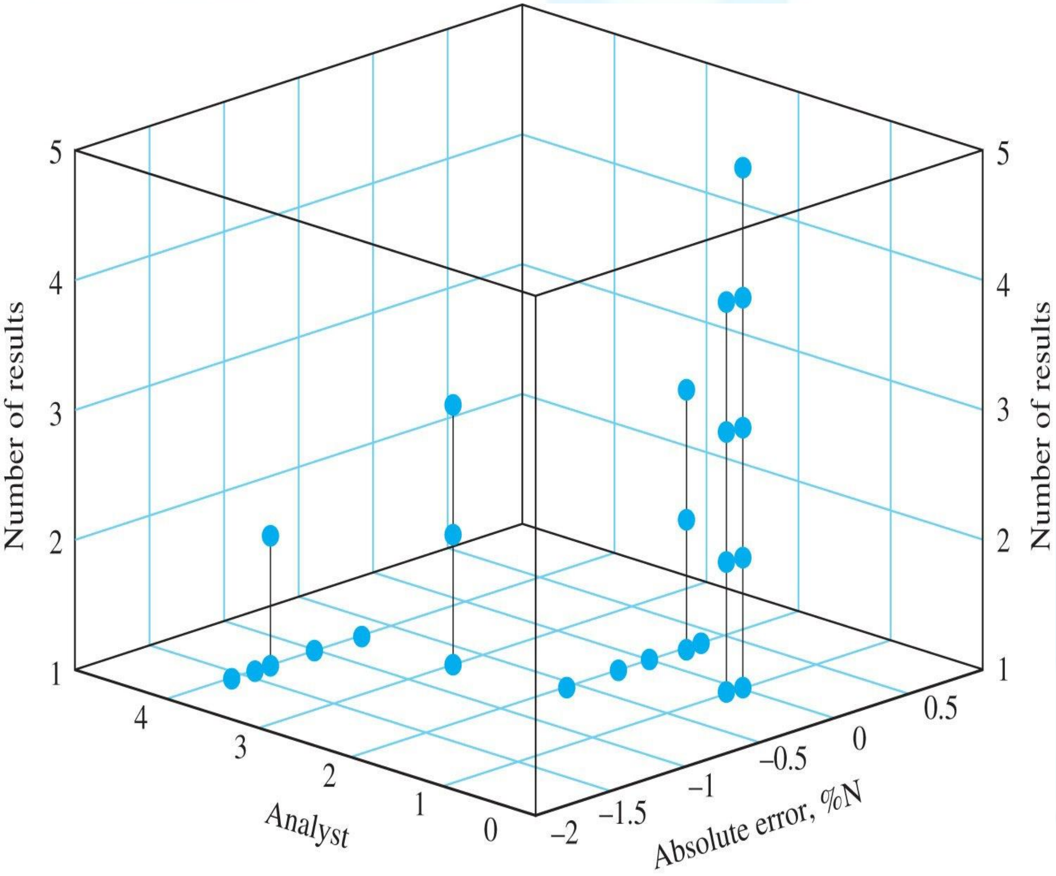
Chapter 6:

Random Errors in Chemical Analysis

Source
slideplayer.com/Fundamentals of Analytical Chemistry, F.J. Holler, S.R.Crouch

- **Random errors are present in every measurement no matter how careful the experimenter.**
- Random, or indeterminate, errors can never be totally eliminated and are often the major source of uncertainty in a determination.
- Random errors are caused by the many uncontrollable variables that accompany every measurement.
- The accumulated effect of the individual uncertainties causes replicate results to fluctuate randomly around the mean of the set.
- In this chapter, we consider the
  - sources of random errors,
  - the determination of their magnitude, and
  - their effects on computed results of chemical analyses.
  - We also introduce the significant figure convention and illustrate its use in reporting analytical results.

6A The nature of random errors



- random error in the results of analysts 2 and 4 is much larger than that seen in the results of analysts 1 and 3.
- The results of analyst 3 show outstanding precision but poor accuracy. The results of analyst 1 show excellent precision and good accuracy.

Figure 6-1 A three-dimensional plot showing absolute error in Kjeldahl nitrogen determinations for four different analysts.

Random Error Sources

- Small undetectable uncertainties produce a detectable random error in the following way.

- Imagine a situation in which just four small random errors combine to give an overall error. We will assume that each error has an equal probability of occurring and that each can cause the final result to be high or low by a fixed amount ±*U.*

- *Table 6.1 gives a*ll the possible ways in which four errors can combine to give the indicated deviations from the mean value.

TABLE 6-1

Possible Combinations of Four Equal-Sized Uncertainties

Combinations of Uncertainties	Magnitude of Random Error	Number of Combinations	Relative Frequency
$+U_1+U_2+U_3+U_4$	$+4U$	1	1/16 = 0.0625
$-U_1+U_2+U_3+U_4$ $+U_1-U_2+U_3+U_4$ $+U_1+U_2-U_3+U_4$ $+U_1+U_2+U_3-U_4$	$+2U$	4	4/16 = 0.250
$-U_1-U_2+U_3+U_4$ $+U_1+U_2-U_3-U_4$ $+U_1-U_2+U_3-U_4$ $-U_1+U_2-U_3+U_4$ $-U_1+U_2+U_3-U_4$ $+U_1-U_2-U_3+U_4$	0	6	6/16 = 0.375
$+U_1-U_2-U_3-U_4$ $-U_1+U_2-U_3-U_4$ $-U_1-U_2+U_3-U_4$ $-U_1-U_2-U_3+U_4$	$-2U$	4	4/16 = 0.250
$-U_1-U_2-U_3-U_4$	$-4U$	1	1/16 = 0.0625

* Note that only 1 combination leads to a deviation of +4 *U*, *4 combinations give a deviation of + 2 U, and 6 give a* deviation of 0 *U.*

* This ratio of 1:4:6:4:1 is a measure of the probability for a deviation of each magnitude

If we make a sufficiently large number of measurements, we can expect a frequency distribution like that shown in Figure below.

