# Teaching modern data analysis with the Royal Australian Chemical Institute's titration competition

**Author/Contributor:**
Hibbert, D. Brynn

**Publication details:**
Australian Journal of Education in Chemistry
v. 66
pp. 5-11
1445-9698 (ISSN)

**Publication Date:**
2006

# Teaching modern data analysis with The Royal Australian Chemical Institute's titration competition

D Brynn Hibbert

School of Chemistry, University of New South Wales, Sydney, NSW 2052, Australia

e-mail: b.hibbert@unsw.edu.au

**Abstract**

The Royal Australian Chemical Institute (RACI), the professional body for chemists in Australia, runs a yearly titration competition for high school students. In 1997, twenty five teams of three students competed in a heat at the University of New South Wales, Sydney. The results are an excellent set of data, showing random and gross errors, that can be used to illustrate many basic aspects of data analysis, including histograms, normal distribution of data, means and standard deviations, robust estimators, hypothesis tests and measurement uncertainty. They also support general observations that ¼ to ⅓ of analytical results might not be fit for purpose, and provide a platform for a discussion of quality in analytical chemistry.

**Key Words**: Data analysis, statistics, normal distribution, measurement uncertainty, acid base titration

# 1 Introduction

Teaching the statistics of data analysis to undergraduate students can be considered quite straight forward. Concepts such as mean and standard deviation might already be familiar, and the formulae for confidence intervals, t-tests and the like are not the hardest to understand. With access to spreadsheets, graphs and calculations have become much easier. What students struggle with, in my experience, is simply why they need to do data analysis, and what information is really being gleaned. Modern approaches to metrology stress the 'uncertainty approach' in which a holistic view is taken, rather than a classical assignment of dispersion of results in terms of random and systematic error (ISO, 2005). Laboratories that are accredited to ISO 17025 (ISO, 1999) must estimate measurement uncertainty and must report it with their results when required by the client and when there is comparison with a limit. A measurement result is a value (number plus unit) and an uncertainty, often expressed as an expanded uncertainty, which is a range in which the value of the measurand is expected to lie with a certain probability (ISO, 1993).

This modern approach is an attempt to improve the quality of analytical results, and in the author's view, this has to start as early as possible during the training of an analytical chemist. There are a number of examples of the high cost of poor analysis, and in recent years pronouncements from national institutes like NIST (National Institute of Standards and Technology, USA) (May, 2001) and the LGC (Laboratory of the Government Chemist, UK) (King, 1995) has implied that a surprisingly large fraction of analytical results are not fit for purpose. For example in a survey of clients of analytical chemistry laboratories carried out in the early 1990s, the LGC found 29% of respondents reported results that did not meet the customer requirements, and of these 12% caused 'very serious loss' to the customer's business (King, 1995).

Courses on data analysis, often a few lectures in an analytical or physical chemistry subject, provide useful examples for their students to study, but these are often chosen piecemeal with a specific illustrative objective for each, and can lead to an incoherent whole. The data reported in this paper happens to have several useful qualities that provide an exemplar for a good part of basic data analysis. An advantage of the use of the data in Australia is that many university chemistry students will have taken part in the competition in their final year at high school, and so feel some ownership and sympathy with the anonymous students, whose results are being picked over and analyzed. Here about 5 or 6 lectures in a second year (of a three year BSc course) subject on analytical chemistry is based on one set of results, from the RACI titration competition of 1997.

## *1.1 The RACI titration competition*

Each year the RACI organizes around the states and territories of Australia a titration competition, open to students attending high schools, usually in their last two years before tertiary education (grade 11 and 12). Winners of regional heats go on to a final, and the whole competition has been a good instrument for raising awareness of chemistry and the need for proper laboratory techniques. The model was the popular schools analysis competition under the auspices of the Royal Society of Chemistry, first organised in London by then Polytechnic of North London (now the London Metropolitan University) in 1982 (RSC, 2005).

In the RACI competition teams of three are given (1) a sodium hydroxide solution, (2) a hydrochloric acid solution of assigned molarity, and (3) individual solutions of acetic acid. Common indicators are available, and glassware is provided (although many students from private schools come with their own calibrated pipettes and burettes). The students are expected to return the three amount concentrations of the acetic acid solutions, and a marking scheme is used that usually adds the absolute error (|assigned value – student's reported value|) of each student, with the team with lowest aggregate being the winner. The expected procedure is that each student will titrate the hydrochloric acid solution with sodium hydroxide and thus calculate the molarity of the sodium hydroxide. Then with a suitable indicator (e.g. phenol phthalein) each will titrate the acetic acid with the, now standardized, sodium hydroxide and thus return the required concentration of their acetic acid. It is recommended that each student perform his or her own set of titrations, although it is known that some teams pool their knowledge of the concentration of sodium hydroxide (and with occasional disastrous results).

# 2 The results of the 1997 UNSW heat

Table 1 gives the results of the 26 teams that took part in the 1997 heat held at the University of New South Wales, Sydney.

Table 1: Results of the analysis of three acetic acid solutions Assigned values: A = 0.1147 M; B = 0.1241 M; C = 0.1340 M.

| Team | Solution concentration of A (M) | concentration of B (M) | concentration of C (M) |
|---|---|---|---|
| 1 | 0.1146 | 0.1242 | 0.1341 |
| 2 | 0.1148 | 0.1238 | 0.1343 |
| 3 | 0.1150 | 0.1241 | 0.1343 |
| 4 | 0.1150 | 0.1243 | 0.1336 |
| 5 | 0.1148 | 0.1247 | 0.1346 |
| 6 | 0.1139 | 0.1244 | 0.1336 |
| 7 | 0.1142 | 0.1244 | 0.1336 |
| 8 | 0.1144 | 0.1227 | 0.1339 |
| 9 | 0.1152 | 0.1245 | 0.1327 |
| 10 | 0.1155 | 0.1256 | 0.1345 |
| 11 | 0.1158 | 0.1252 | 0.1350 |
| 12 | 0.1143 | 0.1243 | 0.1323 |
| 13 | 0.1141 | 0.1255 | 0.1335 |
| 14 | 0.1153 | 0.1262 | 0.1336 |
| 15 | 0.1145 | 0.1231 | 0.1319 |
| 16 | 0.1177 | 0.1249 | 0.1360 |
| 17 | 0.1134 | 0.1246 | 0.1306 |
| 18 | 0.1144 | 0.1281 | 0.1352 |
| 19 | 0.1219 | 0.1299 | 0.1414 |
| 20 | 0.1138 | 0.0908 | 0.1330 |
| 21 | 0.1143 | 0.0855 | 0.1328 |

