Chem 115

Instrumental Analysis and Bioanalytical Chemistry

Lecture 2: Concepts and analysis

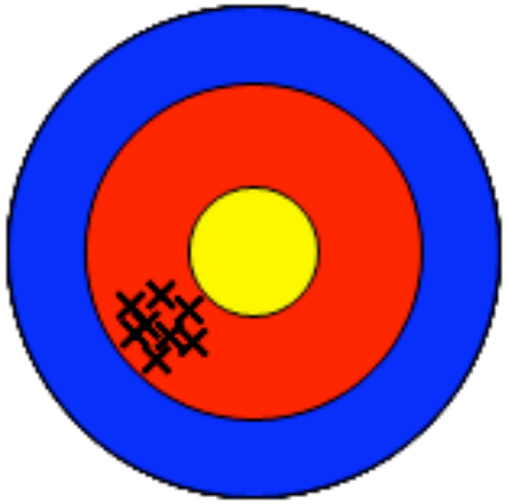
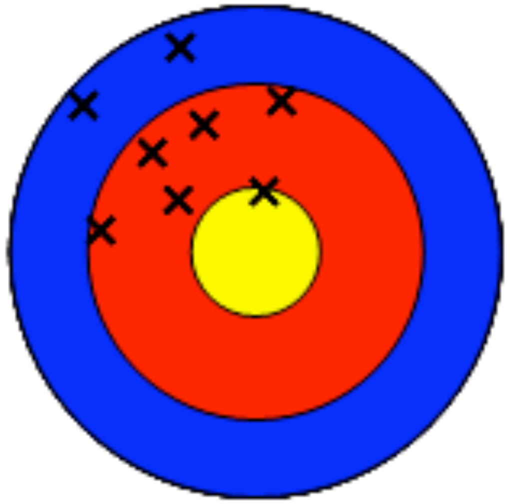
What’s in this lecture?

