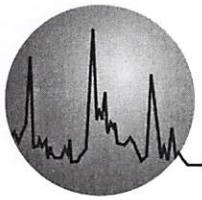
From: Fundamentals of Analytical chemistry, 7th ed.
Skoog, West, Holler
Saunders College Publishing: New York, 1996



CHAPTER

2

Errors in Chemical Analyses

It is impossible to perform a chemical analysis that is totally free of errors, or uncertainties. All one can hope is to minimize these errors and to estimate their size with acceptable accuracy. In this and the next two chapters, we explore the nature of experimental errors and their effects on the results of chemical analyses.

The consequence of errors in analytical data is illustrated in Figure 2-1, which shows results for the quantitative determination of iron(III). Six equal portions of an aqueous solution that contained exactly 20.00 ppm of iron(III) were analyzed in exactly the same way. Note that the results range from a low of 19.4 ppm to a high of 20.3 ppm of iron(III). The average $\bar{x}$ of the data is 19.8 ppm.

Parts per million (ppm), that is, 20.00 parts of iron(III) per million parts of solution.

Every measurement is influenced by many uncertainties, which combine to produce a scatter of results like that in Figure 2-1. Measurement uncertainties can never be completely eliminated, so the true value for any quantity is always unknown.[1] The probable magnitude of the error in a measurement can often be evaluated, however. We can then define limits within which the true value of a measured quantity lies at a given level of probability.

The true value of a measurement is never known exactly.

It is seldom easy to estimate the reliability of experimental data. Nevertheless, we must make such estimates whenever we collect laboratory results *because data of unknown quality are worthless.* On the other hand, results that are not especially accurate may be of considerable value if the limits of uncertainty are known.

Unfortunately, there is no simple and widely applicable method for determining the reliability of data with absolute certainty. The evaluation of experimental data often requires an effort that is comparable to that involved in their acquisition. Reliability can be assessed in several ways. Experiments designed to reveal the presence of errors can be performed. Standards of known composition can be analyzed and the results compared with the known composition. A few minutes in the library to consult the literature of analytical chemistry can be profitable. Calibrating equipment enhances the quality of data. Finally, statistical tests can be applied to the data. None of these options is perfect, so in the end we have to make *judgments* as to the probable accuracy of our results. These judgments tend to become harsher and less optimistic with experience.

[1]Unfortunately, many people do not understand these truths. For example, when asked by a defense attorney in a celebrated homicide investigation what the rate of error in a blood test was, the assistant district attorney replied that their testing laboratories had no percentage of error because "they have not committed any errors" (*San Francisco Chronicle,* June 29, 1994, p. 4).

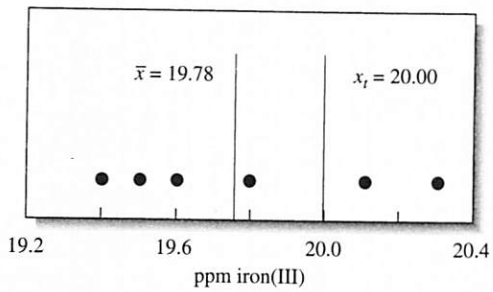


Figure 2-1
Results from six replicate determinations for iron in aqueous samples of a standard solution containing 20.00 ppm of iron(III).

One of the first questions to answer before beginning an analysis is, What is the maximum error that I can tolerate in the result? The answer to this question determines the time required to do the work. For example, a tenfold increase in accuracy may take hours, days, or even weeks of added labor. *No one can afford to waste time generating data that are more reliable than is needed.*

2A DEFINITION OF TERMS

Replicates are samples of about the same size that are carried through an analysis in *exactly* the same way.

Chemists usually carry two to five portions *(replicates)* of a sample through an entire analytical procedure. Individual results from a set of measurements are seldom the same (Figure 2-1), so a central or "best" value is used for the set. We justify the extra effort required to analyze several samples in two ways. First, the central value of a set should be more reliable than any of the individual results. Second, variation in the data should provide a measure of the uncertainty associated with the central result. Either the *mean* or the *median* may serve as the central value for a set of replicated measurements.

2A-1 The Mean and Median

The mean of two or more measurements is their average value.

Mean, arithmetic mean, and *average* ($\bar{x}$) are synonyms for the quantity obtained by dividing the sum of replicate measurements by the number of measurements in the set:

The symbol Σx_i means to add all of the values x_i for the replicates.

