## 1. PREAMBLE

Statistics are essential for rating methods and for reporting analytic results. Evaluating data and methods without proper statistical procedures can lead to costly errors in judgement. Statistical methods are also used in industrial process control. Various types of graphs are used to predict and eliminate manufacturing errors, before they can cause serious problems.

## 2. REPORTING RESULTS

### 2.1 ERRORS

Reporting analytical results without an associated error is useless, and the information given is misleading. Reporting a result without an associated error gives an impression that the result is completely accurate. This assumption is false, considering that no analytical procedure known is totally accurate.

### *2.1.1 THE STANDARDIZED RANDOM ERROR*

To assess the uncertainty associated with a measurement calculate the standardised random error. This error gives the variability arising from random fluctuations in the operating environment.

Most authors, such as Friedmann, use Formula 1 for calculating the standardised random error.

$$S_x = \frac{s}{\sqrt{n}}$$

**Equation 1**

Formula 1 incorporates s, called the Sample Standard Deviation, given in Formula 2.

$$s_x = \sqrt{\frac{\sum(x-\overline{x})^2}{n-1}}$$

**Equation 2**

Formula 2 also uses X bar, which is just the simple arithmetic mean, and is given by Formula 3. n in Formulas 1, 2, and 3 is the number of determinations: if one made three determinations of a sample n would be 3.

$$\bar{x} = \frac{\sum x}{n}$$

**Equation 3**

As n increases, the error decreases towards an asymptotic limit of 0. Although it would be desirable to have a very small error, it is a waste of time and resources to make 200 determinations of a sample. At least three determinations are necessary for an accurate estimate of the error.

Devoe recommends using a modified version of Formula 1, given in Formula 4. He feels that this equation is more valid since it allows reporting at a 95% confidence level. In his experiments, he found that the means sometimes varied outside the predicted error, if the additional factor of t was not included.

$$s_x = \frac{t * s}{\sqrt{n}}$$

**Equation 4**

#### *2.1.1.1 Example 1:*

Calculate the standard randomised error from the average of the following three determinations of percentage chloropyrifos.

| ***Sample*** | ***Concentration (%)*** |
|---|---|
| 1 | 0.3829 |
| 2 | 0.4308 |
| 3 | 0..3763 |

$$\bar{x} = \frac{0.3829 + 0.4308 + 0.3763}{3} = 0.3967$$

$$s = \sqrt{\frac{(0.3829 - 0.3967)^2 (0.4308 - 0.3967)^2 (0.3963 - 0.3967)^2}{2}} = 2.976x10^{-2}$$

$$SE = \frac{4.30 * 2.976x10^{-2}}{2} = 0.1280$$

### *2.1.2 GRAND AVERAGES AND ERRORS*

A grand average calculated from averages with errors must also have an associated error. The best remedy for this problem is to simply average the errors, giving a good estimate of the grand average's error.

### *2.1.3 ESTIMATING THE ERROR*

Determining uncertainty, by calculating the standardized random error, i8 quite accurate. There are, however, times when one wishes to make a quick measurement of some physical property, such as density or pH. In this case, an estimation of the error would be sufficient, since only one measurement is being taken. Table II1 gives a list of the errors associated with common laboratory equipment.

**Table I[1]: Errors associated with common analytic instruments**

| ***Instrument*** | ***Error*** |
|---|---|
| Burettes (mL) | ±0.02 per reading |
| Pipettes (mL) | 5±0.01<br>10±0.02<br>25±0.03 |

---

[1] F.W. Fifield and D. Kealey, <u>Principles and Practice of Analytic Chemistry</u> 2nd edition, 1983.-With additions.

| | |
|---|---|
| | 50±0.05<br>100±0.08 |
| Volumetric flasks (mL) | 25±0.03<br>50±0.05<br>100±0.08<br>250±0.12<br>500±0.15<br>1000±0.30<br>2000±0.50 |
| Graduated cylinders (mL) | 100±0.10 |
| Analytic balance (g) | ±0.0002 per reading |
| Small balance (g) | ±0.05 per reading |
| Specific gravity balance | ±0.0005 per reading |
| pH meter | ±0.005 per reading |

- Note: Estimate the precision of an analytic instrument if not already known. The best approximation is five times the smallest reading. If a pH meter gave readings to two decimal places, then its error would be 0.05.

