

# Chem 321 Lecture 5 - Experimental Errors and Statistics

9/10/13

## Student Learning Objectives

### Experimental Errors and Statistics

#### Calibration Results for a 2.0-mL Transfer Pipet

1.998 mL	1.991 mL
2.001 mL	1.999 mL
2.003 mL	1.998 mL
1.997 mL	2.002 mL

On what basis should one or more of the data points be rejected?

Usually in a series of measurements one or more values deviate significantly from the general cluster. Take, for example, the 1.991 mL pipet calibration result above. It is very appealing to reject this data point because of the lack of agreement with the other values. However, data should never be discarded simply to reduce the standard deviation or RMD and make the measurements appear to be more reproducible. Often such outlying values can be attributed to some flaw in the experiment. For example, if you know that some precipitate was lost during the analysis of one sample or that the meniscus was just below the calibration line when filling the transfer pipet for one delivery, rejection of the result in such cases is justified. However, if there is no obvious explanation for the disparate result, it is useful to apply a statistical test to decide whether to retain or to reject the suspect data point. A number of statistical tests have been suggested and used to determine whether a measurement should be rejected, however, data rejection tests must be used judiciously, especially for small data sets such as your typical lab results. One such test that is easy to apply and that works reasonably well for rejecting one data point from a small data set is the Q test. For this test a quotient (Q) is calculated, then  $Q_{\text{calc}}$  is compared to tabulated values ( $Q_{\text{tab}}$ ); see table below for values of  $Q_{\text{tab}}$  (referred to as  $Q_{\text{crit}}$  in the table) at various confidence levels. If  $Q_{\text{calc}} > Q_{\text{tab}}$ , you can be confident that the suspect data point should be rejected.

$Q_{\text{calc}}$  is determined in the following way.

$$Q_{\text{calc}} = \frac{|\text{suspect value} - \text{closest value}|}{\text{largest value} - \text{smallest value}}$$

Critical Values for the Rejection Quotient, $Q^*$			
Number of Observations	$Q_{crit}$ (Reject if $Q > Q_{crit}$ )		
	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

\*Reprinted with permission from D. B. Rorabacher, *Anal. Chem.*, **1991**, *63*, 139. Copyright 1991 American Chemical Society.

© 2004 Thomson - Brooks/Cole

For the 1.991 mL value that is suspect,

$$Q_{calc} = \frac{|1.991 - 1.997|}{2.003 - 1.991} = 0.5$$

At the 90% confidence level,  $Q_{tab} = 0.468$  for 8 measurements, so  $Q_{calc} > Q_{tab}$  and this value should be rejected. However, it cannot be rejected at the 95% confidence level. Remember that only one data point can be discarded per data set by the Q test.

Another statistical test that is often used with larger data sets is based on the normal error curve. First, calculate the average and standard deviation without the suspect value. Then determine if the suspect point is within  $\pm 3\sigma$  of the mean. If it is not, reject the value.

#### Check for Understanding 2.4

Solutions

1. What is the average, RMD and 90% confidence interval for the remaining 7 pipet calibration results?
2. What is the statistical basis for rejecting a point by a  $\pm 3\sigma$  test?

## Comparing Replicate Measurements

Imagine that you have calibrated your 1-mL autopipet and from 5 measurements you get a mean value of 0.998 mL with a standard deviation of 0.004 mL. Another student calibrates this same pipet and from 7 measurements gets a mean value and standard deviation of  $1.007 \pm 0.006$  mL. Are these two results statistically different, or could these results occur by chance? This question can be answered using a  $t$  test in which you compare  $t_{\text{calc}}$  with  $t_{\text{table}}$ . If  $t_{\text{calc}} > t_{\text{table}}$  then the difference is significant with a certain level of confidence. For two sets of data consisting of  $n_1$  and  $n_2$  measurements with averages  $\bar{x}_1$  and  $\bar{x}_2$  and standard deviations  $s_1$  and  $s_2$ ,  $t_{\text{calc}}$  is obtained by:

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

where

$$s_{\text{pooled}} = \sqrt{\frac{\sum(x_i - \bar{x}_1)^2 + \sum(x_j - \bar{x}_2)^2}{n_1 + n_2 - 2}} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$$

and  $x_i$  and  $x_j$  represent individual values for the first and second set of data, respectively.

