McGraw-Hill Higher Education
A Division of The McGraw-Hill Companies

DATA REDUCTION AND ERROR ANALYSIS FOR THE PHYSICAL SCIENCES
THIRD EDITION

Published by McGraw-Hill, a business unit of The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY 10020.

This book is printed on acid-free paper.

12 13 DOC/DOC 10 9 8 7 6 5 4 3

ISBN-13: 978-0-07-247227-1
ISBN-10: 0-07-247227-8

Publisher: *Kent A. Peterson*
Sponsoring editor: *Daryl Bruflodt*
Developmental editor: *Spencer J. Cotkin, Ph.D.*
Marketing manager: *Debra B. Hash*
Senior project manager: *Mary E. Powers*
Senior production supervisor: *Laura Fuller*
Senior media project manager: *Stacy A. Patch*
Lead media technology producer: *Judi David*
Coordinator of freelance design: *Rick D. Noel*
Cover designer: *John Rokusek/Rokusek Design*
Cover diagrams provided by: *D. Keith Robinson*
Compositor: *GAC—Indianapolis*
Typeface: *10/12 Times Roman*
Printer: *R. R. Donnelley & Sons Company/Crawfordsville, IN*

Library of Congress Cataloging-in-Publication Data

Bevington, Philip R., 1933-1980.
Data reduction and error analysis for the physical sciences / Philip R. Bevington, D. Keith Robinson.—3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-07-247227-8
1. Multivariate analysis. 2. Error analysis (Mathematics). 3. Least squares. 4. Data reduction. I. Robinson, D. Keith. II. Title.

QA278 .B48 2003
511'.43—dc21 2002070896
CIP

www.mhhe.com

CONTENTS

PREFACE TO THE THIRD EDITION

In his 1969 Preface to the first edition of this book, the late Philip Bevington aptly stated his purpose, " to provide an introduction to the techniques of data reduction and error analysis commonly employed by individuals doing research in the physical sciences and to present them in sufficient detail and breadth to make them useful for students throughout their undergraduate and graduate studies. The presentation is developed from a practical point of view, including enough derivation to justify results, but emphasizing the methods more than the theory." This third edition continues Phil's original mission, updated to reflect the ready availability of modern computers.

The first four chapters introduce the concepts of measuring uncertainties, error analysis, and probability distributions, with a new section on probabilities in low-statistics experiments. Chapter 5 provides an introduction to Monte Carlo methods for simulating experimental data, methods that are applied in later chapters to generate data for examples and to study and evaluate the statistical significance of experimental results. In chapters 6 through 9, the least-squares method is applied to problems of increasing complexity, from analytic straight-line fits to nonlinear fits that require iterative solutions. Chapter 10 provides an introduction to the direct application of the maximum-likelihood method, and chapter 11 includes a discussion of χ^2-probability, confidence intervals, and correlation coefficients. Exercises at the end of the chapters range in complexity from simple statistical calculations to minor projects such as least-squares fitting and Monte Carlo calculations. Answers to selected exercises are provided.

The appendixes from previous editions have been retained. Appendix A includes a brief section on basic differential calculus but is devoted mainly to numerical methods that are useful in analyzing data on the computer. Determinants and matrices are discussed in appendix B. Appendix C provides tables and graphs of statistical functions, augmented by computer routines on the website for calculating probabilities. Appendix D sets forth some guidelines for the preparation of effective graphs. Appendix E provides listings of computer routines that illustrate the text.

COMPUTER ROUTINES

Simple, illustrative computer routines that were a useful feature of the original book have been retained and are listed in Fortran77 in appendix E. Fortran was chosen because it has proved to be the most durable of languages over many decades. (Pascal, which was provided in the second edition, has vanished, displaced by C++.) With the help of the comments at the beginning of appendix E, students should be able to read the Fortran programs and follow their logic without special expertise in the language. To simplify the listed routines and to clarify their main objectives, we have deleted most of the calls to graphics routines.