### *2.1.4 USING ESTIMATES OF PRECISION TO REPORT ERROR IN A CALCULATED RESULT*

There are times when a single replicate result has a number of errors contributing to its uncertainty. In this case it is necessary to combine the errors, as shown in Example 2. Convert errors into relative, or percentage, errors before they combining them.

#### ***2.1.4.1 Example 2:***

A 100 mL aliquot of Diazinon in a 100mL graduated cylinder weighs 56.78g find the total error.

$$E = 100 * \sqrt{(\frac{0.1}{100})^2 + (\frac{0.05}{56.78})^2} = 1.23\%$$
$$56.78 \pm (56.78 * 1.23) = 56.78 \pm 0.6984g$$

## 2.2 BIAS

In most cases, a method will have a bias associated with its results  This bias consists of a fixed and concentration dependant, also called relative, portion. An equation compensating for the bias is usually determined during method development.

### ***2.2.1.1 Example 3:***

Adjust for the bias if the equation compensating for it is given by Formula 5, the average is 0.3967, and the error is 0.04760.

$$x = \frac{1}{b}(\frac{1}{y} - a)$$

**Equation 5**

$$x = \frac{1}{6.25}\left(\frac{1}{0.3967 \pm 0.04760} - 0.065\right) = 0.625\left(\frac{1}{0.3967 \pm 0.04760}\right) - 0.065 = 0.3929 \pm 0.04715$$

- ♦ Note:  The relative error, which is the error divided by the result, remains constant unless two results are multiplied together or the result is raised to a power. If results with errors are raised to a power,  then their relative errors are multiplied by that power.

## 2.3 OUTLIERS

It is possible to determine if an excessively large, or small, value can be discarded, so that it no longer skews the average. Anderson describes a set of calculations suitable for this purpose. Formula 6 is useful for 3-7 determinations, Formula 7 for 8-10 determinations, and Formula 8 for 11-13 determinations.

$$R = \frac{x_n - x_{n-1}}{x_n - x_1}$$

**Equation 6**

$$R = \frac{x_n - x_{n-1}}{x_n - x_2}$$

**Equation 7**

$$R = \frac{x_n - x_{n-2}}{x_n - x_2}$$

**Equation 8**

Where, the values being arranged in ascending order,

$X_n$ = The suspected value.

$X_{n-1}$ = The value closest to the suspected value.

$X_{n-2}$ = The value second closest to the suspected value.

$X_1$ = The value furthest from the suspected value.

$X_2$ = The value second furthest from the suspected value.

**Table II[2]: Critical values for R in the test for outliers.**

| | Significance Level | | |
|---|---|---|---|
| Number in Set | 90 % | 95 % | 99 % |
| 3 | 0.885 | 0.941 | 0.998 |
| 4 | 0.679 | 0.765 | 0.889 |
| 5 | 0.557 | 0.643 | 0.780 |
| 6 | 0.482 | 0.560 | 0.698 |
| 7 | 0.434 | 0.507 | 0.637 |
| 8 | 0.479 | 0.554 | 0.683 |
| 9 | 0.441 | 0.512 | 0.635 |

---

[2] W.J. Dixon, Processing Data for Outliers, Biometrics, Vol. 9, No. 1, 74-89, 1953.

| | | | |
|---|---|---|---|
| 10 | 0.409 | 0.477 | 0.597 |
| 11 | 0.517 | 0.576 | 0.679 |
| 12 | 0.490 | 0.546 | 0.642 |
| 13 | 0.467 | 0.521 | 0.615 |

#### *2.3.1.1 Example 4:*

Given that a test for chloropyrifos concentration yielded three values of 0.3763%, 0.4308%, and 0.3829%, is 0.4308% an outlier?

Using Formula (6), and arranging the values in ascending order:

$$R = \frac{0.4308 - 0.3829}{0.4308 - 0.3763} = 0.8789$$

Since $R_{95,\,3}$=0.941, and the calculated R value does not exceed it, 0.4368% is not an outlier at 95% confidence.

- Note: The lag-1 autocorrelation plot, found in section D, can be used for the graphical evaluation of outliers.