The value of  $t_{\text{table}}$  is gotten from the  $t$  table for  $n_1 + n_2 - 2$  degrees of freedom. These calculations assume that  $s_1$  and  $s_2$  are not significantly different from each other. This is determined by using the  $F$  test, where  $F$  is defined as:

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

Note that  $s_1$  is chosen to be larger than  $s_2$  so that  $F$  is always  $\geq 1$ .  $F_{\text{calc}}$  is then compared to  $F_{\text{table}}$  (see  $F$  Table below), and if  $F_{\text{calc}} < F_{\text{table}}$ , then the standard deviations are not significantly different. However, if  $F_{\text{calc}} > F_{\text{table}}$ , the standard deviations are significantly different, and a different formula must be used to get  $t_{\text{calc}}$ .

**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	$\infty$
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
$\infty$	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

© 2003 W. H. Freeman and Company

For the two sets of autopipet calibration data,

$$F_{calc} = \left( \frac{0.006}{0.004} \right)^2 = 2.2 < F_{table} = 6.16 \quad \text{(for 6 degrees of freedom in the numerator and 4 degrees of freedom in the denominator)}$$

so the standard deviations are not significantly different and

$$s_{pooled} = \sqrt{\frac{0.006^2(6) + 0.004^2(4)}{7 + 5 - 2}} = 0.005_3$$

$$t_{calc} = \frac{|1.007 - 0.998|}{0.005_3} \sqrt{\frac{7 \cdot 5}{7 + 5}} = 2.9 > t_{table} = 2.228$$

Since  $t_{\text{calc}} > t_{\text{table}}$  this means that the difference in the two calibration values is significant at the 95% confidence level.

### Propagation of Errors

Generally an analysis result is based on several different measurements, each with its own error. The overall result should be reported to a number of significant figures that properly reflects the errors associated with the various measurements. Two rules of error propagation apply.

Addition/Subtraction Rule:

If only the operations of addition and subtraction are involved, the absolute error in the result ( $e$ ) is calculated as in the following example.

$$(A \pm e_A) + (B \pm e_B) - (C \pm e_C) = (A + B - C) \pm e$$

$$e = \sqrt{e_A^2 + e_B^2 + e_C^2}$$

Note that the absolute error (uncertainty) in the result is generally reported with one significant figure. The result is then rounded off to the same place as this figure.)

Multiplication/Division Rule:

If only the operations of multiplication and division are involved, the absolute error in the result ( $e$ ) is calculated as in the following example.

$$\frac{(A \pm e_A)(B \pm e_B)}{(C \pm e_C)} = \frac{A \times B}{C} \pm e$$

$$e = \left( \frac{A \times B}{C} \right) \sqrt{\left( \frac{e_A}{A} \right)^2 + \left( \frac{e_B}{B} \right)^2 + \left( \frac{e_C}{C} \right)^2}$$

Recall that the relative error equals the absolute error divided by the quantity associated with the absolute error. For example,

$$e_A^{rel} = \frac{e_A}{A}$$

Consequently, the relative error associated with a result gotten by multiplication/division can be calculated by:

$$e^{rel} = \sqrt{(e_A^{rel})^2 + (e_B^{rel})^2 + (e_C^{rel})^2}$$

This expression emphasizes the importance of relative errors when rounding off results to the proper number of significant figures. The measurement with the largest relative error primarily determines the relative error in the overall result. In general, an analysis result should be rounded off so that the implied relative error is comparable to the largest relative error associated with the experiment. For a calculation involving mixed operations (both multiplication/division and addition/subtraction), apply one of the above rules, as appropriate, to each step as you work through the calculation. Do not round off the intermediate results, however, take note of the appropriate number of significant figures for the intermediate result.

### Check for Understanding 2.5

Solutions

1. How should an experimental result of 29.6518% be rounded off to reflect at least 4 ppt relative error in the measuring process?
2. For the following, calculate the absolute error and use this to round off the result to the appropriate number of significant figures.
  - a)  $(36.2 \pm 0.4)/(27.1 \pm 0.6)$
  - b)  $(25.0 \pm 0.1)(0.0215 \pm 0.0003) - (1.02 \pm 0.01)$