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

where x_i represents the individual values of x making up a set of N replicate measurements.

The *median* is the middle result when replicate data are arranged in order of size. There are equal numbers of data that are larger and smaller than the median. For an odd number of results, the median can be evaluated directly. For an even number, the mean of the middle pair is used.

The **median** is the middle value in a set of data that has been arranged in order of size. The median is used advantageously when a set of data contains an *outlier*—that is, a result that differs significantly from the rest of the data in the set. An outlier can have a significant effect on the mean of the set but has a lesser effect on the median.

EXAMPLE 2-1

Calculate the mean and median for the data shown in Figure 2-1.

$$\text{mean} = \bar{x} = \frac{19.4 + 19.5 + 19.6 + 19.8 + 20.1 + 20.3}{6}$$

$$= 19.78 \approx 19.8 \text{ ppm Fe}$$

Because the set contains an even number of measurements, the median is the average of the central pair:

$$\text{median} = \frac{19.6 + 19.8}{2} = 19.7 \text{ ppm Fe}$$

See *Mathcad Applications for Analytical Chemistry*, **pp. 15–18.**

Ideally, the mean and median are identical. Frequently they are not, however, particularly when the number of measurements in the set is small.

2A-2 Precision

Precision describes the reproducibility of measurements—that is, the closeness of results that have been obtained *in exactly the same way*. Generally, the precision of a measurement is readily determined by simply repeating the measurement.

Precision is the closeness of data to other data that have been obtained in exactly the same way.

Three terms are widely used to describe the precision of a set of replicate data: *standard deviation, variance,* and *coefficient of variation.* All of these terms are a function of the *deviation from the mean,* which is defined as

$$\text{deviation from the mean} = d_i = |x_i - \bar{x}| \qquad (2\text{-}2)$$

The relationship between deviation from the mean and the three precision terms is given in Section 3B.

Note that deviations from the mean are calculated without regard to sign.

2A-3 Accuracy

Figure 2-2 illustrates the difference between accuracy and precision. *Accuracy* indicates the closeness of the measurement to its true or accepted value and is expressed by the *error*. Note the basic difference between accuracy and precision. Accuracy measures agreement between a result and its true value. Precision describes the agreement among several results that have been measured in the same way. Precision is determined by simply replicating a measurement. On the other hand, accuracy can never be determined exactly because the true value of a quantity can never be known exactly. An accepted value must be used instead.

Accuracy is the closeness of a result to its true or accepted value.

Accuracy is expressed in terms of either absolute or relative error.

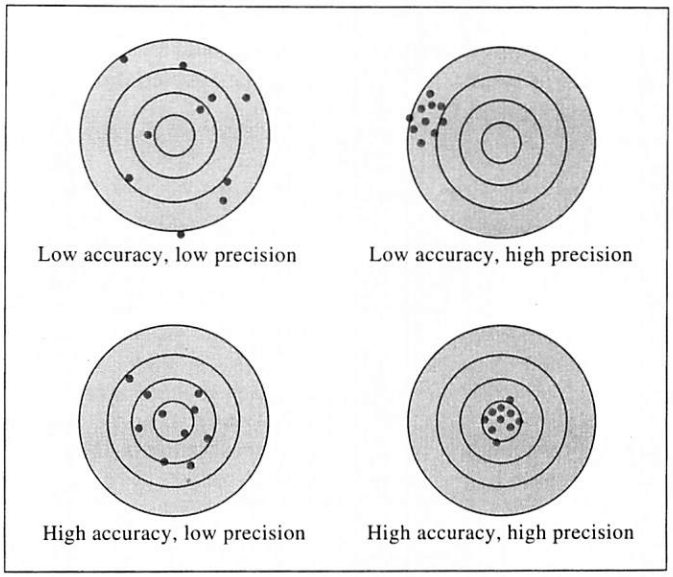


Figure 2-2
Accuracy and precision.

Absolute Error

The term "absolute" has a different meaning here than it does in mathematics. An absolute value in mathematics means the magnitude of a number *ignoring its sign.* As we shall use it, the absolute error is the difference between an *experimental result and the accepted value including its sign.*

The *absolute error E* in the measurement of a quantity x_i is given by the equation

$$E = x_i - x_t \tag{2-3}$$

The absolute error of a measurement is the difference between the measured value and the true value. It bears a sign.

where x_t is the true, or accepted, value of the quantity. Returning to the data displayed in Figure 2-1, the absolute error of the result immediately by the left of the true value of 20.00 ppm is -0.2 ppm Fe; the result at 20.10 ppm is in error by $+0.1$ ppm Fe. Note that we retain the sign in stating the error. Thus, the negative sign in the first case shows that the experimental result is smaller than the accepted value.