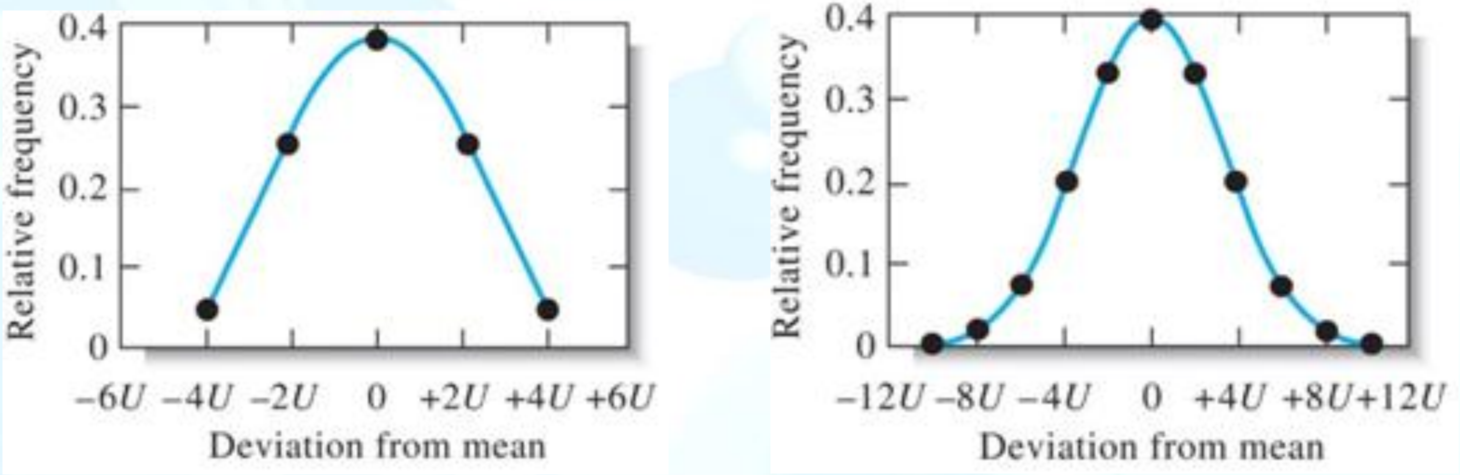


Figure 6-2 Frequency distribution for measurements containing
(a) Four random uncertainties, (b) ten random uncertainties,

The most frequent occurrence is zero deviation from the mean.

At the other extreme, a maximum deviation of 10 U occurs only about once in 500 results.

When the same procedure is applied to a very large number of random uncertainties, a bell-shaped curve results. Such a plot is called a Gaussian curve or a normal error curve.

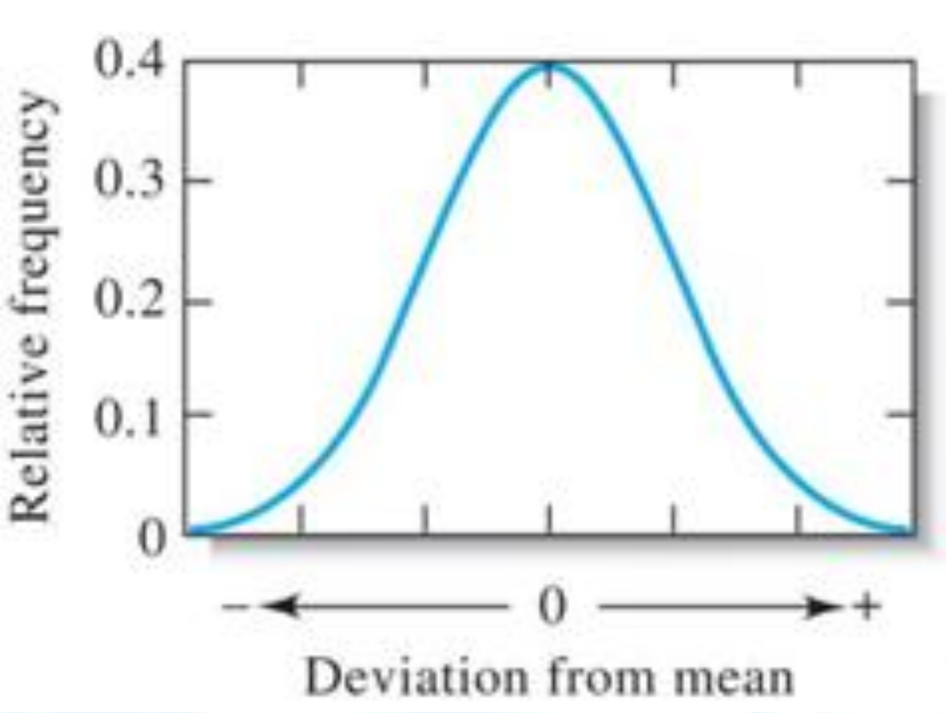


Figure 6-2-c Frequency distribution for measurements containing a very large number of random uncertainties,

Distribution of Experimental Results

The distribution of replicate data from most quantitative analytical experiments approaches that of the Gaussian curve.

Exp: Calibration of a 10-mL pipet.

In this experiment a small flask and stopper were weighed.

Ten milliliters of water were transferred to the flask with the pipet, and the flask was stoppered. The flask, the stopper, and the water were then weighed again.

The temperature of the water was also measured to determine its density.

The mass of the water was then calculated by taking the difference between the two masses.

The mass of water divided by its density is the volume delivered by the pipet. The experiment was repeated 50 times.

TABLE 6-2†

	A	B	C	D	E	F	G	H
1	Replicate Data for the Calibration of a 10-mL Pipet*							
2	**Trial**	**Volume, mL**		**Trial**	**Volume, mL**		**Trial**	**Volume, mL**
3	1	9.988		18	9.975		35	9.976
4	2	9.973		19	9.980		36	9.990
5	3	9.986		20	9.994		37	9.988
6	4	9.980		21	9.992		38	9.971
7	5	9.975		22	9.984		39	9.986
8	6	9.982		23	9.981		40	9.978
9	7	9.986		24	9.987		41	9.986
10	8	9.982		25	9.978		42	9.982
11	9	9.981		26	9.983		43	9.977
12	10	9.990		27	9.982		44	9.977
13	11	9.980		28	9.991		45	9.986
14	12	9.989		29	9.981		46	9.978
15	13	9.978		30	9.969		47	9.983
16	14	9.971		31	9.985		48	9.980
17	15	9.982		32	9.977		49	9.984
18	16	9.983		33	9.976		50	9.979
19	17	9.988		34	9.983			
20	*Data listed in the order obtained							
21	Mean	9.982		Maximum	9.994			
22	Median	9.982		Minimum	9.969			
23	Std. Dev.	0.0056		Spread	0.025			

†For Excel calculations of the statistical quantities listed at the bottom of Table 6-2, see S. R. Crouch and F. J. Holler, *Applications of Microsoft® Excel in Analytical Chemistry*, 2nd ed., Belmont, CA: Brooks/Cole, 2014, ch. 2.