### 2.4 IS THE PRECISION OF TWO SETS OF DATA THE SAME?

There are many occasions, such as a method being evaluated, when it would be advantageous to know whether two sets of data have the same precision. The statistical method that allows this to be determined is known as the F-Test and is given by Formula (9), with the critical values in Appendix D.

$$F = \frac{s_a^{\ 2}}{s_b^{\ 2}}$$

**Equation 9**

Where:

$s_a$= The larger standard deviation.

$s_b$= The smaller standard deviation.

#### *2.4.1.1 Example 5:*

Two different concentrations of chloropyrifos were determined, with the averages having standard deviations of 2.976X10-2 and 0.1002, and 3 and 2 determinations respectively. Is the precision of these two sets of data different?

$$F = \frac{0.1002^2}{0.02976^2} = 11.34$$

From the F table, the value for 2 by 1 degrees of freedom is 18.5. Since F=11.34 , the two sets of data have the same precision at a 95% significance level.

### 2.5 ARE TWO AVERAGES THE SAME?

If the analytical result from a sample appears much higher than would be expected, another analysis would be in order. When the second analysis yields a different average, one can perform a statistical test, called a two-tailed Students t-Test in order to determine whether they are the same.

The first thing that must be done is an F-Test, ascertaining whether the variance of the two sets is the same. If they are of homogeneous variation, Formulas 10 and 11 apply, otherwise Formulas 12 and 13 are used. The critical values for t are given in Appendix D.

$$t = \frac{\left|\bar{x}_1 - \bar{x}_2\right|}{s_p}\sqrt{\frac{n_1 n_2}{n_1 + n_2}}; df = n_1 + n_2 - 2$$

**Equation 10**

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

**Equation 11**

$$t = \frac{\left|\bar{x}_1 - \bar{x}_2\right|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

**Equation 12**

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1+1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2+1}} - 2$$

**Equation 13**

### *2.5.1.1 Example 6:*

Concentration 1 was 1.05196, s-8.891X10~2, n-9, and concentration 2 was 2,328%, s=1.308x10-1 Are these two averages statistically different?

First, use the F-Test to determine whether the two runs had the same precision.

$$F = \frac{0.1308^2}{0.08891^2} = 1.47$$

Since the calculated F is less than F 38=8.85, from the table of critical values in Appendix D, it Is possible to say that the two sets of results are of homogenous variance.

Now, using equation (10) and (11), calculate the t value.

$$s_p = \sqrt{\frac{(4-1)(0.1308)^2 + (9-1)\left(8.891x10^{-2}\right)^2}{4-1+9-1}} = 0.1021$$

$$t = \frac{|2.3276 - 1.0513|}{0.1021}\sqrt{\frac{9*4}{9+4}} = 20.80$$

$$df = 9 + 4 - 2 = 11$$

From the two-tailed t-Table, the critical value is t $_{95,\ 11}$ = 2.20.  The two averages are not statistically the same

When the averages are less than t, and/or there are very few measurements, the t-Test may give a false result.   Appendix D contains a table that gives the number of measurements necessary to ensure a specified sensitivity of the t-Test.

### 2.6 STANDARDIZING A PRIMARY REFERENCE

In order to standardize a primary reference, at least ten determinations must be made. The average is calculated and reported, with the randomized standard error (4), as the purity of the standard. The purity should be reported to three significant figures. If the error is unacceptably high, then more determinations should be made.

## 3. METHOD EVALUATION

### 3.1 ESTIMATING THE PRECISION OF A METHOD

The standard approach, for estimating the precision of a method, is to calculate the standard deviation from a large number of determinations. Another, much less labour intensive, technique is to pool the standard deviations from a set of determinations.

#### *3.1.1 WHEN CAN VARIANCES BE POOLED?*

One can only pool variances if they are homogenous. Two tests exist for this purpose. Cochranes Test for sets of data with equivalent degrees of freedom, and Bartletts Test for sets with different degrees of freedom. It is best to design a method evaluation so that the number of determinations remains constant considering that Bartletts Test is much more tedious to calculate. One could also use the F-Test to determine whether each variance is equivalent—the ultimate in tedium

##### ***3.1.1.1 THE COCHRAN TEST***

Like the F-Test, Cochran’s test, given by equation 14, uses a ratio of variances to calculate a critical value, called the g statistic. If the value of g, for k standard deviations and n observations, is greater than the calculated g, the variances are homogenous. Critical values for g can be found in the Appendix.

$$g = \frac{\left(s_{\max}\right)^2}{s_1^2 + s_2^2 + \ldots + s_k^2}$$

**Equation 14**

- Note: The test only works if all n’s are equal.