- Significant figures
- Statistical analysis

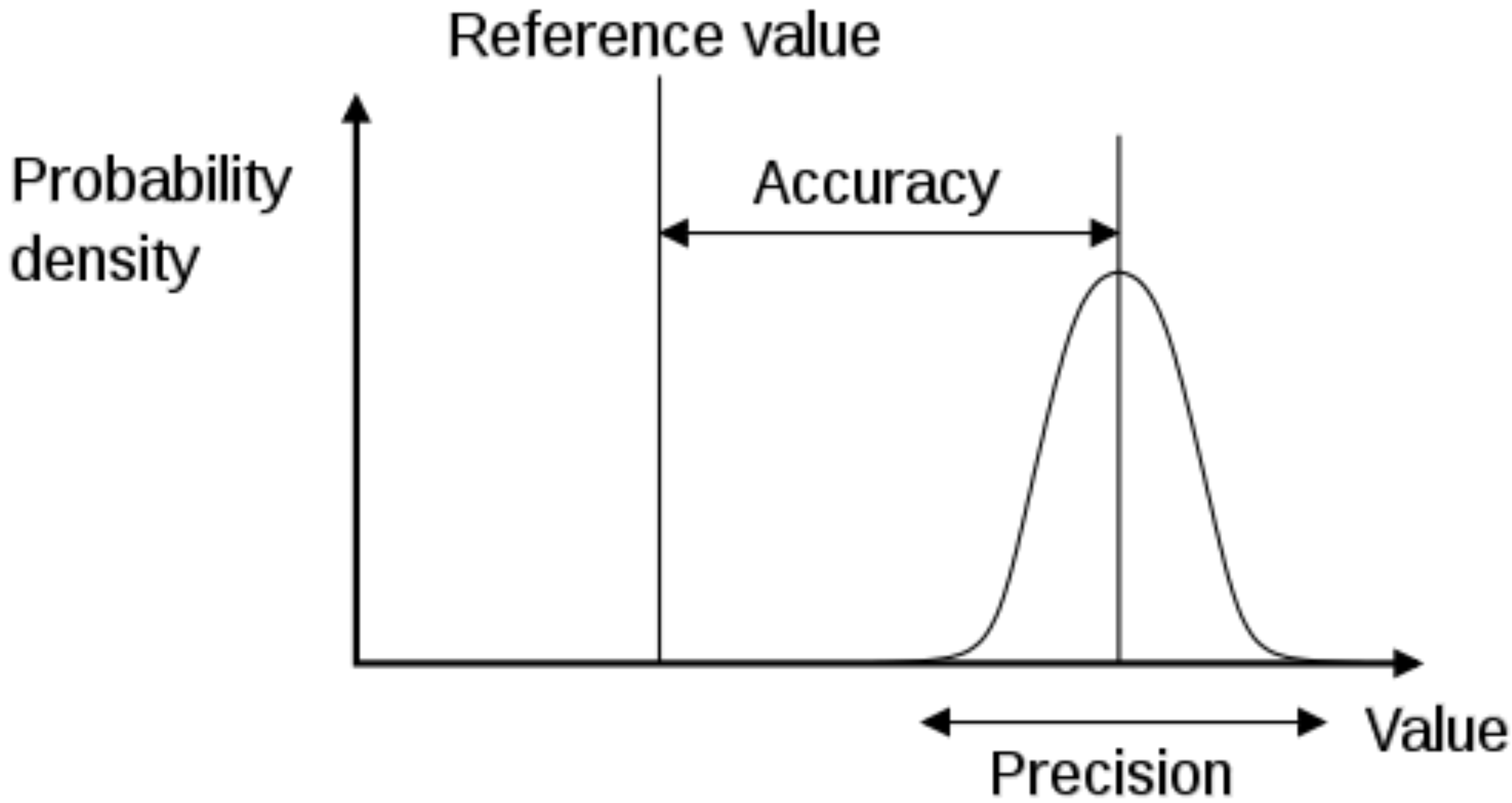
Rules for significant figures

1. Report all significant figures, such that the only the last digit is uncertain.
2. Reject all other uncertain digits, rounding in the process.
3. For addition and subtraction, round off after the largest absolute uncertainty.
4. For multiplication and division, the number of sig. figs. is determined by the value with the smallest number of sig. figs.
5. For logarithms, the mantissa has the same number of sig. figs. as the original number.
6. For a series of operations, keep track of the sig. figs, and round at the end.

Accuracy vs. Precision

	Accurate	Inaccurate (systematic error)
Precise		
Imprecise (reproducibility error)		

Accuracy vs. Precision



Accuracy vs. Precision

	% Analyte			
	Analyst 1	Analyst 2	Analyst 3	Analyst 4
Sample 1	10.0	8.1	13.0	13.0
Sample 2	10.2	8.0	10.2	8.0
Sample 3	10.0	8.3	10.3	7.9
Sample 4	10.2	8.2	11.1	12.4
Sample 5	10.1	8.0	13.1	10.3
Sample 6	10.1	8.0	9.3	9.0
Mean	10.1	8.1	11.2	10.1
Error	0.0	-2.0	1.1	0.0
Std. Dev.	0.089	0.13	1.57	2.2

True value = 10.1 ± 0.2%

Types of errors

- Determinate or systematic
  - Faults with procedure or instrument.
  - All measurements shifted in 1 direction.
  - Constant vs. proportional.
- Indeterminate or random
  - Due to limitations with instruments and/or noise.
  - Sources cannot be prevented, corrected, avoided, or even identified.
  - Random, can be taken into account.

Systematic errors

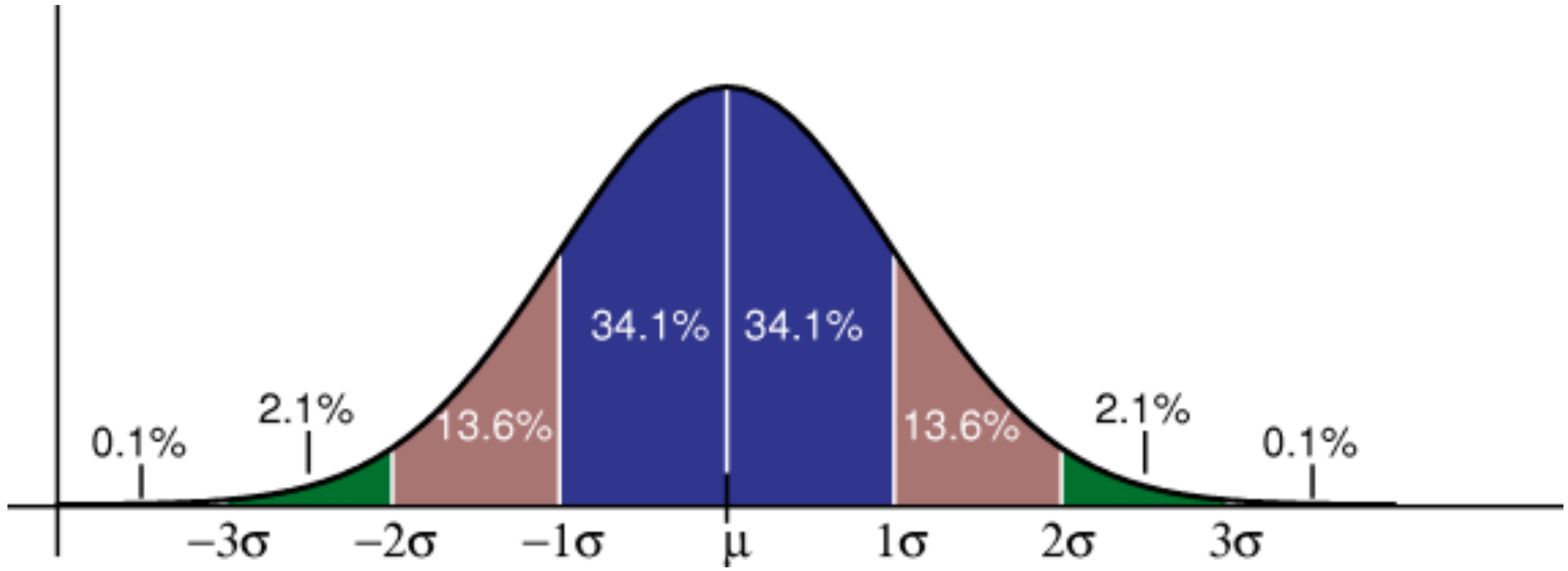
- Constant errors
  - Measurement off by same absolute amount.
  - Generally due to instrument not zeroed correctly.
- Proportional errors
  - Measurement off by same proportional amount.
  - Generally due to poor calibration over measurement range.
- Drift - measurement changes over time