| | | | |
|---|---|---|---|
| 22 | 0.0920 | 0.0840 | 0.1278 |
| 23 | 0.1222 | 0.1212 | 0.0850 |
| 24 | 0.1556 | 0.1645 | 0.1231 |
| 25 | 0.0936 | 0.0854 | 0.0818 |
| 26 | 0.9083 | 0.8589 | 0.7746 |

Apart from the hapless teams 25 and 26 many of the students appear to have made a good attempt at the analysis. When teaching with these data, I point out that tabulated values are not always easy to interpret and form an impression of the nature of the results. I also observe that out of 75 results only one person (team 3 B) actually returned a value that was 'correct', and recall the famous saying by Berzelius who is quoted as saying about analysis "…… not to obtain results that are absolutely exact – which I consider only to be obtained by accident – but to approach as near accuracy as chemical analysis can go."

## *2.1 Analysis of the data*

### 2.1.1 Plots

It is always recommended to graph the data in some way. A simple plot against team number highlights the real problem team (Figure 1a) , but also shows that one great outlier compresses the rest of the data.

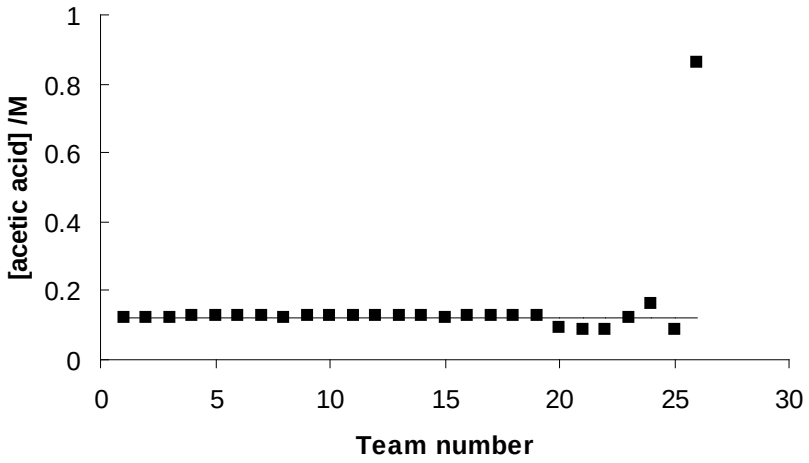


Figure 1: Results of the RACI titration competition for sample B. Dashed line is the assigned value.

The data is sorted and the greatest point left out, now showing some more potential outliers (Figure 2). Homing in on the plot of Figure 3 shows the data that we can demonstrate (later) to be normally distributed. Students' attention must be drawn to the scales on the *y*-axes to emphasise the differences between the core data and the outliers.

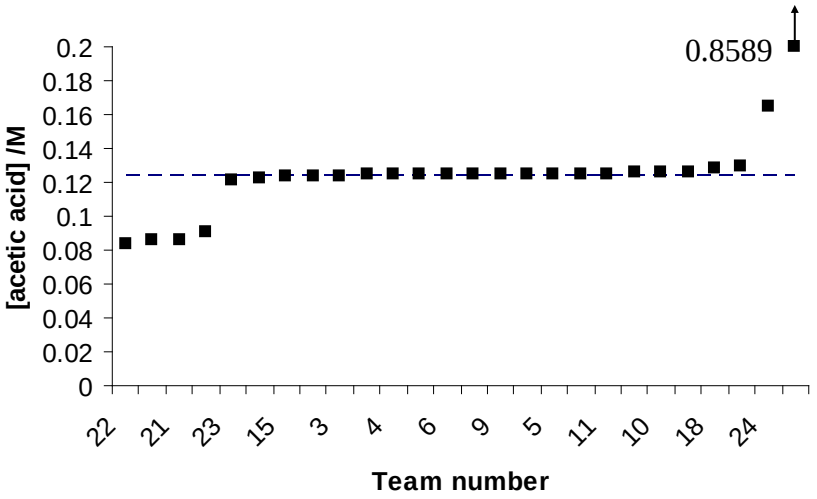


Figure 2: The data of Figure 1, ordered.

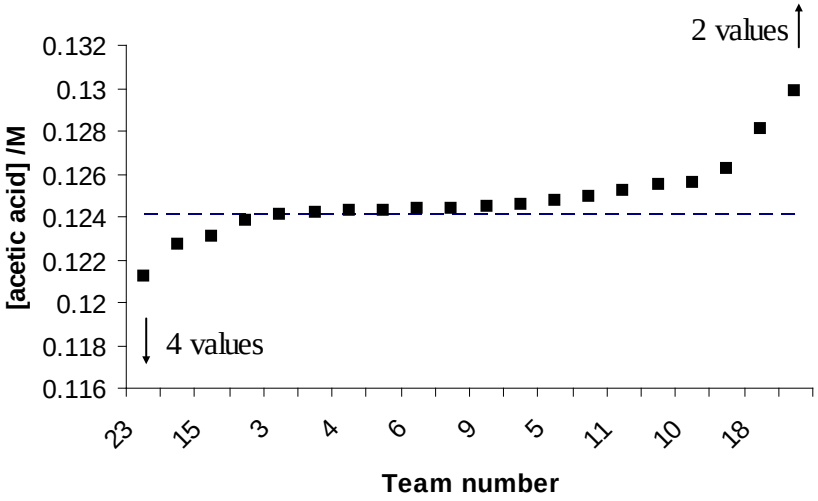


Figure 3: The data of Figure 1 with outliers removed.