### *3.1.1.2 Example 7:*

For each of the following standard deviations, n was 3. Are the variances of these sets equal?

**Table III: Standard deviations of six standards.**

| ***Sample #*** | ***Standard Deviation*** |
|---|---|
| 1 | $1.339X10^{-2}$ |
| 2 | $2.974X10^{-2}$ |
| 3 | $5.119X10^{-2}$ |
| 4 | $3.747X10^{-2}$ |
| 5 | $1.027X10^{-2}$ |
| 6 | $3.409X10^{-2}$ |

With the number of determinations uniform between variances, Cochran’s Test would be suitable.

$$g = \frac{\left(5.119X10^{-2}\right)^2}{\left(1.339X110^{-2}\right)^2 + \left(2.974X10^{-2}\right)^2 + \left(5.119X10^{-2}\right)^2 + \left(3.747X10^{-2}\right)^2 + \left(1.027X10^{-2}\right)^2 + \left(3.409X10^{-2}\right)^2} = 0.4122$$

Since, from Appendix D, $g_{0.05,\ k=6,\ n=3}$ = 0.6161, is larger than the calculated g, all sets of data have the same variance.

#### *3.1.1.3 THE BARTLETT TEST*

Bartlett’s Test, given by Equations (15) to (21), is used to calculate an F value. The Bartlett Tests only fallback is its erroneous result of non-homogeneity, when the data is not normally distributed. It can, however, still be used if the variances approximate a normal distribution.

$$s_p^2 = \sum \frac{(n_i - 1)s_1^2}{(N - k)}$$

**Equation 15**

$$M = (N - k)\ln\left(s_p^2\right) - \sum (n_i - 1)\ln\left(s_i^2\right)$$

**Equation 16**

$$df_1 = k - 1$$

**Equation 17**

$$A = \frac{1}{3(k-1)}\left(\sum\left(\frac{1}{n_i - 1}\right) - \frac{1}{N - k}\right)$$

**Equation 18**

$$df_2 = \frac{k + 1}{A}$$

**Equation 19**

$$b = \frac{df_2}{1 - A + \frac{2}{df_2}}$$

**Equation 20**

$$F = \frac{df_2 M}{df_1(b - M)}$$

**Equation 21**

Where:

$n_i$ = Number of observations associated with each variance $s_i$.

N = Sum of all n’s

k = Number of variances being compared.

$df_1$ = Degrees of freedom for the numerator.

$df_2$ = Degrees of freedom for the denominator.

### *3.1.1.4 Example 8:*

Given the following standard deviations and observations determine whether the variances of these data ssets are equivalent.

**Table IV[3]: Standard Deviations of Iron Tests.**

| ***Sample #*** | ***Number of Determinations $n_i$*** | ***Standard Deviation $s_i$*** |
|---|---|---|
| 1 | 10 | 4.11X10-3 |
| 2 | 6 | 1.353X10-2 |
| 3 | 5 | 4.445X10-2 |
| 4 | 6 | 1.976X10-2 |

Since the number of determinations differs between samples it is necessary to use Bartlett's test.

---

[3] M. Fernandez, unpublished work, 1992

$$s_p^2 = \frac{(10-1)\left(4.11x10^{-3}\right)^2 + (6-1)\left(1.353x10^{-2}\right)^2 + (5-1)\left(4.445x10^{-2}\right)^2 + (6-1)\left(1.967x10^{-2}\right)^2}{27-4}$$

$$s_p^2 = 1.285x10^{-2}$$

$$M = (27-4)\ln\left(1.1285x10^{-2}\right)^2 - (10-1)\ln\left(4.11x10^{-3}\right)^2 - (6-1)\ln\left(1.353x10^{-2}\right)^2 -$$
$$(5-1)\ln\left(4.445x10^{-2}\right)^2 - (6-1)\ln\left(1.967x10^{-2}\right)^2 = 0.202403$$

$$df_1 = 4-1 = 3$$

$$A = \frac{1}{3(4-1)}\left[\frac{1}{10-1} + \frac{1}{6-1} + \frac{1}{5-1} + \frac{1}{6-1}\right] - \frac{1}{27-4} = 0.079737$$

$$df_2 = \frac{4+1}{0.079737^2} = 786.41$$

$$b = \frac{786.41}{1 - 0.079737 + \frac{2}{786.41}} = 852.1968$$

$$F = \frac{786.41(-0.202403)}{3(852.1968 + 0.202403)} = -0.0622447$$

Since $3.81 < F_{95,\,786,\,3} < 3.82$, all sets have equal variances.