Consider the data in the table for the calibration of a 10-mL pipet.

The results vary from a low of 9.969 mL to a high of 9.994 mL.

This 0.025-mL spread of data results directly from an accumulation of all random uncertainties in the experiment.

The spread in a set of replicate measurements can be defined as the difference between the highest and lowest result.

The frequency distribution data is shown in the table.

TABLE 6-3

Frequency Distribution of Data from Table 6-2

Volume Range, mL	Number in Range	% in Range
9.969 to 9.971	3	6
9.972 to 9.974	1	2
9.975 to 9.977	7	14
9.978 to 9.980	9	18
9.981 to 9.983	13	26
9.984 to 9.986	7	14
9.987 to 9.989	5	10
9.990 to 9.992	4	8
9.993 to 9.995	1	2
	Total = 50	Total = 100%

The frequency distribution data are plotted as a bar graph, or histogram.

As the number of measurements increases, the histogram approaches the shape of the continuous Gaussian curve.

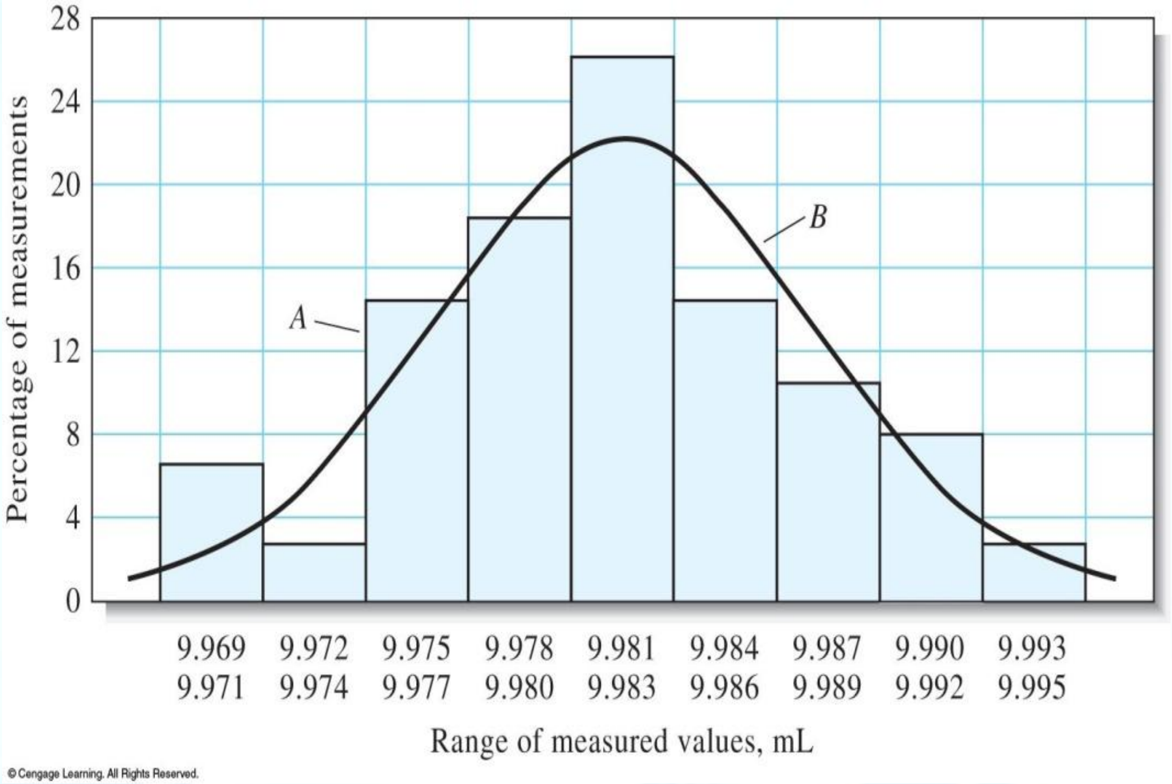


Figure 6-3 A histogram (A) showing distribution of the 50 results in Table 6-3 and a Gaussian curve (B) for data having the same mean and standard deviation as the data in the histogram.

6B Statistical treatment of random errors

- Statistical analysis only reveals information that is already present in a data set.

-Statistical methods, do allow us to categorize and characterize data and to make objective and intelligent decisions about data quality and interpretation.

-Samples and Populations

A population is the collection of all measurements of interest to the experimenter, while a sample is a subset of measurements selected from the population.

Statistical sample is different from the analytical sample.

Properties of Gaussian Curves

Gaussian curves can be described by an equation that contains two parameters, the population mean μ and the population standard deviation σ.

The term parameter refers to quantities such as μ and σ that define a population or distribution. Data values such as x are variables.

The term statistic refers to an estimate of a parameter that is made from a sample of data.