Relative Error

The relative error of a measurement is the absolute error divided by the true value.

Often, the *relative error* E_r is a more useful quantity than the absolute error. The percent relative error is given by the expression

$$E_r = \frac{x_i - x_t}{x_t} \times 100\% \tag{2-4}$$

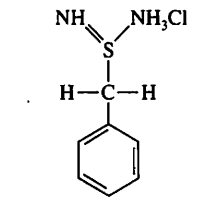
Relative error is also expressed in parts per thousand (ppt). Thus, the relative error for the mean of the data in Figure 2-1 is

$$E_r = \frac{19.8 - 20.00}{20.00} \times 100\% = -1\% \text{ or } -10 \text{ ppt}$$

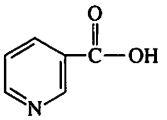
2A-4 Types of Errors in Experimental Data

The precision of a measurement is readily determined by comparing data from carefully replicated experiments. Unfortunately, an estimate of the accuracy is not so easy to obtain. To determine the accuracy, we have to know the true value, and ordinarily this is exactly what we are looking for.

It is tempting to assume that if we know the answer precisely, then we also know it accurately. The danger of this assumption is illustrated in Figure 2-3, which summarizes the results for the determination of nitrogen in the two pure compounds shown in the margin. The dots show the absolute errors of replicate results obtained by four analysts. Note that analyst 1 obtained relatively high precision and high accuracy. Analyst 2 had poor precision but good accuracy. The results of analyst 3 are surprisingly common. The precision is excellent, but there is significant error in the numerical average for the data. Both the precision and the accuracy are poor for the results of analyst 4.



benzyl isothiourea hydrochloride



nicotinic acid

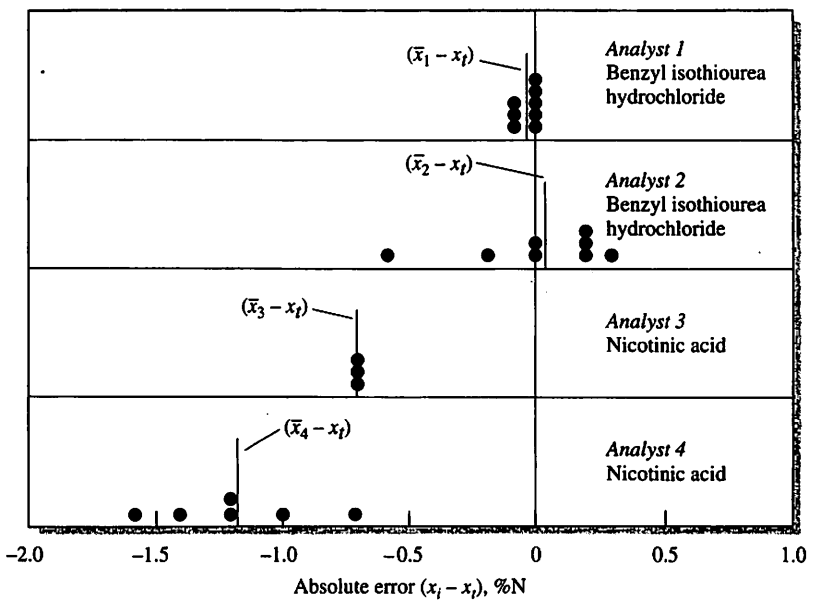


Figure 2-3

Absolute error in the micro-Kjeldahl determination of nitrogen. Each dot represents the error associated with a single determination. Each vertical line labeled $(\bar{x}_i - x_t)$ is the absolute average deviation of the set from the true value. (Data from C. O. Willits and C. L. Ogg, *J. Assoc. Offic. Anal. Chem.*, **1949**, *32*, 561. With permission.)

Random, or indeterminate, errors are errors that affect the precision of measurement.

Figures 2-1 and 2-3 suggest that chemical analyses are affected by at least two types of errors. One type, called *random* (or *indeterminate*) *error,* causes data to be scattered more or less symmetrically around a mean value. Refer again to Figure 2-3, and notice that the scatter in the data, and thus the random error, for analysts 1 and 3 is significantly less than that for analysts 2 and 4. In general, then, the random error in a measurement is reflected by its precision.

