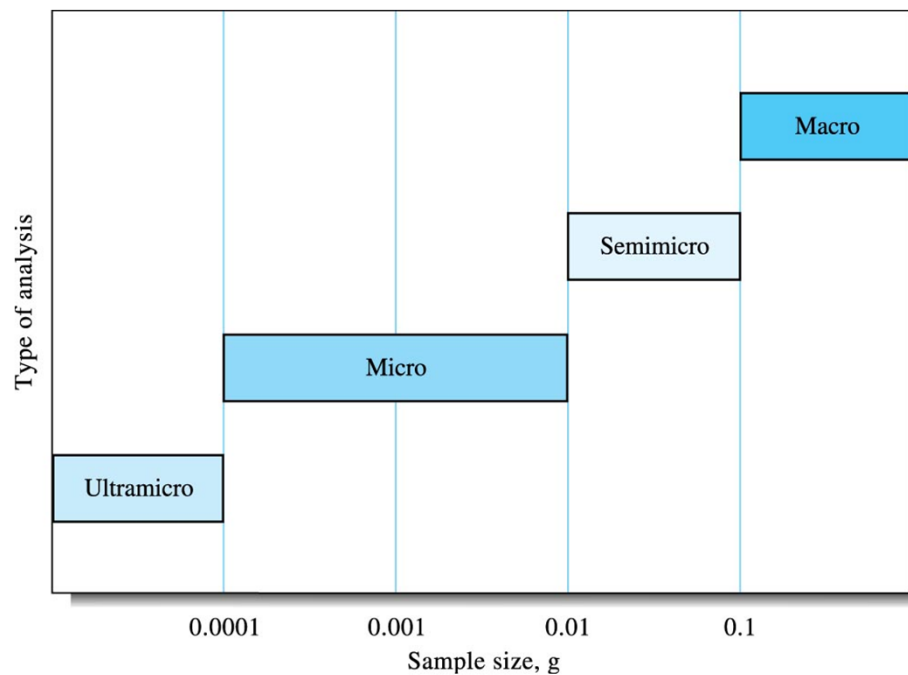
Critical Values for Rejection Quotient, Q			
Number of Observations	Q_{crit} (Reject if $Q > Q_{\text{crit}}$)		
	90% Confidence	95% Confidence	99% Confidence
3	0.941	0.970	0.994
4	0.765	0.829	0.926
5	0.642	0.710	0.821
6	0.560	0.625	0.740
7	0.507	0.568	0.680
8	0.468	0.526	0.634
9	0.437	0.493	0.598
10	0.412	0.466	0.568

The Q-test is not recommended for small sets of data.

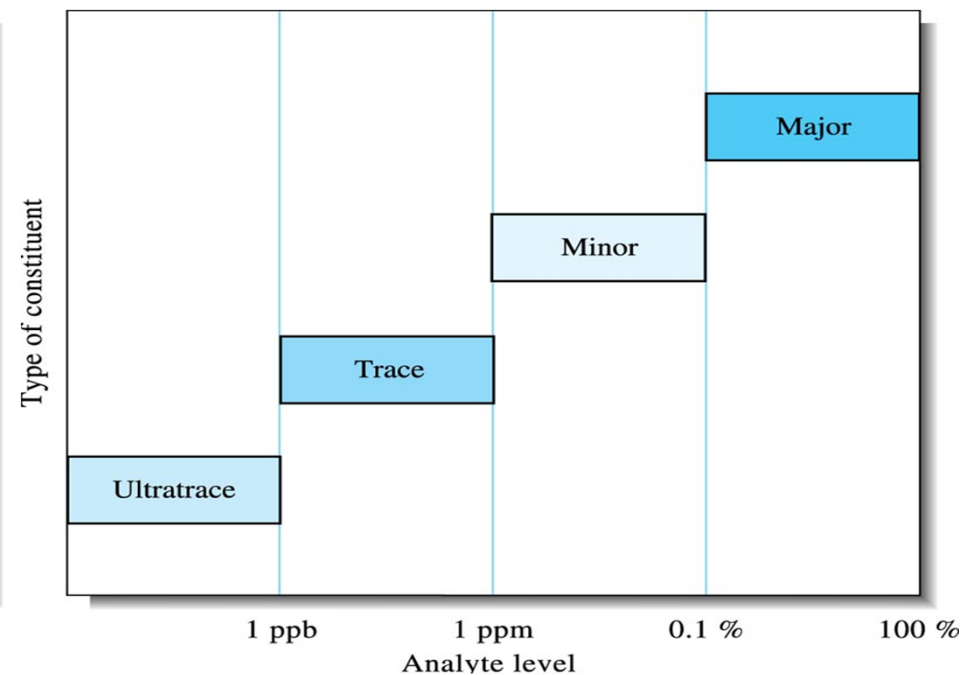
Rather obtain more data!

What precision can you expect?