### 2.1.2 Histogram

A problem with displaying data using histograms, is that unless there is a reasonable amount of data, they rarely look convincingly normal. Here the 78 data have been expressed as a % error = 100 × (value – assigned value)/assigned value, to allow for the different assigned values and a histogram is shown in Figure 4. This results in two groups, one of extreme values (the 21 in the ‘less’ and ‘more’ categories) and the remaining majority which seem to cluster about the correct answer (i.e. error = 0).

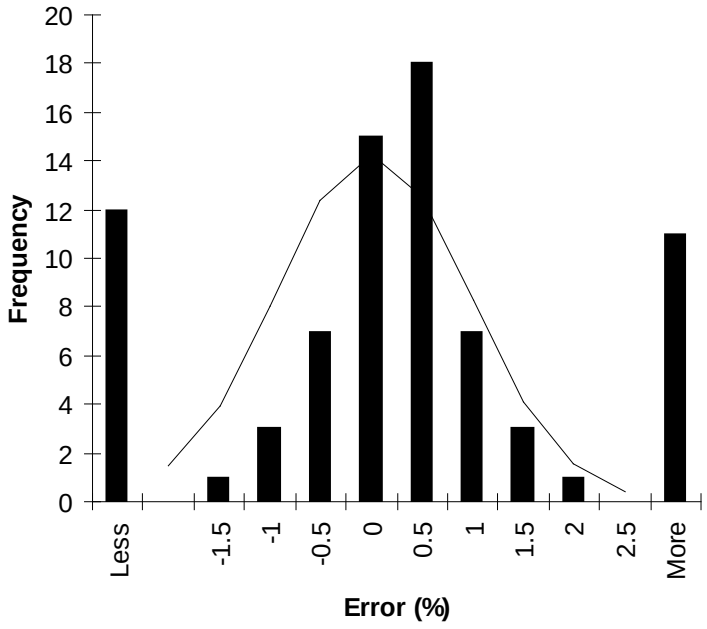


Figure 4: A histogram of the errors of the results of Table 1. A bar contains values between the value shown for the previous bin and its value, e.g. the bar labelled 0.5 counts results between 0 and 0.5 %. The solid line is the normal distribution with mean 0.01% and standard deviation 0.94 %.

### 2.1.3 Rankit

There is not always time in a short course of data analysis to cover testing of data for normality, but the Rankit method can be quickly implemented in Excel and can provide a platform for a discussion of distributions. The Rankit method is as follows, with each step creating data in an adjacent column.

1. Sort the data into ascending order. In Excel this is done via a command in the Data menu, or icon on the Standard icon bar.

2. Write the cumulative frequency of the data, that is how many data have values equal to or less than the ranked value. This definition means that ties get the higher rank. (data 7.1, 7.3, 7.3, 7.5 is ranked 1, 3, 3, 4)In Excel this becomes = COUNT($*range*) + 1 – RANK(*cell* ,$*range*), where $*range* is the range of cells containing the ordered data with reference fixed to allow copying down the column (e.g. $A$1:$A$27), and *cell* is the cell containing the value to be ranked (e.g. A1).

3. Calculate the normalized cumulative frequency as $f$ = cumulative frequency/$n$+1, where $n$ is the number of data.

4. Calculate the point on the normal distribution corresponding to the normalized cumulative frequency, $z$ =NORMSINV($f$)

5. Plot $z$ against the data.

A straight line through $f(z) = 0$ indicates normality. Outliers are displaced to the left or right . Table 2 has example Rankit calculations, and Figure 5 is the Rankit plot for all the data. As before, team 26 distorts the plot, but by reducing the data set, data that gives a good straight line indicating a normal distribution is easily found (Figures 6 and 7).

Table 2. Rankit calculations for the 78 results of the titration competition. Error = 100 × (result – assigned value)/assigned value. (The data is copied directly from an Excel spreadsheet and no precision of the values is implied).

| Error (%) | Rank | z |
|---|---|---|
| -38.9552 | 1 | -2.23654 |
| -36.5672 | 2 | -1.95458 |
| -32.3127 | 3 | -1.77469 |
| -31.1845 | 4 | -1.63875 |
| … | … | … |
| 32.55439 | 74 | 1.527719 |
| 35.65824 | 75 | 1.638748 |
| 478.0597 | 76 | 1.774688 |
| 592.1031 | 77 | 1.954577 |
| 691.8919 | 78 | 2.236538 |