Computer routines and programs are available for downloading in both Fortran and C++ from the www.mhhe.com/bevington website, along with supporting routines to facilitate the construction of complete programs for Monte Carlo generation, least-squares fitting, and probability calculations. A "Read Me" file on the site describes the organization of the programs and provides instructions for using them.

ACKNOWLEDGMENTS

I am most indebted to the late Philip R. Bevington for his original book, which formed the basis for these revisions. I am grateful to the Case Western Reserve Physics Department for its support, and to my undergraduate laboratory students for providing several of the examples and much inspiration. I thank Spencer Cotkin and his colleagues at McGraw-Hill for their encouragement.

I would also like to thank readers and, in particular, the following reviewers, for their many helpful comments and suggestions: Jingsong Zhang, University of California, Riverside; Gary Schmidt, University of Arizona; Herbert Strauss, University of California, Berkeley; Daniel Suson, Texas A&M, Kingsville.

Finally, I wish to thank my wife Margi for her remarkable patience and support.

D. Keith Robinson

ABOUT THE AUTHORS

The late Philip R. Bevington was a professor of physics at Case Western Reserve University. He graduated from Harvard University in 1954 and received his Ph.D. from Duke University in 1960. He taught at Duke University for five years and was an assistant professor at Stanford University from 1963 to 1968 before coming to Case Western Reserve University. He was involved in research in nuclear structure physics with Van de Graaff accelerators. While at Stanford he was active in computer applications for nuclear physics and was responsible for development of the SCANS system.

D. Keith Robinson is an emeritus professor of physics at Case Western Reserve University in Cleveland, Ohio. He received his B.Sc. in physics from Dalhousie University in Canada in 1954 and his D.Phil. from the University of Oxford in 1960. He was a member of the staff at Brookhaven National Laboratory from 1960 until 1966 when he joined CWRU. His research in experimental particle physics has included studies of boson resonances, K-meson properties, antiproton-proton interactions, and the radiative decay of hyperons. He has been strongly involved in developing computer-based laboratories for the introductory physics courses at CWRU.

CHAPTER 1

UNCERTAINTIES IN MEASUREMENTS

1.1 MEASURING ERRORS

It is a well-established fact of scientific investigation that the first time an experiment is performed the results often bear all too little resemblance to the "truth" being sought. As the experiment is repeated, with successive refinements of technique and method, the results gradually and asymptotically approach what we may accept with some confidence to be a reliable description of events. We may sometimes feel that nature is loath to give up her secrets without a considerable expenditure of effort on our part, and that first steps in experimentation are bound to fail. Whatever the reason, it is certainly true that for all physical experiments, errors and uncertainties exist that must be reduced by improved experimental techniques and repeated measurements, and those errors remaining must always be estimated to establish the validity of our results.

Error is defined by Webster as "the difference between an observed or calculated value and the true value." Usually we do not know the "true" value; otherwise there would be no reason for performing the experiment. We may know approximately what it should be, however, either from earlier experiments or from theoretical predictions. Such approximations can serve as a guide but we must always determine in a systematic way from the data and the experimental conditions themselves how much confidence we can have in our experimental results.

There is one class of error that we can deal with immediately: errors that originate from mistakes or blunders in measurement or computation. Fortunately, these errors are usually apparent either as obviously incorrect data points or as results that are not reasonably close to expected values. They are classified as *illegitimate errors* and generally can be corrected by carefully repeating the operations. Our interest is

in *uncertainties* introduced by random fluctuations in our measurements, and *systematic errors* that limit the precision and accuracy of our results in more or less well-defined ways. Generally, we refer to the uncertainties as the *errors* in our results, and the procedure for estimating them as *error analysis.*