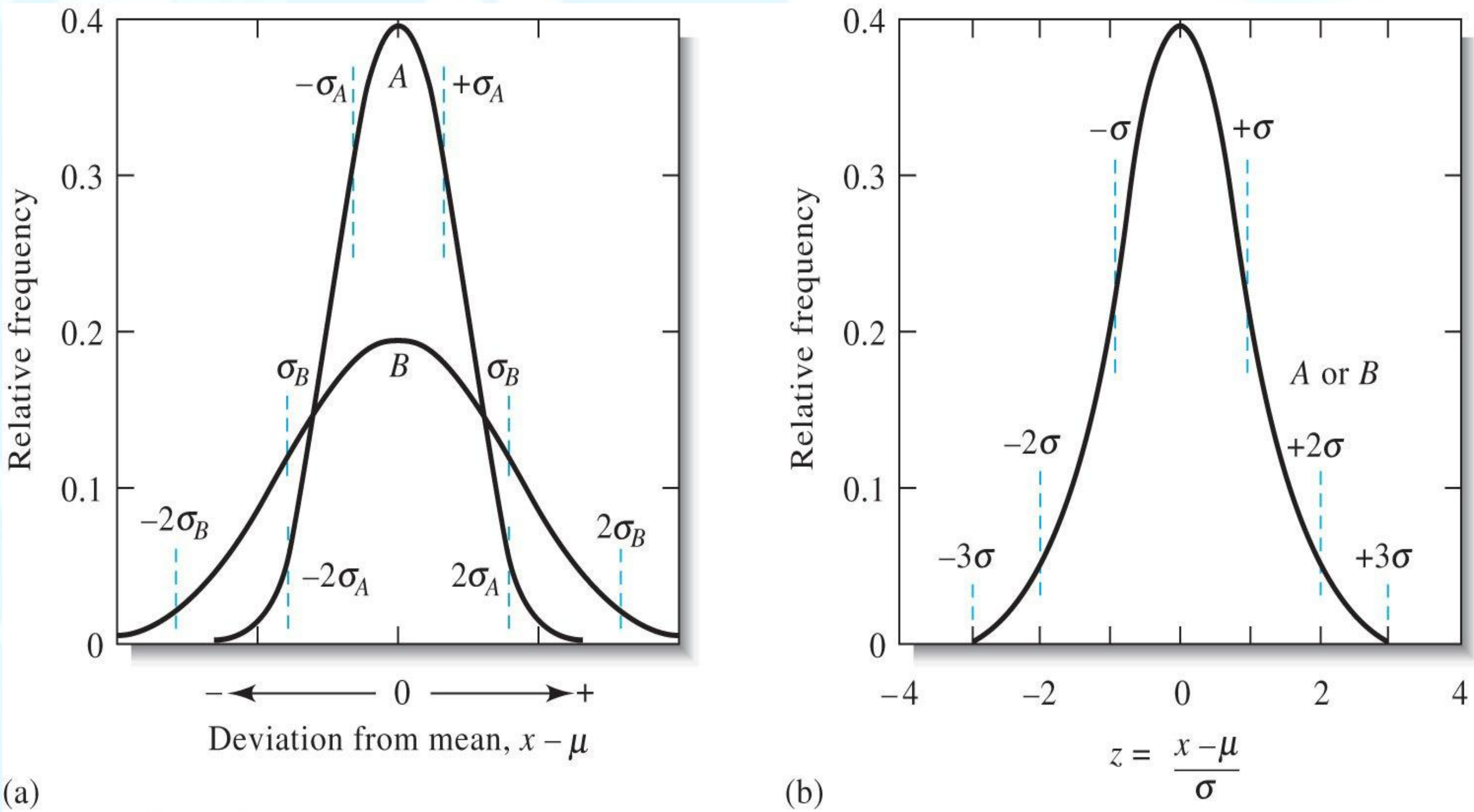


Figure 6-4 Normal error curves.

The equation for a normalized Gaussian curve is as follows:

$$y = \frac{e^{-(x-\mu)^{2/2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

The Population Mean μ *and the Sample Mean Ẍ*

➢ The sample mean x is the arithmetic average of a limited sample drawn from a population of data. The sample mean is defined as the sum of the measurement values divided by the number of measurements.

$$\bar{x} = \frac{\sum_{i=1}^{N} x_1}{N}$$

N is the is the no. of measurements in the sample set

➢ The population mean μ is expressed as:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

where N is the total number of measurements in the population.

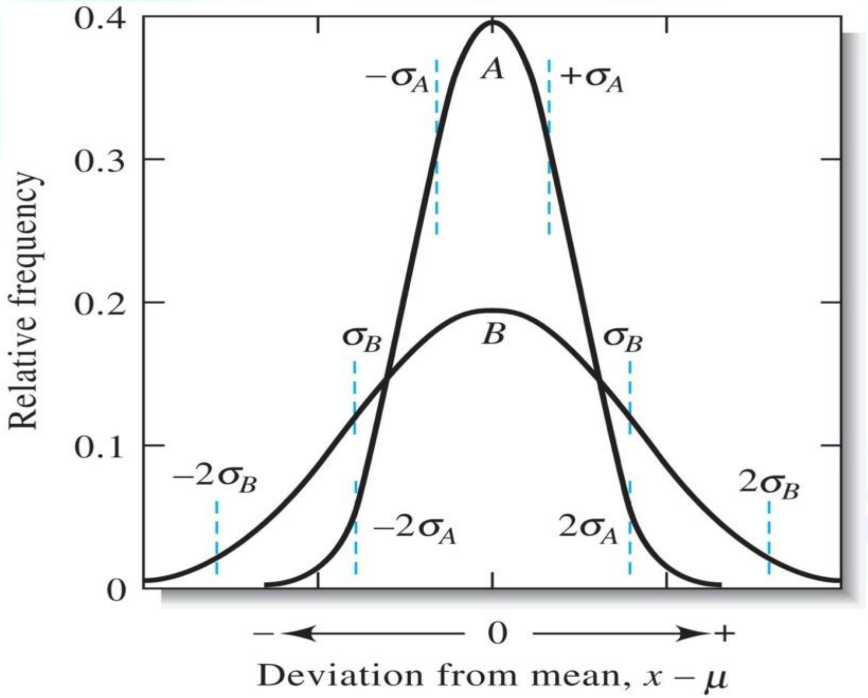
The Population Standard Deviation σ

It is a measure of the precision of the population and is expressed as:

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

Where N is the no. of data points making up the population.

The two curves are for two populations of data that differ only in their standard deviations.



(a)

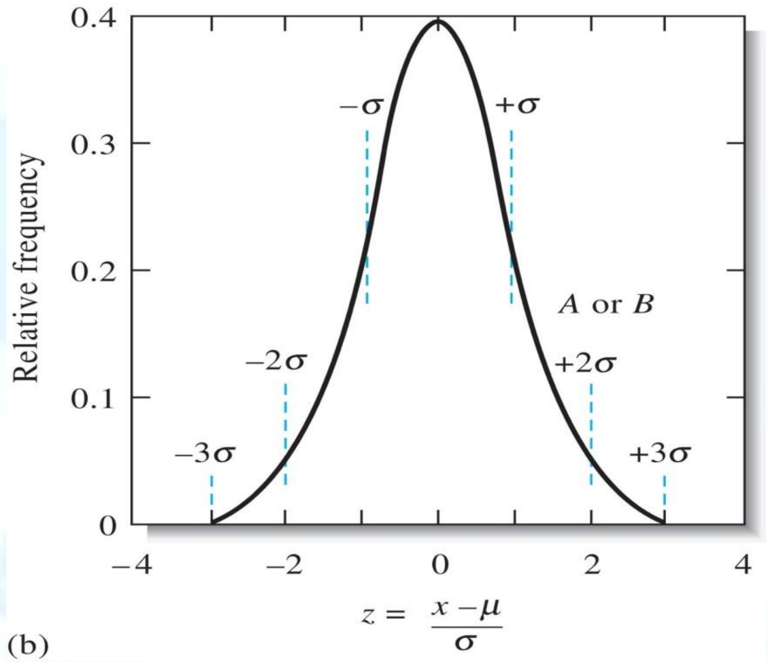
Another type of normal error curve in which the x axis is a new variable z, the quantity z represents the deviation of a result from the population mean relative to the standard deviation.