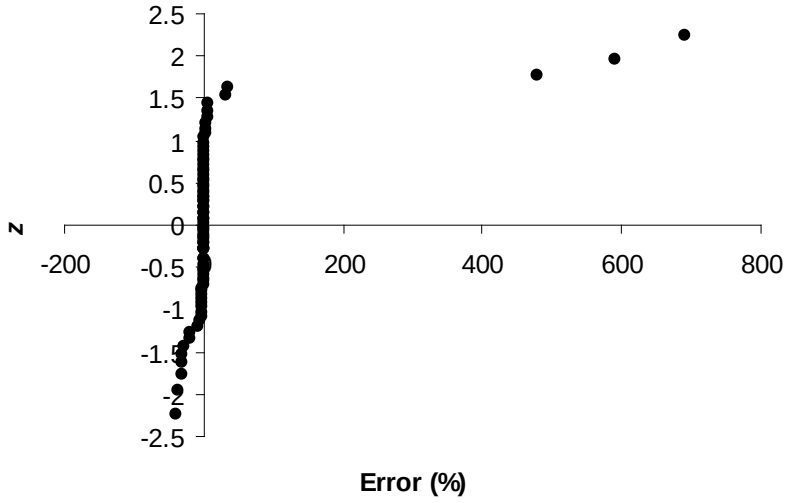


Figure 5: Rankit plot for the 78 results of the titration competition

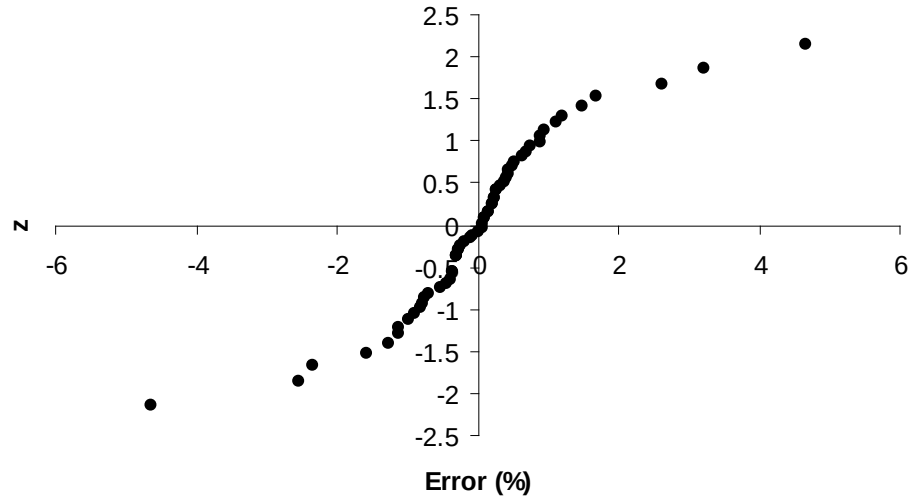


Figure 6: Rankit plot of the titration data without the highest 9 and lowest 8 data

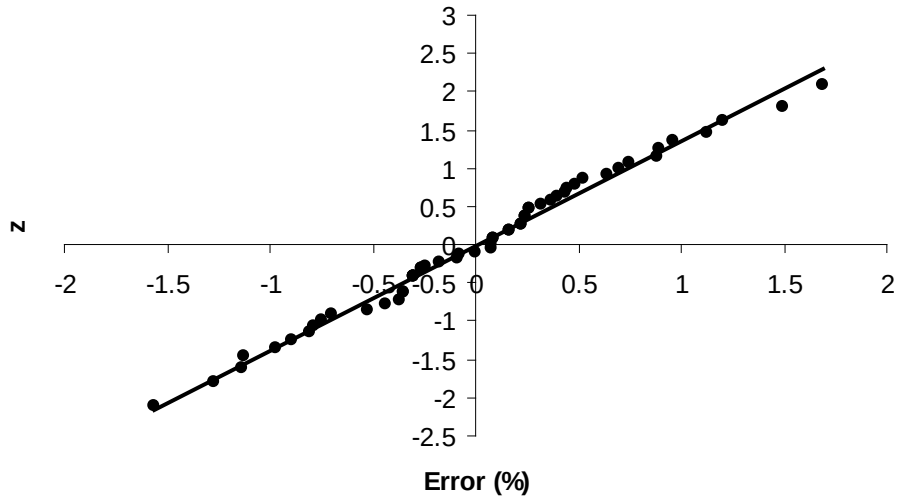


Figure 7: Rankit plot of the titration data with the 'more' and 'less' data of the histogram of Figure 4 removed. The solid line is a Trendline generated by Excel.

As with the raw data, mention must be made of the *x*-axis scales of these plots. With this kind of data, an outlier test for single outliers such as a Grubbs' test, now recommended by ISO and IUPAC, or a Dixon's Q-test, should not be used (Miller and Miller, 2000).

### 2.1.4 Means, medians and standard deviations

The basic measures of a normal distribution are the mean ($\mu$) and standard deviation ($\sigma$). Using our data we can estimate these parameters by the arithmetic average ($\overline{x}$ for *N* data)

$$\bar{x} = \frac{\sum_{i=1}^{i=N} x_i}{N} \quad (1)$$

and the sample standard deviation ($s$)

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

However these estimates are only valid if the data on which they are based are a random sample of normally distributed data. The titration data illustrates the dangers of calculating sample means and standard deviations before checking for outliers and normality. So-called 'robust' estimators are used when a data set is not perfectly normally distributed to provide reasonable values for mean and standard deviation. The median, which is the middle value of data when they are arranged in order (or the average of the two middle values if there are an even number of data), is a robust estimator of the mean. A basis of the equivalent estimator of standard deviation is the interquartile range or IQR. The data is ordered and the median determined. Then the medians of each half of the data give the IQR. The IQR is thus the range of data containing the middle 50% of the data. Calculation of the IQR requires sufficient data to bin into quartiles, and so cannot be used on less than about one dozen data. As ± 1 standard deviation encompasses 68% of the data it is easy to show that the IQR × 0.75 is an estimate of the standard deviation. This is called the normalised interquartile range, or NIQR. An alternative robust estimator of the standard deviation, the normalised median absolute deviation (NMAD), has been supported by Miller and Miller (Miller and Miller, 2000) as a more useful robust estimate for small data sets. The median absolute deviation (MAD) is the median of the absolute differences between each value and the median of the data. Divided by 0.68 the MAD becomes the normalised MAD (NMAD). Table 3 has the mean, median, standard deviation, NIQR and NMAD for the titration data.

Table 3: Statistics of the titration results for team member B. Assigned value = 0.1241 M. NIQR = interquartile range × 0.75, NMAD = median$|x_i - \text{median}(x)|$ /0.68

| statistic (unit) | All data | Data without greatest 4 and least 4 values |
|---|---|---|
| Mean (M) | 0.1486 | 0.1243 (assigned value 0.1241) |
| *s* (M) | 0.1458 | 0.00114 |
| median (M) | 0.1244 | 0.1244 |
| NIQR (M) | 0.00161 | 0.00054 |
| NMAD (M) | 0.00169 | 0.00059 |

Courtesy of team 26, the raw data is skewed high, as has been seen, and the mean and standard deviation of the original data does not give a useful estimate of the centre of the data nor the spread. The median and NIQR and NMAD for the whole data do return values that are in keeping with the mean and standard deviation of the normally distributed data. The message that goes with these calculations is that wherever possible data that can be demonstrated to be normally distributed should be identified and the sample mean and sample standard deviation calculated. When this is not possible or desirable, and the data is known to have outliers or be skewed in some way, robust estimates of mean and standard deviation are to be used. An example is in interlaboratory trials where each result must be preserved and robust *z*-scores, $z_i = [x_i - \text{median}(x)]/\text{NIQR}$ , used (Hibbert, 2005). At this stage it is worth reminding students that all of these statistics have the units of the measurand (here M), and that the symbol for sample standard deviation is *s*, not sd, s.d., SD, std dev, and so on.