Accuracy Versus Precision

It is important to distinguish between the terms *accuracy* and *precision.* The accuracy of an experiment is a measure of how close the result of the experiment is to the true value; the precision is a measure of how well the result has been determined, without reference to its agreement with the true value. The precision is also a measure of the reproducibility of the result in a given experiment. The distinction between accuracy and precision is illustrated by the two sets of measurements in Figure 1.1 where the straight line on each graph shows the expected relation between the dependent variable y and the independent variable x. In both graphs, the scatter of the data points is a reflection of uncertainties in the measurements, consistent with the error bars on the points. The data in Figure 1.1(a) have been measured to a high degree of precision as illustrated by the small error bars, and are in excellent agreement with the expected variation of y with x, but are clearly inaccurate, deviating from the line by a constant offset. On the other hand, the data points in Figure 1.1(b) are rather imprecise as illustrated by the large error bars, but are scattered about the predicted distribution.

It is obvious that we must consider the accuracy and precision simultaneously for any experiment. It would be a waste of time and energy to determine a result with high precision if we knew that the result would be highly inaccurate. Conversely, a

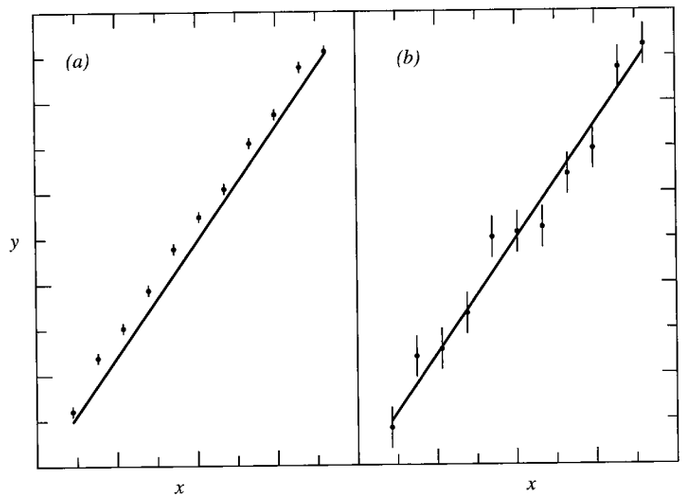


FIGURE 1.1
Illustration of the difference between precision and accuracy. (*a*) Precise but inaccurate data. (*b*) Accurate but imprecise data. True values are represented by the straight lines.

result cannot be considered to be extremely accurate if the precision is low. In general, when we quote the *uncertainty* or *error* in an experimental result, we are referring to the precision with which that result has been determined. *Absolute* precision indicates the magnitude of the uncertainty in the result in the same units as the result; *relative* precision indicates the uncertainty in terms of a fraction of the value of the result.

Systematic Errors

The accuracy of an experiment, as we have defined it, is generally dependent on how well we can control or compensate for systematic errors, errors that will make our results different from the "true" values with reproducible discrepancies. Errors of this type are not easy to detect and not easily studied by statistical analysis. They may result from faulty calibration of equipment or from bias on the part of the observer. They must be estimated from an analysis of the experimental conditions and techniques. A major part of the planning of an experiment should be devoted to understanding and reducing sources of systematic errors.

EXAMPLE 1.1 A student measures a table top with a steel meter stick and finds that the average of his measurements yields a result of (1.982 ± 0.001)m for the length of the table. He subsequently learns that the meter stick was calibrated at 25 °C and has an expansion coefficient of 0.0005 $°C^{-1}$. Because his measurements were made at a room temperature of 20°C, they are systematically too small. To correct for this effect, he multiplies his results by $1 + 0.0005 \times (20 - 25) = 0.9975$ so that his new determination of the length is 1.977 m.

When the student repeats the experiment, he discovers a second systematic error, his technique for reading the meter stick was faulty in that he did not always read the divisions from directly above. By experimentation he determines that this consistently resulted in a reading that was 2 mm short. The corrected result is 1.979 m.