It is commonly given as a variable in statistical tables since it is a dimensionless quantity.

z is defined as

$$z = \frac{(x - \mu)}{\sigma}$$

A plot of relative frequency versus z yields a single Gaussian curve that describes all populations of data regardless of standard deviation.



The equation for the Gaussian error curve is:

$$y = \frac{e^{-(x-\mu)^{2/2\sigma^2}}}{\sigma\sqrt{2\pi}} = \frac{e^{-z^2/2}}{\sigma\sqrt{2\pi}}$$

This curve has several general properties:

(a) The mean occurs at the central point of maximum frequency,

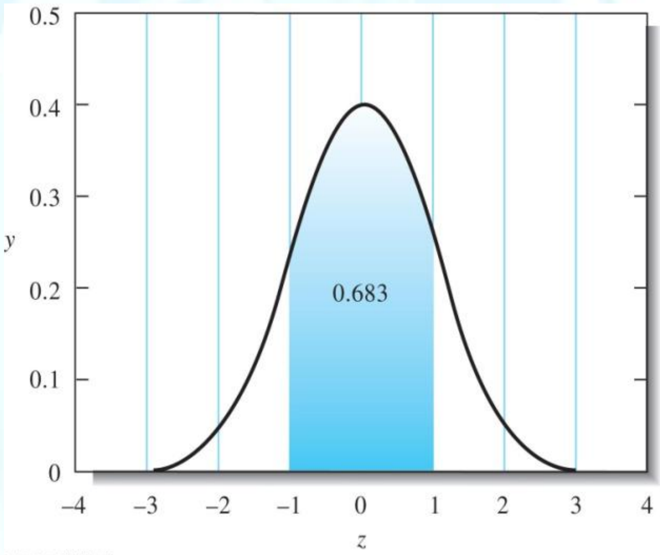
(b) there is a symmetrical distribution of positive and negative deviations about the maximum, and

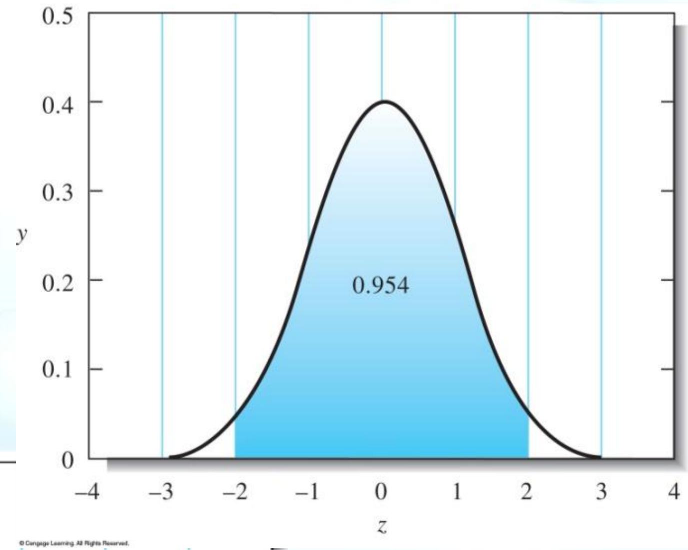
(c) there is an exponential decrease in frequency as the magnitude of the deviations increases.

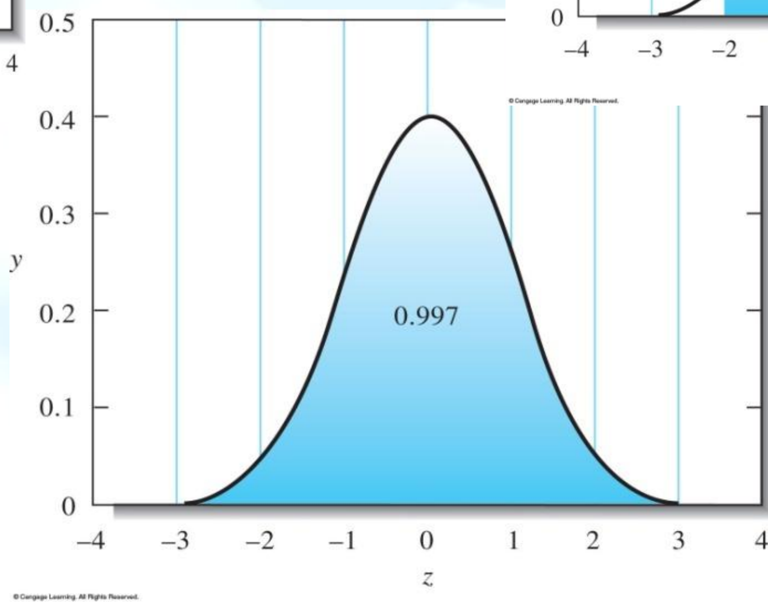
Areas under a Gaussian Curve

Regardless of its width, 68.3% of the area beneath a Gaussian curve for a population lies within one standard deviation (1s) of the mean m.

Approximately 95.4% of all data points are within $\pm 2\sigma$ of the mean and 99.7% within $\pm 3\sigma$.







The Sample Standard Deviation: A Measure of Precision

The sample standard deviation s is expressed as:

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

- This equation applies to small sets of data.

-The number of degrees of freedom indicates the number of independent results that enter into the computation of the standard deviation.

- When number of degrees of freedom, (N - 1) is used instead of N, s is said to be an unbiased estimator of the population standard deviation s.

- The sample variance s^2 is an estimate of the population variance σ^2.

An Alternative Expression for Sample Standard Deviation

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

- Any time you subtract two large, approximately equal numbers, the difference will usually have a relatively large uncertainty.

- Hence, you should never round a standard deviation calculation until the end.

- Because of the uncertainty in x, a sample standard deviation may differ significantly from the population standard deviation.

- As N becomes larger, x and s become better estimators of μ, and σ.

EXAMPLE 6-1

The following results were obtained in the replicate determination of the lead content of a blood sample: 0.752, 0.756, 0.752, 0.751, and 0.760 ppm Pb. Find the mean and the standard deviation of this set of data.

Solution

To apply Equation 6-5, we calculate $\sum x_i^2$ and $(\sum x_i)^2/N$.