Systematic, or determinate, errors affect the accuracy of results.

A second type of error, called *systematic* (or *determinate*) *error,* causes the mean of a set of data to differ from the accepted value. For example, the results of analysts 1 and 2 in Figure 2-3 have little systematic error, but the data of analysts 3 and 4 show determinate errors of about -0.7 and -1.2% nitrogen. In general, a systematic error causes the results in a series of replicate measurements to be all high or all low.

A third type of error is *gross error.* Gross errors differ from indeterminate and determinate errors. They usually occur only occasionally, are often large, and may cause a result to be either high or low. Gross errors lead to *outliers*—results that appear to differ markedly from all other data in a set of replicate measurements. There is no evidence of a gross error in Figures 2-1 and 2-3. Had one of the results shown in Figure 2-1 occurred at 22.5 ppm Fe, it might have been an outlier.

An outlier is an occasional result in replicate measurements that obviously differs significantly from the rest of the results.

2B SYSTEMATIC ERRORS

Systematic errors have a definite value, an assignable cause, and are of the same magnitude for replicate measurements made in the same way. Systematic errors lead to *bias* in measurement technique. Note that bias affects all of the data in a set in the same way and that it bears a sign.

Bias measures the systematic error associated with an analysis. It has a negative sign if it causes the results to be low and a positive sign otherwise.

2B-1 Sources of Systematic Errors

There are three types of systematic errors: (1) *Instrument errors* are caused by imperfections in measuring devices and instabilities in their power supplies. (2) *Method errors* arise from nonideal chemical or physical behavior of analytical systems. (3) *Personal errors* result from the carelessness, inattention, or personal limitations of the experimenter.

Instrument Errors

All measuring devices are sources of systematic errors. For example, pipets, burets, and volumetric flasks may hold or deliver volumes slightly different from those indicated by their graduations. These differences may arise from using glassware at a temperature that differs significantly from the calibration temperature, from distortions in container walls due to heating while drying, from errors in the original calibration, or from contaminants on the inner surfaces of the containers. Calibration eliminates most systematic errors of this type.

Electronic instruments are subject to instrumental systematic errors. These uncertainties have many sources. For example, errors emerge as the voltage of a battery-operated power supply decreases with use. Errors result from increased resistance in circuits because of dirty electrical contacts. Temperature changes

cause variations in resistors and standard potential sources. Currents induced from 110-V power lines affect electronic instruments. Errors from these and other sources are detectable and correctable.

Method Errors

The nonideal chemical or physical behavior of the reagents and reactions upon which an analysis is based often introduces systematic method errors. Such sources of nonideality include the slowness of some reactions, the incompleteness of others, the instability of some species, the nonspecificity of most reagents, and the possible occurrence of side reactions that interfere with the measurement process. For example, a common method error in volumetric methods results from the small excess of reagent required to cause an indicator to undergo the color change that signals completion of the reaction. The accuracy of such an analysis is thus limited by the very phenomenon that makes the titration possible.

Another example of method error is illustrated by the data in Figure 2-3, in which the results by analysts 3 and 4 show a negative bias that can be traced to the chemical nature of the sample, nicotinic acid. The analytical method used involves the decomposition of the organic samples in hot concentrated sulfuric acid, which converts the nitrogen in the samples to ammonium sulfate. The amount of ammonia in the ammonium sulfate is then determined in the measurement step. Experiments have shown that compounds containing a pyridine ring such as nicotinic acid (see page 15) are incompletely decomposed by the sulfuric acid unless special precautions are taken. Without these precautions, low results are obtained. It is highly likely that the negative errors, $(\bar{x}_3 - x_t)$ and $(\bar{x}_4 - x_t)$ in Figure 2-3 are systematic errors that can be blamed on incomplete decomposition of the samples.

Errors inherent in a method are often difficult to detect and are thus the most serious of the three types of systematic error.

Of the three types of systematic errors encountered in a chemical analysis, method errors are usually the most difficult to identify and correct.

Personal Errors

Many measurements require personal judgments. Examples include estimating the position of a pointer between two scale divisions, the color of a solution at the end point in a titration, or the level of a liquid with respect to a graduation in a pipet or buret (see Figure 3-5, page 41). Judgments of this type are often subject to systematic, unidirectional errors. For example, one person may read a pointer consistently high, another may be slightly slow in activating a timer, and a third may be less sensitive to color changes. An analyst who is insensitive to color changes tends to use excess reagent in a volumetric analysis. Physical handicaps are often sources of personal determinate errors.