Sources of systematic errors

- Analyst error
- Reagents
- Instrumentation
- Method
- Contamination

Random errors

- Always present in measurements.
- Overcome by taking many measurements.
- Approximately follow a normal distribution.



Statistics

- True value: T or x_t
- Observed value: x_i
- Sample mean: $\overline{x} = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{x_1 + x_2 + x_3 + \cdots + x_N}{N}$
- Error: $x_i - x_t$
- Absolute error: $|x_i - x_t|$
- Relative error: $(x_i - x_t)/x_t$
- Relative absolute error: $|x_i - x_t|/x_t$
- Sample standard deviation: $s = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \overline{x})^2}{N-1}}$

Sample mean vs. population mean

$$\overline{x} = \sum_{i=1}^{N} \frac{x_i}{N} \qquad \mu = \lim_{N \to \infty} \sum_{i=1}^{N} \frac{x_i}{N}$$

In reality, $\overline{x} \rightarrow \mu$ rapidly after 20-30 measurements

Sample standard deviation vs. population standard deviation

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

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

- 68.3% of values lie within 1σ of μ
- 95.5% of values lie within 2σ of μ
- 99.7% of values lie within 3σ of μ

Student's t statistic

Measure of how confident we are the population mean lies within a certain range.

$$\mu = \overline{x} \pm \frac{ts}{\sqrt{N}}$$

Student's t table

Degrees of Freedom	80%	90%	95%	99%
1	3.078	6.314	12.706	63.657
2	1.886	2.92	4.303	9.925
3	1.638	2.353	3.182	5.841
4	1.533	2.132	2.776	4.604
5	1.476	2.015	2.571	4.032
6	1.44	1.943	2.447	3.707
7	1.415	1.895	2.365	3.5
8	1.397	1.86	2.306	3.355
9	1.383	1.833	2.262	3.25
10	1.372	1.812	2.228	3.169
11	1.363	1.796	2.201	3.106
12	1.356	1.782	2.179	3.055
13	1.35	1.771	2.16	3.012
14	1.345	1.761	2.145	2.977
15	1.341	1.753	2.131	2.947
16	1.337	1.746	2.12	2.921
17	1.333	1.74	2.11	2.898
18	1.33	1.734	2.101	2.878
19	1.328	1.729	2.093	2.861
20	1.325	1.725	2.086	2.845
21	1.323	1.721	2.08	2.831
22	1.321	1.717	2.074	2.819
23	1.319	1.714	2.069	2.807
24	1.318	1.711	2.064	2.797
25	1.316	1.708	2.06	2.787
26	1.315	1.706	2.056	2.779
27	1.314	1.703	2.052	2.771
28	1.313	1.701	2.048	2.763
29	1.311	1.699	2.045	2.756
30	1.31	1.697	2.042	2.75
infinity	1.282	1.645	1.96	2.576

Degrees of freedom = N-1

A word of caution...

