Testing a Hypothesis (Significance Tests)

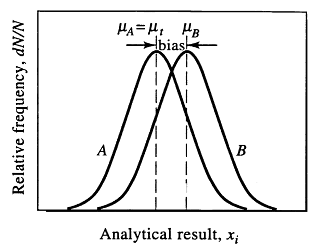
Carry out measurements on an accurately known standard.

Experimental value is different from the true value.

Is the difference due to a systematic error (bias) in the method - or simply to random error?

Assume that there is *no* bias (**NULL HYPOTHESIS**), and calculate the probability that the experimental error is due to random errors.

Figure shows (A) the curve for the true value ($\mu_A = \mu_t$) and (B) the experimental curve (μ_B)



Bias = μ_B - μ_A = μ_B - x_t.

Test for bias by comparing $\bar{x} - x_t$ with the difference caused by random error

Remember confidence limit for μ (assumed to be x_t, i.e. *assume no bias*) is given by:

CL for $\mu = \bar{x} \pm \frac{ts}{\sqrt{N}}$

$\therefore$ at desired confidence level, random errors can lead to:

$\bar{x} - x_t = \pm \frac{ts}{\sqrt{N}}$

$\therefore$ if $\bar{x} - x_t > \left| \frac{ts}{\sqrt{N}} \right|$, then at the desired confidence level bias (systematic error) is likely (and vice versa).

Detection of Systematic Error (Bias)

A standard material known to contain 38.9% Hg was analysed by atomic absorption spectroscopy. The results were 38.9%, 37.4% and 37.1%. At the 95% confidence level, is there any evidence for a systematic error in the method?

$\bar{x} = 37.8\% \qquad \therefore \bar{x} - x_t = -1.1\%$

$\sum x_i = 113.4 \qquad \sum x_i^2 = 4208.30$

$\therefore s = \sqrt{\frac{4208.30 - (113.4)^2/3}{2}} = 0.943\%$

Assume null hypothesis (no bias). Only reject this if

$$t_{calc} = (\bar{x} - \mu)\frac{\sqrt{N}}{s} \qquad \text{Reject if } t_{calc} > t_{table}$$

But t (from Table) = 4.30, s (calc. above) = 0.943% and N = 3

$$2.02 < 4.30$$

Therefore the null hypothesis is maintained, and there is no evidence for systematic error at the 95% confidence level.

Are two sets of measurements significantly different?

Suppose two samples are analysed under identical conditions.

Sample 1 $\rightarrow \bar{x}_1$ from N_1 replicate analyses

Sample 2 $\rightarrow \bar{x}_2$ from N_2 replicate analyses

Are these significantly different?

Using definition of pooled standard deviation, the equation on the last overhead can be re-arranged:

$$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled}\sqrt{\frac{N_1 + N_2}{N_1 N_2}}}$$

Again reject if $t_{calc} > t_{table}$

Detection of Gross Errors

A set of results may contain an outlying result - out of line with the others.
Should it be retained or rejected?
There is no universal criterion for deciding this.
One rule that can give guidance is the ***Q test***.

The parameter Q_{calc} is defined as follows:

$$Q_{calc} = \left| \frac{outlier - nearest\_neighbor}{total\_spread} \right|$$

Again reject if $Q_{calc} > Q_{table}$

No. of observations	90%	95%	99% confidencelevel
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

Rejection of outlier recommended if $Q_{calc} > Q_{table}$ for the desired confidence level.

Note:1. The higher the confidence level, the less likely is rejection to be recommended.

2. Rejection of outliers can have a marked effect on mean and standard deviation, esp. when there are only a few data points. *Always try to obtain more data*.

Q Test for Rejection of Outliers

The following values were obtained for the concentration of nitrite ions in a sample of river water: 0.403, 0.410, 0.401, 0.380 mg/l. Should the last reading be rejected?

$$Q_{calc} = |0.380 - 0.401| / (0.410 - 0.380) = 0.7$$

But $Q_{table} = 0.829$ (at 95% level) for 4 values

Therefore, $Q_{calc} < Q_{table}$, and we cannot reject the suspect value.

Suppose 3 further measurements taken, giving total values of: 0.403, 0.410, 0.401, 0.380, 0.400, 0.413, 0.411 mg/l. Should 0.380 still be retained?

$$Q_{calc} = |0.380 - 0.400| / (0.413 - 0.380) = 0.606$$

But $Q_{table} = 0.568$ (at 95% level) for 7 values

Therefore, $Q_{calc} > Q_{table}$, and rejection of 0.380 is recommended.

But note that 5 times in 100 it will be wrong to reject this suspect value!
Also note that if 0.380 is retained, s = 0.011 mg/l, but if it is rejected, s = 0.0056 mg/l, i.e. precision appears to be twice as good, just by rejecting one value.