Color blindness is a good example of a handicap that amplifies personal errors in a volumetric analysis. A famous color-blind analytical chemist enlisted his wife to come to the laboratory to help him detect color changes at end points of titrations.

A universal source of personal error is prejudice. Most of us, no matter how honest, have a natural tendency to estimate scale readings in a direction that improves the precision in a set of results, or we may have a preconceived notion of the true value for the measurement. We then subconsciously cause the results to fall close to this value. Number bias is another source of personal error that varies considerably from person to person. The most common number bias encountered in estimating the position of a needle on a scale involves a preference for the digits 0 and 5. Also prevalent is a prejudice favoring small digits over large and even numbers over odd.

Digital readouts on pH meters, laboratory balances, and other electronic instruments eliminate number bias because no judgment is involved in taking a reading.

Persons who make measurements must guard against personal bias to preserve the integrity of the collected data.

Constant errors are independent of the size of the sample being analyzed. **Proportional errors** decrease or increase in proportion to the size of the sample.

2B-2 The Effect of Systematic Errors upon Analytical Results

Systematic errors may be either *constant* or *proportional.* The magnitude of a constant error does not depend on the size of the quantity measured. Proportional errors increase or decrease in proportion to the size of the sample taken for analysis.

Constant Errors

Constant errors become more serious as the size of the quantity measured decreases. The effect of solubility losses on the results of a gravimetric analysis illustrates this behavior.

EXAMPLE 2-2

Suppose that 0.50 mg of precipitate is lost as a result of being washed with 200 mL of wash liquid. If the precipitate weighs 500 mg, the relative error due to solubility loss is $-(0.50/500) \times 100\% = -0.1\%$. Loss of the same quantity from 50 mg of precipitate results in a relative error of -1.0%.

The excess of reagent required to bring about a color change during a titration is another example of constant error. This volume, usually small, remains the same regardless of the total volume of reagent required for the titration. Again, the relative error from this source becomes more serious as the total volume decreases. One way of minimizing the effect of constant error is to use as large a sample as possible.

Proportional Errors

A common cause of proportional errors is the presence of interfering contaminants in the sample. For example, a widely used method for the determination of copper is based upon the reaction of copper(II) ion with potassium iodide to give iodine. The quantity of iodine is then measured and is proportional to the amount of copper in the sample. Iron(III), if present, also liberates iodine from potassium iodide. Unless steps are taken to prevent this interference, high results are observed for the percentage of copper because the iodine produced will be a measure of the copper(II) *and* iron(III) in the sample. The size of this error is fixed by the *fraction* of iron contamination, which is independent of the size of sample taken. If the sample size is doubled, for example, the amount of iodine liberated by both the copper and the iron contaminant is also doubled. Thus, the magnitude of the reported percentage of copper is independent of sample size.

2B-3 Detection of Systematic Instrument and Personal Errors

After entering a reading into the laboratory notebook, many scientists habitually make a second reading to assure the correctness of the entry.

Systematic instrument errors are usually found and corrected by calibration. Periodic calibration of equipment is always desirable because the response of most instruments changes with time as a result of wear, corrosion, or mistreatment.

Most personal errors can be minimized by care and self-discipline. It is a good habit to check instrument readings, notebook entries, and calculations systemati-

cally. Errors that result from a known physical handicap can usually be avoided by a careful choice of method.

2B-4 Detection of Systematic Method Errors

Bias in an analytical method is particularly difficult to detect. We may take one or more of the following steps to recognize and adjust for a systematic error in an analytical method.

Analysis of Standard Samples

The best way of estimating the bias of an analytical method is by the analysis of *standard reference materials*—materials that contain one or more analytes with exactly known concentration levels. Standard reference materials are obtained in several ways.

Standard materials can sometimes be prepared by synthesis. Here, carefully measured quantities of the pure components of a material are measured out and mixed so as to produce a homogeneous sample whose composition is known from the quantities taken. The overall composition of a synthetic standard material must approximate closely the composition of the samples to be analyzed. Great care must be taken to ensure that the concentration of analyte is known exactly. Unfortunately, the synthesis of such standard samples is often impossible or so difficult and time-consuming that this approach is not practical.