Two sided

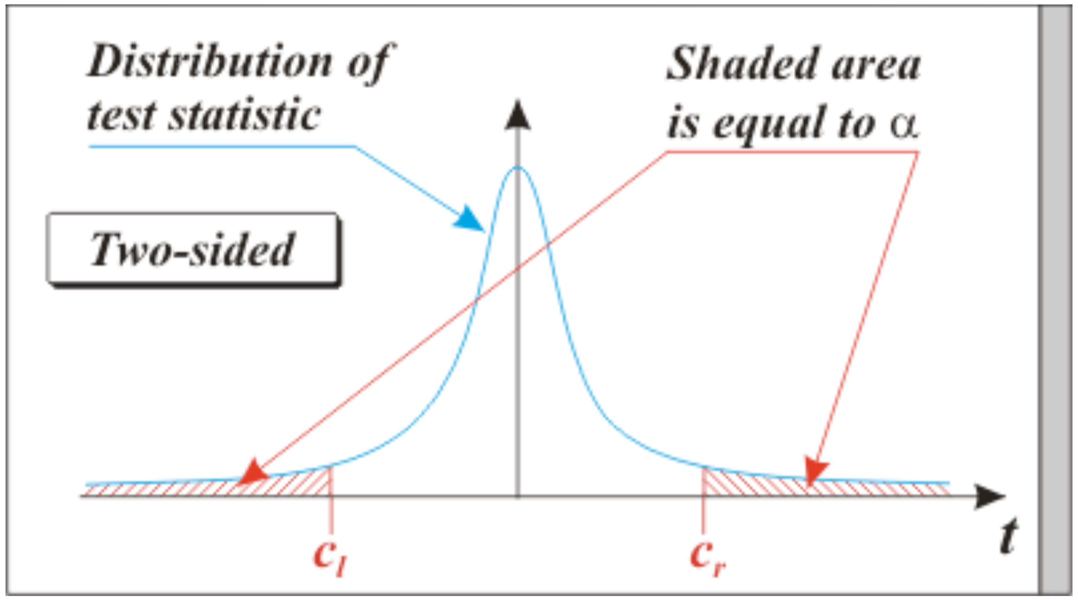
Degrees of Freedom	80%	90%	95%	99%
1	3.078	6.314	12.706	63.657
2	1.886	2.92	4.303	9.925
3	1.638	2.353	3.182	5.841
4	1.533	2.132	2.776	4.604
5	1.476	2.015	2.571	4.032
6	1.44	1.943	2.447	3.707
7	1.415	1.895	2.365	3.5
8	1.397	1.86	2.306	3.355
9	1.383	1.833	2.262	3.25
10	1.372	1.812	2.228	3.169
11	1.363	1.796	2.201	3.106
12	1.356	1.782	2.179	3.055
13	1.35	1.771	2.16	3.012
14	1.345	1.761	2.145	2.977
15	1.341	1.753	2.131	2.947
16	1.337	1.746	2.12	2.921
17	1.333	1.74	2.11	2.898
18	1.33	1.734	2.101	2.878
19	1.328	1.729	2.093	2.861
20	1.325	1.725	2.086	2.845
21	1.323	1.721	2.08	2.831
22	1.321	1.717	2.074	2.819
23	1.319	1.714	2.069	2.807
24	1.318	1.711	2.064	2.797
25	1.316	1.708	2.06	2.787
26	1.315	1.706	2.056	2.779
27	1.314	1.703	2.052	2.771
28	1.313	1.701	2.048	2.763
29	1.311	1.699	2.045	2.756
30	1.31	1.697	2.042	2.75
infinity	1.282	1.645	1.96	2.576