## 2.2 Did the students get the right answer?

If the course includes hypothesis testing then the data can be used to answer the question, are the means of the students' data significantly different from the assigned values? A Student-*t* test is used for each set of results (A, B, C).

$$t = \frac{|ARV - \bar{x}|}{\left(s/\sqrt{n}\right)}, \text{ with } n-1 \text{ degrees of freedom} \qquad (1)$$

where *ARV* is the assigned reference value and $\bar{x}$ and *s* the sample mean and standard deviation of the normally distributed data. The equation for the *t* value is written as in (1) in order to emphasize that *t* is just the difference between mean and assigned value expressed in standard deviations of the mean. The null hypothesis, $H_0$, is that the experimental data come from a population with mean $\mu = ARV$. The *t* values from (1) may be compared with 95% two-tailed *t*, generated in Excel by =TINV(0.05,*n* – 1), or the probability associated with *t* calculated by *P* =TDIST(*t*, *n* – 1,2). This allows discussion of what is being tested – not the probability of $H_0$, but the probability of the data given the truth of $H_0$. Later a more formal definition of *P* as the probability of finding a *t* value more extreme than *t* in repeated experiments, can be given together with the knowledge that *P* is also the probability of making a Type I error if $H_0$ is rejected.

Table 4: Student-*t* tests on the means of the titration results

| | Solution | | |
|---|---|---|---|
| | A | B | C |
| Assigned value /M | 0.1147 | 0.1241 | 0.1340 |
| mean /M | 0.1146 | 0.1243 | 0.1338 |
| s /M | 0.0006 | 0.0011 | 0.0010 |
| n | 19 | 18 | 19 |
| standard deviation of the mean /M | 0.00014 | 0.00027 | 0.00024 |
| *t* | 0.7493 | 0.8058 | 0.7831 |
| P(*T*>*t*) | 0.4634 | 0.4315 | 0.4438 |
| $t_{0.05",n-1}$ | 2.1009 | 2.1098 | 2.1009 |

So the answer is clear, analysis of the solutions by team members who obtained results that were not classed as outliers gave mean results that were consistent with coming from populations with the assigned values. (At this stage we might lapse from statistical orthodoxy and suggest 'they did get the answer right').

# 3 Measurement uncertainty

A major change in our understanding of measurement results has come with the rise of the field of 'metrology in chemistry'. This has brought understanding of measurement uncertainty as something more than a 95% confidence interval calculated from a few repeated measurements. The ISO-approved method for estimating uncertainty is given in the "Guide to the expression of uncertainty in measurement" (ISO, 1993) and referred to by everyone as 'the GUM'. A full GUM calculation is a serious business (as an example see (Saed Al-Deen, Hibbert et al., 2004)) but the principles of auditing what factors might contribute to the uncertainty of a result can be instilled. The approach I have taken in my course is to try and predict the relative standard deviations of the students' results which were, for the normally distributed data, A: 0.0053, B: 0.0092, and C: 0.0077, i.e. between 0.5 to 1 %.

Also recorded with the results are the students' actual titration volumes. If the relative standard deviations are averaged (as squared RSDs) over the top ten teams an estimate of the repeatability of a single titration as an RSD is 0.0033, or for a volume of 25 mL $s_r$ = 0.083 mL . Conventional wisdom has that the reproducibility (precision under conditions of changing analysts, equipment, reagents, time) is two to three times the repeatability (precision under conditions in which the experiment is replicated by the same analyst with the same equipment over a short period of time), and this is followed here.

Approaches to estimating measurement uncertainty suggest construction of a cause-and-effect diagram based on the formulae used to calculate the result. A cause and effect diagram, also called after its early protagonist Ishikawa, or its fish-bone shape, is a way of displaying and connecting information about a particular outcome. The amount concentration

$$M_{\mathrm{NaOH}} = \frac{M_{\mathrm{HCl}} V_{\mathrm{HCl}}}{V_{\mathrm{NaOH}}} \qquad (2)$$

$$M_{\mathrm{AcOH}} = \frac{M_{\mathrm{NaOH}} V'_{\mathrm{NaOH}}}{V_{\mathrm{AcOH}}} \qquad (3)$$

where $M_{\mathrm{NaOH}}$ and $M_{\mathrm{AcOH}}$ are the molarities of sodium hydroxide and acetic acid respectively and the $V$ are the volumes of the subscripted solutions. Most students kept the sodium hydroxide in the burette and titrated first 25 mL of the standard hydrochloric acid, and then 25 mL of acetic acid. As the assigned molarity of the hydrochloric acid solution was 0.1068 M, the volumes of the titrations would be around 25 mL, and the uncertainty estimations need only be done once for this volume.

A cause and effect diagram is evolved by considering first equation 3 (figure 8), and then expanding $M_{\mathrm{NaOH}}$ in equation 2 (figure 9), and finally collecting all the precision terms to give the repeatability (figure 10).