Standard reference materials can be purchased from a number of governmental and industrial sources. For example, the National Institute of Standards and Technology (NIST) (formerly the National Bureau of Standards) offers over 900 standard reference materials including rocks and minerals, gas mixtures, glasses, hydrocarbon mixtures, polymers, urban dusts, rainwaters, and river sediments.[2] The concentration of one or more of the components in these materials has been determined in one of three ways: (1) through analysis by a previously validated reference method, (2) through analysis by two or more independent, reliable measurement methods, or (3) through analysis by a network of cooperating laboratories, technically competent and thoroughly knowledgeable with the material being tested.

Several commercial supply houses also offer analyzed materials for method testing.[3]

One of the problems you will encounter in using standard reference materials to establish the presence or absence of bias is that the mean of your replicate analysis of the standard will ordinarily differ somewhat from the theoretical result. Then you are faced with the question whether this difference is due to random error of your measurements or to bias in the method. In Section 4B-1, we demonstrate a statistical test that can be applied to aid your judgment in answering this question.

Figure 2-4
Standard reference materials from NIST. (Photo courtesy of the National Institute of Standards and Technology.)

A standard reference material (SRM) is a substance prepared and sold by the National Institute of Standards and Technology and certified to contain specified concentrations of one or more analytes.

In using SRMs it is often difficult to separate bias from ordinary random error.

[2]See U.S. Department of Commerce, *NIST Standard Reference Materials Catalog,* 1992–1993 ed., NIST Special Publication 260. Washington: Government Printing Office, 1992. For a description of the reference material programs of the NIST, see R. A. Alvarez, S. D. Rasberry, and G. A. Uriano, *Anal. Chem.,* **1982,** *54,* 1226A; G. A. Uriano, *ASTM Standardization News,* **1979**, *7,* 8.

[3]For sources of biological and environmental reference materials containing various elements, see C. Veillon, *Anal. Chem.,* **1986**, *58,* 851A.

Independent Analysis

If standard samples are not available, a second independent and reliable analytical method can be used in parallel with the method being evaluated. The independent method should differ as much as possible from the one under study. This minimizes the possibility that some common factor in the sample has the same effect on both methods. Here again, a statistical test must be used to determine whether any difference is a result of random errors in the two methods or to bias in the method under study (see Section 4B-2).

Blank Determinations

A **blank** solution contains the solvent and all of the reagents in an analysis, but none of the sample.

Blank determinations are useful for detecting certain types of constant errors. In a blank determination, or *blank*, all steps of the analysis are performed in the absence of a sample. The results from the blank are then applied as a correction to the sample measurements. Blank determinations reveal errors due to interfering contaminants from the reagents and vessels employed in analysis. Blanks also allow the analyst to correct titration data for the volume of reagent needed to cause an indicator to change color at the end point.

Variation in Sample Size

Example 2-2 demonstrates that as the size of a measurement increases, the effect of a constant error decreases. Thus, constant errors can often be detected by varying the sample size.

2C QUESTIONS AND PROBLEMS

2-1. Explain the difference between
- *(a) constant and proportional error.
- (b) random and systematic error.
- *(c) mean and median.
- (d) absolute and relative error.

*2-2. Suggest some sources of random error in measuring the width of a 3-m table with a 1-m metal rule.

*2-3. Name three types of systematic errors.

2-4. How are systematic method errors detected?

*2-5. What kind of systematic errors are detected by varying the sample size?

2-6. A method of analysis yields weights for gold that are low by 0.3 mg. Calculate the percent relative error caused by this uncertainty if the weight of gold in the sample is
- *(a) 800 mg.
- (b) 500 mg.
- *(c) 100 mg.
- (d) 25 mg.

2-7. The method described in Problem 2-6 is to be used for the analysis of ores that assay about 1.2% gold. What minimum sample weight should be taken if the relative error resulting from a 0.3-mg loss is not to exceed
- *(a) −0.2%?
- (b) −0.5%?
- *(c) −0.8%?
- (d) −1.2%?

2-8. The color change of a chemical indicator requires an overtitration of 0.03 mL. Calculate the percent relative error if the total volume of titrant is
- *(a) 50.00 mL.
- (b) 10.0 mL.
- *(c) 25.0 mL.
- (d) 40.0 mL.

2-9. A loss of 0.4 mg of Zn occurs in the course of an analysis for that element. Calculate the percent relative error due to this loss if the weight of Zn in the sample is
- *(a) 40 mg.
- (b) 175 mg.
- *(c) 400 mg.
- (d) 600 mg.