One sided

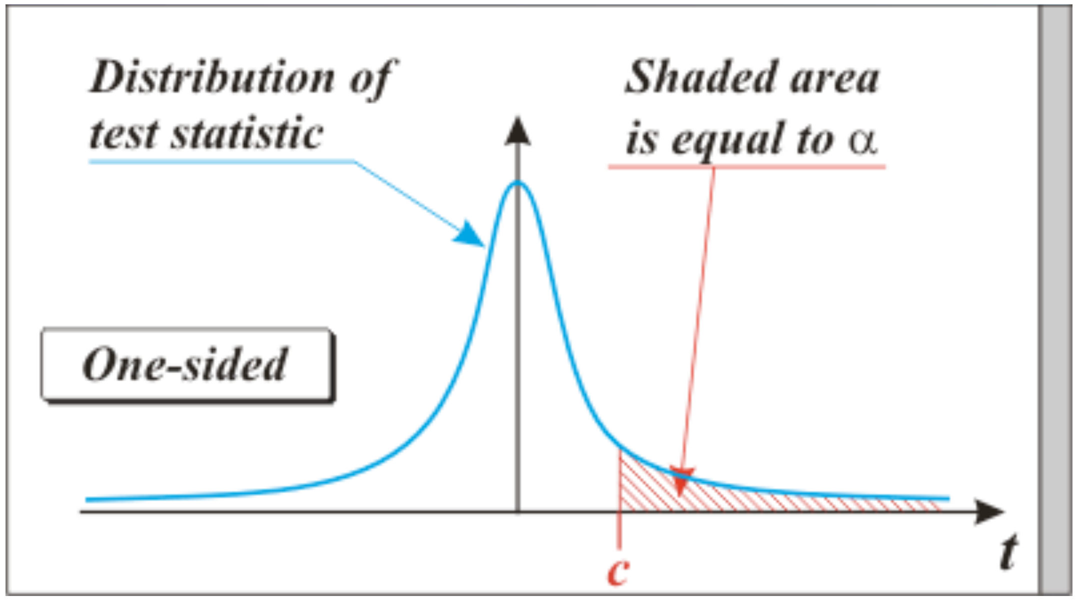
Degrees of Freedom	80%	90%	95%	99%
1	1.376	3.078	6.314	31.82
2	1.061	1.886	2.92	6.965
3	0.978	1.638	2.353	4.541
4	0.941	1.533	2.132	3.747
5	0.92	1.476	2.015	3.365
6	0.906	1.44	1.943	3.143
7	0.896	1.415	1.895	2.998
8	0.889	1.397	1.86	2.896
9	0.883	1.383	1.833	2.821
10	0.879	1.372	1.812	2.764
11	0.876	1.363	1.796	2.718
12	0.873	1.356	1.782	2.681
13	0.87	1.35	1.771	2.65
14	0.868	1.345	1.761	2.624
15	0.866	1.341	1.753	2.602
16	0.865	1.337	1.746	2.583
17	0.863	1.333	1.74	2.567
18	0.862	1.33	1.734	2.552
19	0.861	1.328	1.729	2.539
20	0.86	1.325	1.725	2.528
21	0.859	1.323	1.721	2.518
22	0.858	1.321	1.717	2.508
23	0.858	1.319	1.714	2.5
24	0.857	1.318	1.711	2.492
25	0.856	1.316	1.708	2.485
26	0.856	1.315	1.706	2.479
27	0.855	1.314	1.703	2.473
28	0.855	1.313	1.701	2.467
29	0.854	1.311	1.699	2.462
30	0.854	1.31	1.697	2.457
Infinity	0.842	1.282	1.645	2.326

A word of caution...

Two sided



One sided



Comparing two means

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}}$$

Comparing precision

$$F = \frac{\sigma_1^2}{\sigma_2^2}$$

F values at 95% C.L.

Numerator D.F.

Denominator D.F.

	1	2	3	4	5	7	10
1	161.45	199.5	215.71	224.58	230.16	236.77	241.88
2	18.513	19	19.164	19.247	19.296	19.353	19.396
3	10.128	9.5522	9.2766	9.1172	9.0135	8.8867	8.7855
4	7.7086	6.9443	6.5915	6.3882	6.256	6.0942	5.9644
5	6.6078	5.7862	5.4095	5.1922	5.0504	4.8759	4.7351
7	5.5914	4.7375	4.3469	4.1202	3.9715	3.7871	3.6366
10	4.9645	4.1028	3.7082	3.478	3.3259	3.1354	2.9782

Rejecting results

1. Carefully reexamine all data, procedures and observations to determine if a gross error could be the cause.
2. Estimate the precision to check if the outlier is questionable.
3. Repeat the analysis if sample and time are available.
4. Apply Q test to determine if the outlier can be rejected on statistical grounds.
5. Consider reporting median instead of mean.

Q-test

$$Q_{exp} = \frac{|x_n - x_o|}{x_{max} - x_{min}}$$

If $Q_{exp} > Q_{crit}$, reject x_o

# of observations	90% conf.	95% conf	99% conf
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

Example Time...

	% Analyte			
	Analyst 1	Analyst 2	Analyst 3	Analyst 4
Sample 1	10.0	8.1	13.0	13.0
Sample 2	10.2	8.0	10.2	8.0
Sample 3	10.0	8.3	10.3	7.9
Sample 4	10.2	8.2	11.1	12.4
Sample 5	10.1	8.0	13.1	10.3
Sample 6	10.1	8.0	9.3	9.0
Mean	10.1	8.1	11.2	10.1
Std. Dev.	0.089	0.13	1.57	2.2