## 3.2 POOLING THE STANDARD DEVIATIONS

Standard deviations can be pooled if they have been classified as homogeneous in variance by either the Cochran Test or Bartlett Test. The more standard deviations pooled, the closer the pooled standard deviation (s) will be to the true standard deviation (σ). A formula for calculating $s_p$ is given by Equation (22).

$$s_p = \sqrt{\frac{\sum\left(s_i^2 df_i\right)}{\sum df_i}}$$
$$df_p = \sum df_i$$

**Equation 22**

Where:

$df_p$ = Degrees of freedom associated with $s_p$.

$S_i$ = Standard deviation of set 1.

$Df_i$ = Degrees of freedom associated with that set.

- Note: See Appendix A for a definition of degrees of freedom.

#### *3.2.1.1 Example 9:*

Given the data from Table IV, calculate the pooled standard deviation.

$$s_p = \sqrt{\frac{\left(4.11x10^{-3}\right)^2 * 9 + \left(1.353x10^{-2}\right)^2 * 6 + \left(4.445x10^{-2}\right)^2 * 5 + \left(1.976x10^{-2}\right)^2}{10+6+5+6=27}} = 2.233x10^{-2}$$

$$df = 10+6+5+6=27$$

### 3.3 ESTIMATING CONFIDENCE LIMITS

When a standard deviation is calculated, it is merely an estimate of the true standard deviation (σ). The only case in which s equals σ occurs when infinite determinations are made. Since this never happens in reality, it is necessary to report a standard deviation with its uncertainty.

The Chi Squared Distribution, given by Equation (23), is rearranged, given by Equation (24), to calculate confidence limits for the standard deviation. Critical values for Chi Squared are given in Appendix D.

$$c^2 = \frac{(df)s^2}{s^2}$$

**Equation 23**

$$\sqrt{\frac{(df)s^2}{c_a^2}} < s < \sqrt{\frac{(df)s^2}{c_{1-a}^2}}$$

**Equation 24**

#### *3.3.1.1 Example 10:*

Using the pooled standard deviation calculated in Example 8, compute the confidence limits for σ, with α=0.05.

$$\sqrt{\frac{27*\left(2.233x10^{-2}\right)^2}{42.6_{0.05}}} < s < \sqrt{\frac{27*\left(2.233x10^{-2}\right)^2}{16.2_{0.95}}}$$

$$1.778x10^{-2} < s < 2.883x10^{-2}$$

## 3.4 COMPARING A METHOD PRECISION TO A REFERENCE METHOD’S PRECISION

A useful comparison between two methods is their precision. Differences between the precision of two methods can be determined using the Chi Squared Distribution, given by Equation(23).

### *3.4.1.1 Example 10:*

The $s_p$ determined in Example 8 is to be compared to the precision of another method, in which s=0.025. Is the reference method significantly less precise than the new method?

The critical value for Chi Squared at α=0.95%, with 27 degrees of freedom, is 16.2. There i8 a 95% chance that the calculated value would be above 16.2 from random chance. Therefore the two precisions are statistically the same.

$$c^2 = \frac{27*0.02233^2}{0.025^2} = 21.54$$

Conversely the critical value at α=0.05 is 40.1. Therefore, there is a 5% chance that a calculated value for Chi Squared would lie above 40.1 from random chance, but a 99.5% probability that it would be less than 40.1. For this reason the two precisions can also be considered statistically the same.

## 3.5 EXPERIMENTAL DESIGN AND REPORTING PRECISION

### *3.5.1 THE NUMBER OF DETERMINATIONS*

When designing an experiment to determine precision, one should strive to keep the number of determinations the same at each concentration. It is possible to vary the number of determinations between concentration levels, but the calculations become more complex. As a minimum, the method should be evaluated at six different concentrations, with three determinations at each level.