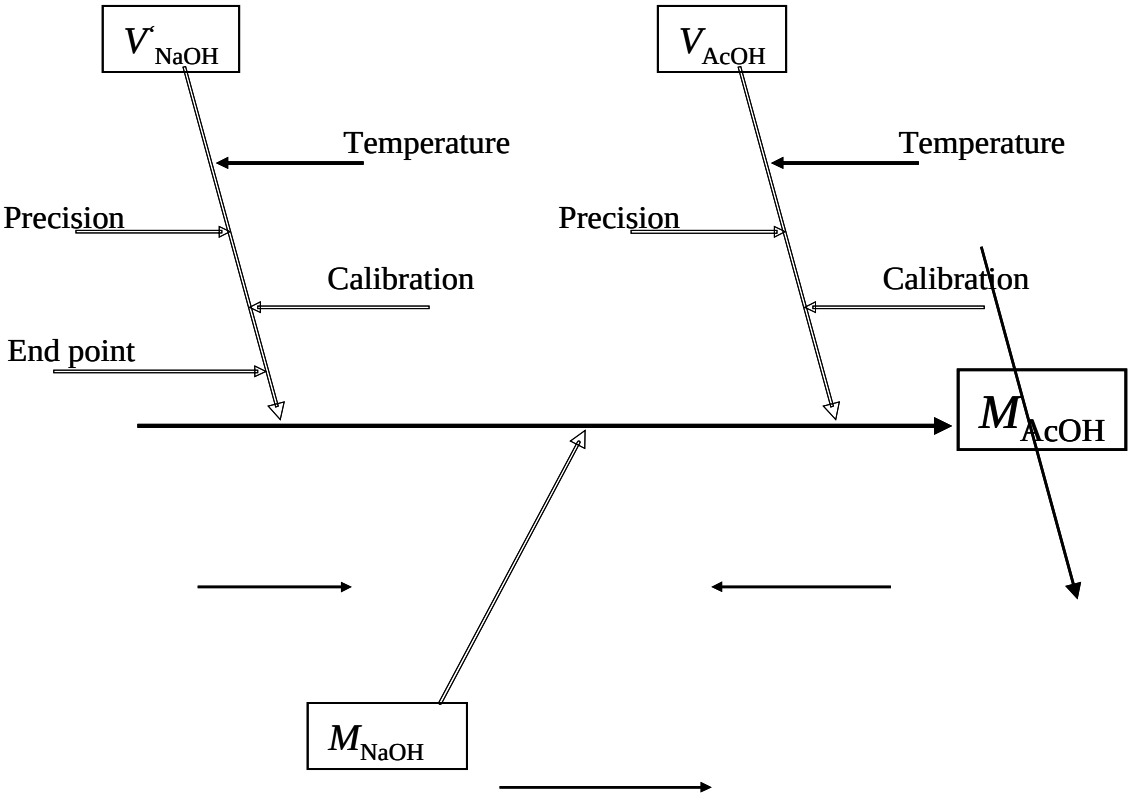


Figure 8: Cause-and-effect diagram for the uncertainty of the concentration of an acetic acid solution by titration

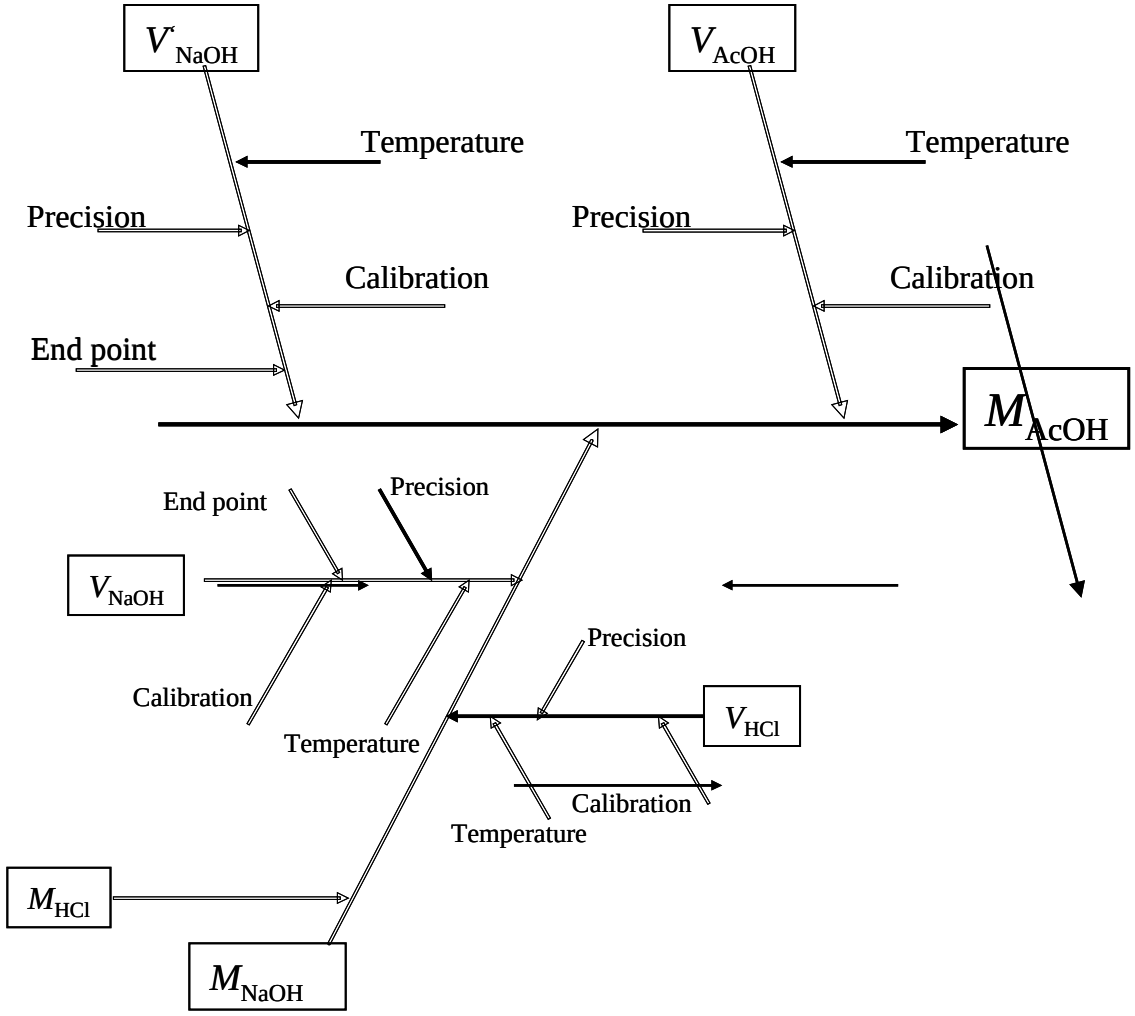


Figure 9: Cause-and-effect diagram for the uncertainty of the concentration of an acetic acid solution by titration including the uncertainty of the sodium hydroxide solution.