Sample	x_i	x_i^2
1	0.752	0.565504
2	0.756	0.571536
3	0.752	0.565504
4	0.751	0.564001
5	0.760	0.577600
	$\sum x_i = 3.771$	$\sum x_i^2 = 2.844145$

$$\bar{x} = \frac{\sum x_i}{N} = \frac{3.771}{5} = 0.7542 \approx 0.754 \text{ ppm Pb}$$

$$\frac{\left(\sum x_i\right)^2}{N} = \frac{(3.771)^2}{5} = \frac{14.220441}{5} = 2.8440882$$

Substituting into Equation 6-5 leads to

$$s = \sqrt{\frac{2.844145 - 2.8440882}{5 - 1}} = \sqrt{\frac{0.0000568}{4}} = 0.00377 \approx 0.004 \text{ ppm Pb}$$

Standard Error of the Mean

- If a series of replicate results, each containing N measurements, are taken randomly from a population of results, the mean of each set will show less and less scatter as N increases.

- The standard deviation of each mean is known as the standard error of the mean, s_m, is expressed as:

$$s_m = \frac{s}{\sqrt{N}}$$

- When N is greater than about 20, s is usually a good estimator of σ, and these quantities can be assumed to be identical for most purposes.

Pooling Data to Improve the Reliability of s

- If we have several subsets of data, a better estimate of the population standard deviation can be obtained by pooling (combining) the data, assuming that the samples have similar compositions and have been analyzed the same way.

- The pooled estimate of s_{pooled}, is a weighted average of the individual estimates.

- To calculate spooled, deviations from the mean for each subset are squared; the squares of the deviations of all subsets are then summed and divided by the appropriate number of degrees of freedom.

$$s_{pooled} = \sqrt{\frac{\sum_{i=1}^{N_1}(x_i - \bar{x}_i)^2 + \sum_{j=1}^{N_2}(x_j - \bar{x}_2)^2 + \sum_{k=1}(x_k - \bar{x}_3)^2 + \ldots}{N_1 + N_2 + N_3 + \ldots - N_t}}$$

EXAMPLE 6-2

Glucose levels are routinely monitored in patients suffering from diabetes. The glucose concentrations in a patient with mildly elevated glucose levels were determined in different months by a spectrophotometric analytical method. The patient was placed on a low-sugar diet to reduce the glucose levels. The following results were obtained during a study to determine the effectiveness of the diet. Calculate a pooled estimate of the standard deviation for the method.

Time	Glucose Concentration, mg/L	Mean Glucose, mg/L	Sum of Squares of Deviations from Mean	Standard Deviation
Month 1	1108, 1122, 1075, 1099, 1115, 1083, 1100	1100.3	1687.43	16.8
Month 2	992, 975, 1022, 1001, 991	996.2	1182.80	17.2
Month 3	788, 805, 779, 822, 800	798.8	1086.80	16.5
Month 4	799, 745, 750, 774, 777, 800, 758	771.9	2950.86	22.2

Total number of measurements = 24   Total sum of squares = 6907.89

Solution

For the first month, the sum of the squares in the next to last column was calculated as follows:

$$\begin{aligned}\text{Sum of squares} = &(1108 - 1100.3)^2 + (1122 - 1100.3)^2 \\ &+ (1075 - 1100.3)^2 + (1099 - 1100.3)^2 + (1115 - 1100.3)^2 \\ &+ (1083 - 1100.3)^2 + (1100 - 1100.3)^2 = 1687.43\end{aligned}$$

The other sums of squares were obtained similarly. The pooled standard deviation is then

$$s_{pooled} = \sqrt{\frac{6907.89}{24 - 4}} = 18.58 \approx 19 \text{ mg/L}$$

Note this pooled value is a better estimate of σ than any of the individual s values in the last column. Note also that one degree of freedom is lost for each of the four sets. Because 20 degrees of freedom remain, however, the calculated value of s can be considered a good estimate of σ.

Steve Horrell/Photo Researchers, Inc

A glucose analyzer.

Variance and Other Measures of Precision

Precision of analytical data may be expressed as:

Variance (s^2) is the square of the standard deviation.

The sample variance s^2 is an estimate of the population variance σ^2 and is given by:

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

Relative Standard Deviation (RSD) and Coefficient of Variation (CV)

IUPAC recommends that the symbol s_r be used for relative sample standard deviation and s_r for relative population standard deviation.

In equations where it is cumbersome to use RSD, use s_r and s_r.

$$RSD = s_r = \frac{s}{\overline{x}}$$

$$RSD_{(in\,ppt)} = \frac{s}{\bar{x}} \times 1000\,ppt$$

The relative standard deviation multiplied by 100% is called the coefficient of variation (CV).

$$CV = RSD\,in\,percent = \frac{s}{\bar{x}} \times 100\%$$

- Spread or Range (w) is used to describe the precision of a set of replicate results. It is the difference between the largest value in the set and the smallest.

EXAMPLE 6-3

For the set of data in Example 6-1, calculate (a) the variance, (b) the relative standard deviation in parts per thousand, (c) the coefficient of variation, and (d) the spread.

Solution

In Example 6-1, we found

$$\bar{x} = 0.754 \text{ ppm Pb} \quad \text{and} \quad s = 0.0038 \text{ ppm Pb}$$

(a) $s^2 = (0.0038)^2 = 1.4 \times 10^{-5}$

(b) $\text{RSD} = \dfrac{0.0038}{0.754} = \times 1000 \text{ ppt} = 5.0 \text{ ppt}$

(c) $\text{CV} = \dfrac{0.0038}{0.754} \times 100\% = 0.50\%$

(d) $w = 0.760 - 0.751 = 0.009 \text{ ppm Pb}$

6C Standard deviation of calculated results

We must estimate the standard deviation of a result that has been calculated from two or more experimental data points, each of which has a known sample standard deviation.

TABLE 6-4

Error Propagation in Arithmetic Calculations

Type of Calculation	Example*	Standard Deviation of y†	
Addition or subtraction	$y = a + b - c$	$s_y = \sqrt{s_a^2 + s_b^2 + s_c^2}$	(1)
Multiplication or division	$y = a \times b/c$	$\frac{s_y}{y} = \sqrt{\left(\frac{s_a}{a}\right)^2 + \left(\frac{s_b}{b}\right)^2 + \left(\frac{s_c}{c}\right)^2}$	(2)
Exponentiation	$y = a^x$	$\frac{s_y}{y} = x\left(\frac{s_a}{a}\right)$	(3)
Logarithm	$y = \log_{10} a$	$s_y = 0.434\frac{s_a}{a}$	(4)
Antilogarithm	$y = \text{antilog}_{10}\, a$	$\frac{s_y}{y} = 2.303\, s_a$	(5)

Standard Deviation of a Sum or Difference

- The variance of a sum or difference is equal to the sum of the variances of the numbers making up that sum or difference.