### *3.5.2 USING THE STANDARD DEVIATION TO REPORT PRECISION*

Reporting the precision as a pooled standard deviation is preferable. This strategy saves labour, and the higher number of degrees of freedom results in a more accurate estimate of s. In this manner, the results of plant quality control tests could be pooled to give an estimate of s. This estimate would require a large number of equivalent determinations if it were to be determined by traditional methods.

The precision should be reported as s with confidence limits, and be based on the greatest possible degrees of freedom. If the variance is calculated from sp, the greatest number of homogenous variances should be used. Otherwise, at least 9 determinations should be made, although more would improve the confidence limits.

### *3.5.3 WHEN TO USE THE COEFFICIENT OF VARIANCE IN REPORTING PRECISION*

There are times when the standard deviation statistically fluctuates over a range of concentrations. In this case, the coefficient of variance, abbreviated CV, i8 used. The CV is given by Formula (25), and is a ratio of the standard deviation to the arithmetic mean.

$$CV = \frac{s}{\bar{x}} * 100$$

**Equation 25**

Worthy of note is that the coefficient of variance i8 a dimensionless number, unlike the standard deviation. The coefficient of variance can also be pooled and have confidence limits estimated for it. This is done by substituting the coefficient of variance, in place of the standard deviation, within the appropriate formulas.

### *3.5.4 WHAT TO DO WHEN NEITHER CV OR S ARE CONSTANT OVER A CONCENTRATION RANGE*

In this case, either measure of precision is equally valid. The only drawback is that a different assessment of variance has to be made at each concentration level. The assessment of variance should be based upon at least five determinations for each level.

## 4. EVALUATING THE ACCURACY OF A METHOD

Accuracy is just as important a measure of a method’s performance as precision. Statistics can answer the question is the methods value the same as a reference value? If no reference value exists, is the value the same as a reference method’s With these concerns answered, the accuracy of a method can be reported.

### 4.1 IS THE METHODS VALUE THE SAME AS AN EXACT REFERENCE VALUE

To solve this problem, a modified t-test must be used, given by Equation (26). This test assumes that there is no uncertainty in the reference material’s concentration.

$$t = \frac{\bar{x} - m_0}{s}\sqrt{n}$$
$$df = n - 1$$

**Equation 26**

Where:

X bar = the arithmatic average of the method values.

$\mu_0$ = the reference value.

S = the methods standard deviation.

N = the methods number of determinations.

#### *4.1.1.1 Example 11:*

The average, from six determinations, of a GCFID method evaluation testing the concentration of Diazinon was 0. 5529%, with s = $3.02X10^{-3}$. The actual concentration of the reference diazinon was 0.4592%. Is the method statistically valid?

$$t = \frac{0.5529 - 0.4592}{3.02X10^{-3}}\sqrt{6} = 76$$
$$df = 6 - 1 = 5$$

1. With $t_{5,\ 0.05} = 2.57$, the method does not yield an accurate result.

- ♦ Note: If an average fails the t-test, then non-random influences cause the deviation from the reference value. This average is not statistically the same as the reference.

## 4.2 IS THE METHOD'S VALUE THE SAME AS A REFERENCE METHOD'S?

In this case, along with the case of a reference solution having an error associated with it, one should refer back to section B 5. The t-tests in that sections will be sufficient to determine whether an average of a method evaluation is equal to a reference concentration or method.

## 4.3 REPORTING THE METHOD'S ACCURACY

At least three determinations, at a representative concentration level, should be made to determine accuracy. Accuracy is reported as the percent difference, given by Equation (27).

When a method is evaluated, the %Difference at each of the concentration levels should be averaged. The magnitude of the %Difference at each level should also be averaged. These two results are reported as the accuracy for the method.

$$\%difference = \frac{(reference\ value - \ Measured\ value)}{reference\ value} * 100$$

**Equation 27**

## 5. EVALUATING THE BIAS OF A METHOD BY LINEAR REGRESSION

Each method has a bias associated with it, and an equation relating the measured and true concentrations can be statistically derived. This equation eliminates the non-random portion of a method s error, making an analytical result much more accurate.

Many equations will fit a set of data points, but only one gives the best estimate. The independent variable, x, is always the actual concentrations whereas y is the measured concentration.

A good statistics package, such as GBStat[4], can be used to calculate the best fit equation for a set of data automatically. Such a program is a worthwhile investment since it can save hours of drudgery.