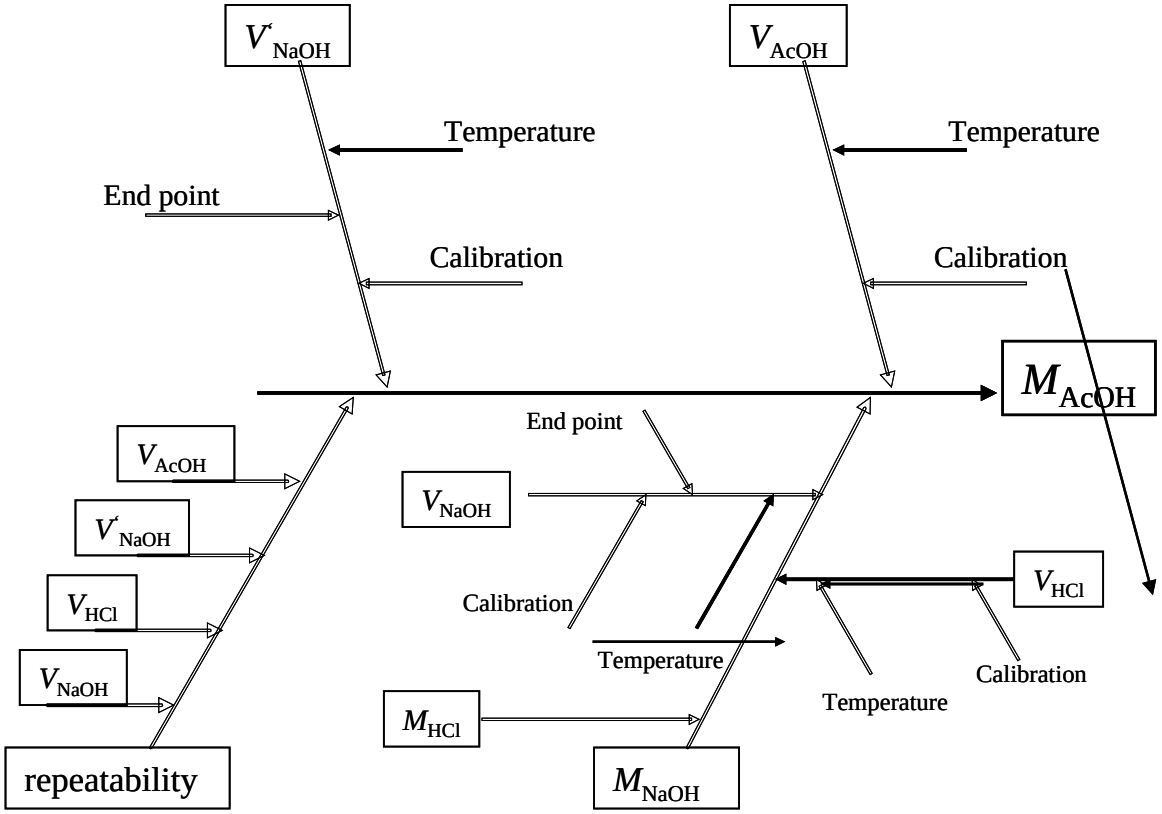


Figure 10: Cause-and-effect diagram for the uncertainty of the concentration of an acetic acid solution by titration with repeatability treated separately.

In the GUM nomenclature, the repeatability of the results is a Type A uncertainty, while any estimates of other effects, such as the calibration of the pipettes, and burettes, errors in estimating the end points and changes in temperature are classed as Type B effects. It comes as a surprise to students that a 25 mL pipette might not deliver 25.00 mL even when correctly filled to the mark. Few courses require students to calibrate their glassware (as I recall doing in the 1970s), so any calibration bias must be included in the uncertainty. The manufacturer gives the tolerance on a 25 mL pipette as ± 0.03 mL, which taken as a rectangular distribution leads to $u = 0.03/\sqrt{3} = 0.017$ mL. The temperature effect can be calculated from an estimate of the temperature fluctuations in a laboratory. The students are told that the 95% confidence interval (i.e. ± 2 $\sigma$) on the temperature in the laboratory was ± 3 °C, and the volume coefficient of water is 0.00021 °C$^{-1}$. Therefore $u = 3 \times 0.00021 \times 25 \times ½ = 0.008$ mL. These combine as

$$u_c = \sqrt{0.017^2 + 0.008^2} = 0.019\,\text{mL} \qquad (4)$$

A similar calculation for the volume delivered by the burette has the tolerance ± 0.05 mL ($u = 0.05/\sqrt{3} = 0.029$ mL). The greatest uncertainty is in the estimation of the end point which after some discussion is agreed to be 0.1 to 0.2 mL. If we take a rectangular distribution with $a = 0.15$ mL, $u = 0.15/\sqrt{3} = 0.087$ mL, and the combined uncertainty is

$$u_c = \sqrt{0.029^2 + 0.087^2 + 0.008^2} = 0.092\,\text{mL} \qquad (5)$$

The relative uncertainty of the molarity of NaOH (Equation 2) is

$$\frac{u_{M_{\mathrm{NaOH}}}}{M_{\mathrm{NaOH}}} = \sqrt{\left(\frac{u_{V_{\mathrm{NaOH}}}}{V_{\mathrm{NaOH}}}\right)^2 + \left(\frac{u_{V_{\mathrm{HCl}}}}{V_{\mathrm{HCl}}}\right)^2 + \left(\frac{u_{M_{\mathrm{HCl}}}}{M_{\mathrm{HCl}}}\right)^2 + \left(\frac{s_r}{25}\right)^2} \tag{6}$$

where the uncertainty in the standard HCl is zero (we have no information on this, and as everyone used the same solution, any error will be as a bias in all results and not contribute to the reproducibility.) Therefore

$$\frac{u_{M_{\mathrm{NaOH}}}}{M_{\mathrm{NaOH}}} = \sqrt{\left(\frac{0.092}{25}\right)^2 + \left(\frac{0.019}{25}\right)^2 + 0 + (0.0033)^2} = 0.0050 \tag{7}$$