- The most probable value for a standard deviation of a sum or difference can be found by taking the square root of the sum of the squares of the individual absolute standard deviations.

$$y = a(\pm s_a) + b(\pm s_b) - c(\pm s_c)$$

$$s_y^2 = s_a^2 + s_b^2 + s_c^2$$

- Hence, the standard deviation of the result s_y is

$$s_y = \sqrt{s_a^2 + s_b^2 + s_c^2}$$

- For a sum or a difference, the standard deviation of the answer is the square root of the sum of the squares of the standard deviations of the numbers used in the calculation.

Standard Deviation of a Product or Quotient

The relative standard deviation of a product or quotient is determined by the relative standard deviations of the numbers forming the computed result.

In case of,

$$y = \frac{a \times b}{c}$$

The relative standard deviation s_y/y of the result by summing the squares of the relative standard deviations of a, b, and c and then calculating the square root of the sum:

$$\frac{s_y}{y} = \sqrt{\left(\frac{s_a}{a}\right)^2 + \left(\frac{s_b}{b}\right)^2 + \left(\frac{s_c}{c}\right)^2}$$

To find the absolute standard deviation in a product or a quotient, first find the relative standard deviation in the result and then multiply it by the result.

EXAMPLE 6-4

Calculate the standard deviation of the result of

$$\frac{[14.3(\pm 0.2) - 11.6(\pm 0.2)] \times 0.050(\pm 0.001)}{[820(\pm 10) + 1030(\pm 5)] \times 42.3(\pm 0.4)} = 1.725(\pm ?) \times 10^{-6}$$

Solution

First, we must calculate the standard deviation of the sum and the difference. For the difference in the numerator,

$$s_a = \sqrt{(\pm 0.2)^2 + (\pm 0.2)^2} = \pm 0.283$$

and for the sum in the denominator,

$$s_b = \sqrt{(\pm 10)^2 + (\pm 5)^2} = 11.2$$

We may then rewrite the equation as

$$\frac{2.7(\pm 0.283) \times 0.050(\pm 0.001)}{1850(\pm 11.2) \times 42.3(\pm 0.4)} = 1.725 \times 10^{-6}$$

The equation now contains only products and quotients, and Equation 6-12 applies. Thus,

$$\frac{s_y}{y} = \sqrt{\left(\pm\frac{0.283}{2.7}\right)^2 + \left(\pm\frac{0.001}{0.050}\right)^2 + \left(\pm\frac{11.2}{1850}\right)^2 + \left(\pm\frac{0.4}{42.3}\right)^2} = 0.107$$

To obtain the absolute standard deviation, we write

$$s_y = y \times 0.107 = 1.725 \times 10^{-6} \times (\pm 0.107) = \pm 0.185 \times 10^{-6}$$

and round the answer to $1.7(\pm 0.2) \times 10^{-6}$.

Standard Deviations in Exponential Calculations

- Consider the relationship: $y = a^x$

where the exponent x can be considered free of uncertainty.

-The relative standard deviation in y resulting from the uncertainty in a is

$$\frac{s_y}{y} = x\left(\frac{s_a}{a}\right)$$

- The relative standard deviation of the square of a number is twice the relative standard deviation of the number, the relative standard deviation of the cube root of a number is one third that of the number, and so forth.

- The relative standard deviation of $y = a^3$ is not the same as the relative standard deviation of the product of three independent measurements

$y = abc$, where $a = b = c$.

EXAMPLE 6-5

The solubility product K_{sp} for the silver salt AgX is 4.0 (±0.4) × 10^{-8}, and the molar solubility is

$$\text{solubility} = (K_{sp})^{1/2} = (4.0 \times 10^{-8})^{1/2} = 2.0 \times 10^{-4}\text{M}$$

What is the uncertainty in the calculated solubility of AgX?

Solution

Substituting y = solubility, $a = K_{sp}$, and $x = ½$ into Equation 6-13 gives

$$\frac{s_a}{a} = \frac{0.4 \times 10^{-8}}{4.0 \times 10^{-8}}$$

$$\frac{s_y}{y} = \frac{1}{2} \times \frac{0.4}{4.0} = 0.05$$

$$s_y = 2.0 \times 10^{-4} \times 0.05 = 0.1 \times 10^{-4}$$

$$\text{solubility} = 2.0\ (\pm 0.1) \times 10^{-4}\ \text{M}$$

Standard Deviations of Logarithms and Antilogarithms

For y = log a

$$s_y = 0.434\frac{s_a}{a}$$

And for y = antilog a

$$\frac{s_y}{y} = 2.303 s_a$$

The absolute standard deviation of the logarithm of a number is determined by the relative standard deviation of the number;

conversely, the relative standard deviation of the antilogarithm of a number is determined by the absolute standard deviation of the number.

EXAMPLE 6-6

Calculate the absolute standard deviations of the results of the following calculations. The absolute standard deviation for each quantity is given in parentheses.

(a) $y = \log[2.00(\pm 0.02) \times 10^{-4} = -3.6990 \pm ?$

(b) $y = \text{antilog}[1.200(\pm 0.003)] = 15.849 \pm ?$

(c) $y = \text{antilog}[45.4(\pm 0.3)] = 2.5119 \times 10^{45} \pm ?$

Solution

(a) Referring to Equation 6-14, we see that we must multiply the *relative* standard deviation by 0.434:

$$s_y = \pm 0.434 \times \frac{0.02 \times 10^{-4}}{2.00 \times 10^{-4}} = \pm 0.004$$

Thus,

$$y = \log[2.00(\pm 0.02) \times 10^{-4}] = -3.699\ (\pm 0.004)$$

(b) Applying Equation 6-15, we have

$$\frac{s_y}{y} = 2.303 \times (0.003) = 0.0069$$

$$s_y = 0.0069y = 0.0069 \times 15.849 = 0.11$$

Therefore,

$$y = \text{antilog}[1.200(\pm 0.003)] = 15.8 \pm 0.1$$

(c) $\dfrac{s_y}{y} = 2.303 \times (0.3) = 0.69$

$$s_y = 0.69y = 0.69 \times 2.5119 \times 10^{45} = 1.7 \times 10^{45}$$

Thus,

6D Reporting computed data

One of the best ways of indicating reliability is to give a confidence interval at the 90% or 95% confidence level.

Another method is to report the absolute standard deviation or the coefficient of variation of the data. A much less satisfactory but more common indicator of the quality of data is the significant figure convention.