### 5.1 TRANSFORMING PAIRS OF DATA

In order to fit an equation to a set of a data by simple means the data must first be transformed to linearity. For this reason there are a limited number of equations that can be used to fit a set of data, given in Table V. The alternative to transforming the data is to use advanced differential calculus for fitting an equation to the data points: a much less pleasant alternative.

**Table V: Equations and their Transformations.**

| ***Curve to fit*** | ***Curve rearranged for estimating the true concentration*** | ***Change y to*** | ***Change x to*** |
|---|---|---|---|
| $y=a+bx$ | $x=(y-a)b^{-1}$ | $y$ | $x$ |
| $y=a+bx^{-1}$ | $x=b(y-a)^{-1}$ | $y^{-1}$ | $x$ |
| $y=x(a+bx)^{-1}$ | $x=a(y^{-1}-b)^{-1}$ | $xy^{-1}$ | $x$ |
| $y=ab^{x}$ | $x=\ln(y)[\ln(ab)]^{-1}$ | $\log y$ | $x$ |
| $y=ae^{bx}$ | $x=\ln(ya^{-1})b^{-1}$ | $\log y$ | $x$ |

---

[4] Dr. Phillip Friedmann, GB-Stat, Dynamic Microsystems Inc., 1988, 89, 90, 91.

| | | | |
|---|---|---|---|
| $y=ax^{b}$ | $x=(y/a)^{(1/b)}$ | log y | log x |
| $y=ax+b^{n}$ | $x=((y-a)/n)^{(1/n)}$ | y | $x^{n}$ |

### 5.1.1.1 Example12:

Transform the following data into a form suitable for linear regression, approximating an equation of the form $y=ax^{b}$.

**Table VI**

| ***Real %*** | ***Measured X Bar*** |
|---|---|
| 0.3016 | 0.2980 |
| 0.3993 | 0.4080 |
| 0.5010 | 0.5054 |
| 0.6020 | 0.5992 |
| 0.7013 | 0.7020 |

Take the logarithm of both the real and the measured values giving the results in Table VII.

**Table VII**

| ***Log real %*** | ***Log measured %*** |
|---|---|
| -0.5206 | -0.5258 |
| -0.3987 | -0.3893 |
| -0.3002 | -0.2964 |
| -0.2204 | -0.2224 |
| -0.1541 | -0.1537 |

## 5.2 THE CORRELATION COEFFICIENT

The correlation coefficient, r, is used to determine how well an equation fits a set of data. Values of r2, the coefficient of determination, should be calculated for all of the possible transformations. The transformation with the largest r2 is then fit with linear regression techniques.

The equation for r is given by Formula (28), and critical values for the function are found in Appendix D. If the magnitude of the calculated value for r is greater than the critical value then the equation is a statistically valid fit. If r=1, the curve has a positive slope, whereas an r value of -1 indicates a negative slope.

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right]\left[\sum y^2 - \frac{\left(\sum x\right)^2}{n}\right]}}$$

**Equation 28**

### ***5.2.1.1 Example 13:***

Determine r for the data in Table VI, thereby evaluating how well an equation of the form y=A+Bx fits the data.

Since the data in Table VI has not been transformed, all calculations performed on it apply to the equation y=A+Bx. For an evaluation of the other relations, the appropriate transformation of the data must first be performed. Only then can the calculations be repeated on this data.

$$r = \frac{1.3588 - \frac{(2.5052)(2.5122)}{5}}{\sqrt{\left(1.3556 - \frac{2.5052}{5}\right)\left(1.3621 - \frac{2.5122}{5}\right)}} = 0.9995$$

$$r^2 = 0.9990$$

$$df = 5 - 2 = 3$$

The calculated value for r is much greater than the table value of $r_{\alpha=0.05,\ df=3} = 0.878$, and can therefore be considered a good statistical; fit.

- Note: Although a particular fit may be statistically valid, it is not necessarily the best fit. $r^2$ values for all of the transformations were calculated, and the linear fit was judged to best describe the data.