A similar calculation for the uncertainty of the molarity of acetic acid is

$$\begin{aligned}\frac{u_{M_{\mathrm{AcOH}}}}{M_{\mathrm{AcOH}}} &= \sqrt{\left(\frac{u_{V'_{\mathrm{NaOH}}}}{V'_{\mathrm{NaOH}}}\right)^2 + \left(\frac{u_{V_{\mathrm{AcOH}}}}{V_{\mathrm{AcOH}}}\right)^2 + \left(\frac{u_{M_{\mathrm{NaOH}}}}{M_{\mathrm{NaOH}}}\right)^2 + \left(\frac{s_r}{25}\right)^2} \\ &= \sqrt{\left(\frac{0.092}{25}\right)^2 + \left(\frac{0.019}{25}\right)^2 + (0.0050)^2 + (0.0033)^2} = 0.0071\end{aligned} \tag{8}$$

A bright student might point out that the calibration error of the burette should cancel between the two titrations. Although adding this component two times overestimates the uncertainty (Hibbert, 2003), the effect is not great and this might be seen as an unnecessary complication. The relative uncertainty of the molarity of the acetic acid is thus estimated to be 0.71%, which agrees well with the range of RSD% found for the teams. The process is very instructive for causing the students to think about sources of error, and it becomes clear that there are only two major variances, the repeatability and the end point uncertainty. Combining two repeatabilities for the two titrations and two end point errors gives:

$$\frac{u_{M_{\mathrm{AcOH}}}}{M_{\mathrm{AcOH}}} \approx \sqrt{2\times(0.0033)^2 + 2\times\left(\frac{0.087}{25}\right)^2} = 0.0068 \tag{9}$$

which is a quick, and entirely appropriate estimate. The moral is that if you are trying to reduce the uncertainty of an analysis, it is only worth tackling the one or two greatest sources.

# 4 Discussion

As has been shown, the data from this competition is an ideal pedagogical tool for introducing nearly all the important concepts in statistical data analysis. (In my course only calibration is not touched by these data). Apart from being used to teach the manipulations of data using a spreadsheet such as Excel, the results can be used as a basis for discussions about the practice of analytical chemistry. Why were 23 of the 78 or 29% (accidentally the same as the LGC figure) outside a reasonable estimate of the uncertainty? Why does this appear to fit in with the ¼ to ⅓ findings of NIST and the LGC? Some of the more errant values, as found in the wider studies are not chemical errors, but misplaced decimal points, or transcription errors, but as I say to

the students there is no help in a pathologist explaining to grieving relatives that the dosage recommended was right, except that it was ten times too much!

The nature of random error can be discussed. What is the 'bell shaped curve' of figure 4 telling us? If the competition were held a week later, we would expect the same distribution of results (this is the raison d'etre of statistics), but would the same teams be best and worst? In other words, is the distribution one of abilities in titration, or is it a random measurement uncertainty on the day? If it is the latter, should all the teams who are within the expected measurement uncertainty receive a prize, as the team that happens to have the best score on the day could find themselves less well placed another time?

With a reproducibility of less than 0.5 %, even in the hands of relatively unskilled chemists, a titration is still one of the most accurate analytical methods. The results are traceable to the international system of units (SI) if appropriately traceable standard solutions are used. Someone might ask how we know the assigned values are correct? This is a good question, as they are determined by repeat titrations by the organisers of the competition, who might be expected to provide accurate results, but they do not go to much greater lengths than the competitors and do not provide uncertainty statements. As the judges' answer is final, this is a good example of assigned values rather than demonstrably traceable results. However, the good agreement between the consensus means and the assigned values suggest they are not too far out.

# 5 Conclusions

High School titration and analysis competitions have been good ambassadors for chemistry and their usefulness can be extended to university courses. Data from the 1997 RACI competition round held at the University of New South Wales has provided a complete environment for teaching modern data analysis. Students who have been exposed to this approach since 2000, have shown a greater motivation in wanting to understand what the data and their interpretation means. The impact of the approach is augmented when coupled with a practical course for which uncertainties have to be estimated, and towards the end of the session, results are analysed in a similar manner. Finally the importance of quality analytical measurements is stressed, hopefully giving students a proper attitude to their work when they join the workforce.

**Acknowledgments**

The author gratefully thanks Dr Peter Chia, who organized the RACI competition at UNSW in 1997, for allowing him to use and publish the data.

**References**

Hibbert, D. (2005), Interlaboratory Studies, *Encyclopedia of Analytical Science*, eds. P. J. Worsfold, A. Townshend and C. F. Poole, Oxford, Elsevier **7:** 449-57.

Hibbert, D. B. (2003), The measurement uncertainty of ratios which share uncertainty components in numerator and denominator, *Accreditation and Quality Assurance*, **8**: 195-199.

ISO (1993), *Guide to the Expression of Uncertainty in Measurement*, Geneva, Switzerland, ISO.

ISO (1999), *ISO 17025 General requirements for the competence of testing laboratories*, Geneva, Switzerland, ISO.

ISO (2005), *International vocabulary of basic and general terms in metrology*, Geneva, Switzerland, ISO.

King, B. (1995), Quality in the Analytical Laboratory, *Quality Assurance and TQM for Analytical Laboratories*, ed. M. Parkany, Cambridge, The Royal Society of Chemistry**:** 8 - 18.

May, W. (2001), E-Health and Technology: Empowering Consumers in a Simpler, More Cost Effective Health Care System, NIST, Gaithersberg, http://www.nist.gov/testimony/2001/wmehealt.htm.

Miller, J. N. and J. C. Miller (2005), *Statistics and Chemometrics for Analytical Chemistry*, 5th edition, Pearson Education Ltd., Harlow.

RSC (2005), Schools Analytical Competition, Royal Society of Chemistry, Cambridge, http://www.rsc.org/lap/rsccom/dab/adcomps.htm.

Saed Al-Deen, T., D. B. Hibbert, et al. (2004), An uncertainty budget for the determination of the purity of glyphosate by quantitative nuclear magnetic resonance (QNMR) spectroscopy, *Accreditation and Quality Assurance*, **9** (1-2): 55 - 63.