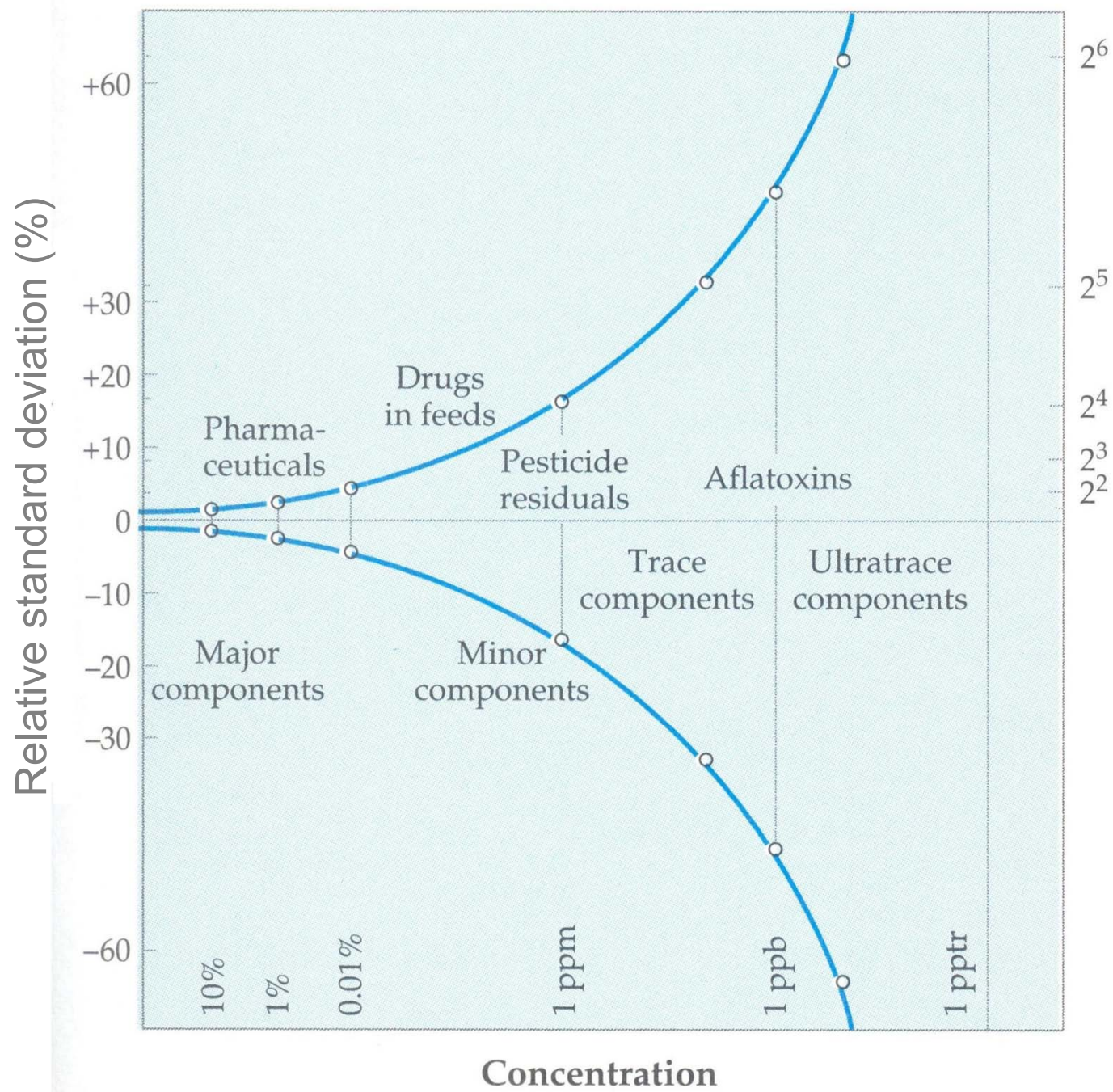
It can be difficult to analyse small samples and trace quantities.



© 2004 Thomson - Brooks/Cole



© 2004 Thomson - Brooks/Cole



How can you be sure about the accuracy of your results?

Certified reference materials (CRM)

- These are commercially available “samples” that have analysed by a number of laboratories using different methods.
- Have a [known composition](#).

<u>Recommended Values</u> <i>(Based on dry weight)</i>			
Analyte	Recommended Value [mg/kg]	95% Confidence Interval [mg/kg]	N*
As	23.6	22.9 – 24.3	47
Cd	0.73	0.68 – 0.78	63
Co	13.7	13.0 – 14.4	50
Cr	84	80 – 88	63
Cu	47.7	46.5 – 48.9	80
Fe	37400	36700 – 38100	64
Hg	0.81	0.77 – 0.85	60

You should be able to:

- Use the Student's t test:
 - to find confidence limits
 - to compare an experimental mean with a true mean
 - to test the accuracy of your results
- Compare two experimental means
- Compare two different analytical methods
- Compare the precision of two different sets of data (F-test)
- Identify gross errors (Q-test)