## 5.3 DETERMINING THE EQUATION OF BEST FIT

Now that the most suitable relationship has been found, all that has to be determined is the actual equation for the fit. The equations corresponding to the various transformations are given in Table V. Therefore, only the coefficients A and B now need to be estimated from the transformed data, using Formulas (29) and (30).

$$B = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}$$

**Equation 29**

$$A = \bar{y} - B\bar{x}$$

**Equation 30**

### *5.3.1.1 Example 15:*

Now calculate the equation, of the form y=A+Bx, fitting the data in Table VI.

Since the best fit for the data is linear, no transformation of the data is necessary. Therefore, the data in Table V should be used.

$$B = \frac{1.3588 - \frac{(2.5052)(2.5122)}{5}}{1.3566 - \frac{(2.5052)^2}{5}} = 0.996915$$

$$A = 0.50244 - (0.996915)(0.50104) = 2.946x10^{-3}$$

The line and plotted data are shown in Figure 1.

- Note: With three data points, a non-linear fit may have a higher co-efficient of determination than a linear fit. In many cases, the non-linear fit will also approximate linearity. The linear fit should be chosen preferentially in such a case, especially if it more accurately describes the chemistry of the method.

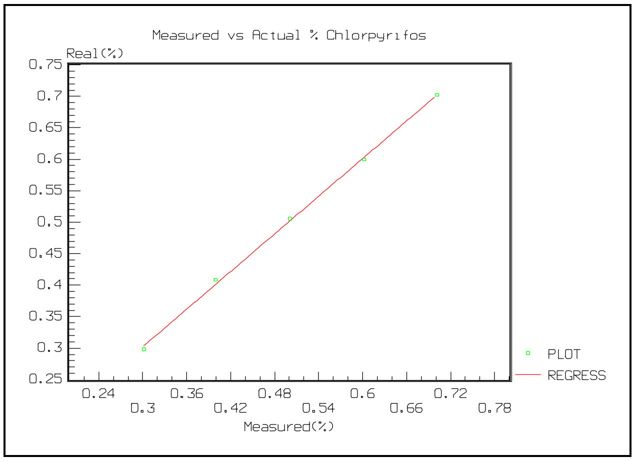

**Figure 1: Plot of Chlorpyrifos Data**

## 6. SUMMARY

1. In order to evaluate a method, the following steps should be taken:

### 6.1 EVALUATING PRECISION

1. Evaluate the variance at six different levels, with at least three replications each. Then determine if the standard deviations between the concentration levels are homogenous, using the Cochran Test. NOTE: KEEP THE NUMBER OF REPLICATIONS THE SAME BETWEEN LEVELS, OTHERWISE THE BARTLETT TEST HAS TO BE USED.
    - If they are homogenous, then the precision at one level is an estimate of the entire method's precision.
        - If routine analytical data is available, use Cochran's or Bartlett's Test to determine their homogeneity. If the analytical data's variances are homogenous, pool them by (22), and use the Chi Squared Distribution (24) to estimate confidence limits for this precision.

    - Otherwise, pool the variances of the six concentration levels using (22), and estimate the confidence limits of the precision using the Chi Squared Distribution (24).
- If the standard deviations are non-homogenous between levels, check to see if the coefficients of variance (24) are homogenous between levels, using Cochran's or Bartlett's Tests.
    - If they are homogenous, pool them in the same manner as the standard deviations see above.
    - If they are also non-homogenous, then two extra determinations per level must be performed. The precision should be reported, at each level, as a standard deviation with confidence limits.

### 6.2 EVALUATING ACCURACY

1. Compare the averages at each of the six concentration levels to the reference value. If the method's value is being compared to an exact reference value, use the t-test (26), otherwise use the t-tests in Section B.
    - If the averages are statistically the same as the reference value, calculate the %Difference at each level using (27). Then average the %Differences at each level. Also average the magnitude of the %Difference at each level. Report these averages as the accuracy of the method.
    - If the averages are not statistically equal to the reference value at each level, then the method gives results that differ from the reference value because of non-random influences. Report the accuracy as above.

### 6.3 EVALUATING BIAS

1. Use a computer program, such as GB-Stat if at all possible.
2. Perform the transformations, for all the curve fitting equations using (28).
3. Calculate $r^2$, the coefficient of determination for each of the curve fitting equations using (28).
4. Fit the equation with the highest $r^2$ to the data using (29) and (30).

5. Plot the average of the measured values V6. the reference values and graph the best fit equation through the data. This pictorial representation gives an excellent idea of how well the equation fits the data.