Significant Figures

The significant figures in a number are all of the certain digits plus the first uncertain digit.

Rules for determining the number of significant figures:

1. Disregard all initial zeros.
2. Disregard all final zeros unless they follow a decimal point.
3. All remaining digits including zeros between nonzero digits are significant.

Significant Figures in Numerical Computations

Determining the appropriate number of significant figures in the result of an arithmetic combination of two or more numbers requires great care.

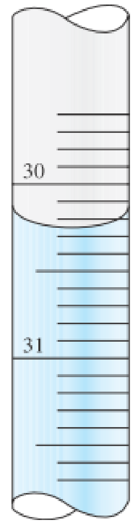


Figure 6-5 A buret section showing the liquid level and meniscus.

Sums and Differences

- For addition and subtraction, the result should have the same number of decimal places as the number with the smallest number of decimal places.

Note that the result contains three significant digits even though two of the numbers involved have only two significant figures.

Products and Quotients

For multiplication and division, the answer should be rounded so that it contains the same number of significant digits as the original number with the smallest number of significant digits.

This procedure sometimes leads to incorrect rounding.

When adding and subtracting numbers in scientific notation, express the numbers to the same power of ten.

The weak link for multiplication and division is the number of significant figures in the number with the smallest number of significant figures.

Use this rule of thumb with caution.

Logarithms and Antilogarithms

The following rules apply to the results of calculations involving logarithms:

1. In a logarithm of a number, keep as many digits to the right of the decimal point as there are significant figures in the original number.
2. In an antilogarithm of a number, keep as many digits as there are digits to the right of the decimal point in the original number.

The number of significant figures in the mantissa, or the digits to the right of the decimal point of a logarithm, is the same as the number of significant figures in the original number.

Thus, $\log (9.57 \times 10^4) = 4.981$.

Since 9.57 has 3 significant figures, there are 3 digits to the right of the decimal point in the result.

Rounding Data

In rounding a number ending in 5, always round so that the result ends with an even number. Thus, 0.635 rounds to 0.64 and 0.625 rounds to 0.62.

It is seldom justifiable to keep more than one significant figure in the standard deviation because the standard deviation contains error as well.

EXAMPLE 6-7

Round the following answers so that only significant digits are retained: (a) $\log 4.000 \times 10^{-5} = -4.3979400$, and (b) $\text{antilog } 12.5 = 3.162277 \times 10^{12}$

Solution

(a) Following rule 1, we retain 4 digits to the right of the decimal point

$$\log 4.000 \times 10^{-5} = -4.3979$$

(b) Following rule 2, we may retain only 1 digit

$$\text{antilog } 12.5 = 3 \times 10^{12}$$

Expressing Results of Chemical Calculations

- If the standard deviations of the values making up the final calculation are known, apply the propagation of error methods.

- However, if calculations have to be performed where the precision is indicated only by the significant figure convention, the result is rounded so that it contains only significant digits.

- It is especially important to postpone rounding until the calculation is completed.

- At least one extra digit beyond the significant digits should be carried through all of the computations in order to avoid a rounding error. This extra digit is sometimes called a “guard” digit.

EXAMPLE 6-8

A 3.4842-g sample of a solid mixture containing benzoic acid, C_6H_5COOH (122.123 g/mol), was dissolved and titrated with base to a phenolphthalein end point. The acid consumed 41.36 mL of 0.2328 M NaOH. Calculate the percent benzoic acid (HBz) in the sample.

Solution

As shown in Section 13C-3, the calculation takes the following form:

$$\%\text{HBz} = \frac{41.36\ \text{mL} \times 0.2328\ \dfrac{\text{mmol NaOH}}{\text{mL NaOH}} \times \dfrac{1\ \text{mmol HBz}}{\text{mmol NaOH}} \times \dfrac{122.123\ \text{g HBz}}{1000\ \text{mmol HBz}}}{3.842\ \text{g sample}}$$

$$\times\ 100\%$$

$$= 33.749\%$$

Since all operations are either multiplication or division, the relative uncertainty of the answer is determined by the relative uncertainties of the experimental data. Let us estimate what these uncertainties are.

1. The position of the liquid level in a buret can be estimated to ± 0.02 mL (Figure 6-5). In reading the buret, two readings (initial and final) must be made so that the standard deviation of the volume s_V will be

$$s_V = \sqrt{(0.02)^2 + (0.02)^2} = 0.028\ \text{mL}$$

The relative uncertainty in volume s_V/V is then

$$\frac{s_V}{V} = \frac{0.028}{41.36} \times 1000\ \text{ppt} = 0.68\ \text{ppt}$$

2. Generally, the absolute uncertainty of a mass obtained with an analytical balance will be on the order of ± 0.0001 g. Thus the relative uncertainty of the denominator s_D/D is

$$\frac{0.0001}{3.4842} \times 1000\ \text{ppt} = 0.029\ \text{ppt}$$

3. Usually we can assume that the absolute uncertainty in the concentration of a reagent solution is ±0.0001, and so the relative uncertainty in the concentration of NaOH, s_c/c is

$$\frac{s_c}{c} = \frac{0.0001}{0.2328} \times 1000\,\text{ppt} = 0.43\,\text{ppt}$$

4. The relative uncertainty in the molar mass of HBz is several orders of magnitude smaller than any of the three experimental values and will not be significant. Note, however, that we should retain enough digits in the calculation so that the molar mass is given to at least one more digit (the guard digit) than any of the experimental data. Thus, in the calculation, we use 122.123 for the molar mass (we are carrying two extra digits in this instance).
5. No uncertainty is associated with 100% and the 1000 mmol HBz since these are exact numbers.

Substituting the three relative uncertainties into Equation 6-12, we obtain

$$\frac{s_y}{y} = \sqrt{\left(\frac{0.028}{41.36}\right)^2 + \left(\frac{0.0001}{3.4842}\right)^2 + \left(\frac{0.0001}{0.2328}\right)^2}$$

$$= \sqrt{(0.00068)^2 + (0.000029)^2 + (0.00043)^2} = 8.02 \times 10^{-4}$$

$$s_y = 8.02 \times 10^{-4} \times y = 8.02 \times 10^{-4} \times 33.749 = 0.027$$

Therefore, the uncertainty in the calculated result is 0.03% HBz, and we should report the result as 33.75% HBz, or better 33.75 (± 0.03)% HBz.

Suggested Problems

- 6.1, 6.2, 6.5, 6.7, 6.8, 6.10(odd), 6.12, 6.15, 6.18