In this example, the first result was given with a fairly high precision, approximately 1 part in 2000. The corrections to this result were meant to improve the accuracy by compensating for known sources of deviation of the first result from the best estimate possible. These corrections did not improve the precision at all, but did in fact worsen it, because the corrections were themselves only estimates of the exact corrections. Before quoting his final result, the student must reexamine his error analysis and take account of any additional uncertainties that may have been introduced by these corrections.

Random Errors

The precision of an experiment depends upon how well we can overcome random errors, fluctuations in observations that yield different results each time the experiment is repeated, and thus require repeated experimentation to yield precise results. A given accuracy implies an equivalent precision and, therefore, also depends to some extent on random errors.

The problem of reducing random errors is essentially one of improving the experimental method and refining the techniques, as well as simply repeating the

experiment. If the random errors result from instrumental uncertainties, they may be reduced by using more reliable and more precise measuring instruments. If the random errors result from statistical fluctuations in a limited number of measurements, they may be reduced by making more measurements. There are practical limits to these improvements. In the measurement of the length of the table of Example 1.1, the student might attempt to improve the precision of his measurements by using a magnifying glass to read the scale, or he might attempt to reduce statistical fluctuations in his measurements by repeating the measurement several times. In neither case would it be useful to reduce the random errors much below the systematic errors, such as those introduced by the calibration of the meter stick or the correction for his initial faulty reading of the scale. The limits imposed by systematic errors are important considerations in planning and performing experiments.

Significant Figures and Roundoff

The precision of an experimental result is implied by the number of digits recorded in the result, although generally the uncertainty should be quoted specifically as well. The number of *significant figures* in a result is defined as follows:

1. The leftmost nonzero digit is the most significant digit.
2. If there is no decimal point, the rightmost nonzero digit is the least significant digit.
3. If there is a decimal point, the rightmost digit is the least significant digit, even if it is a 0.
4. All digits between the least and most significant digits are counted as significant digits.

For example, the following numbers each have four significant digits: 1234, 123,400, 123.4, 1001, 1000., 10.10, 0.0001010, 100.0. If there is no decimal point, there are ambiguities when the rightmost digit is 0. Thus, the number 1010 is considered to have only three significant digits even though the last digit might be physically significant. To avoid ambiguity, it is better to supply decimal points or to write such numbers in *scientific notation*, that is, as an argument in decimal notation multiplied by the appropriate power of 10. Thus, our example of 1010 would be written as 1010. or 1.010×10^3 if all four digits are significant.

When quoting an experimental result, the number of significant figures should be approximately one more than that dictated by the experimental precision. The reason for including the extra digit is to avoid errors that might be caused by rounding errors in later calculations. If the result of the measurement of Example 1.1 is $L = 1.979$ m with an uncertainty of 0.012 m, this result could be quoted as $L = (1.979 \pm 0.012)$ m. However, if the first digit of the uncertainty is large, such as 0.082 m, then we should probably quote $L = (1.98 \pm 0.08)$ m. In other words, we let the uncertainty define the precision to which we quote our result.

When insignificant digits are dropped from a number, the last digit retained should be rounded off for the best accuracy. To round off a number to fewer significant digits than were specified originally, we truncate the number as desired and treat the excess digits as a decimal fraction. Then:

1. If the fraction is greater than ½, increment the new least significant digit.
2. If the fraction is less than ½, do not increment.
3. If the fraction equals ½, increment the least significant digit only if it is odd.

The reason for rule 3 is that a fractional value of ½ may result from a previous rounding up of a fraction that was slightly less than ½ or a rounding down of a fraction that was slightly greater than ½. For example, 1.249 and 1.251 both round to three significant figures as 1.25. If we were to round again to two significant figures, both would yield the same value, either 1.2 or 1.3, depending on our convention. Choosing to round up if the resulting last digit is odd and to round down if the resulting last digit is even, reduces systematic errors that would otherwise be introduced into the average of a group of such numbers. Note that it is generally advisable to retain all available digits in intermediate calculations and round only the final results.

1.2 UNCERTAINTIES

Uncertainties in experimental results can be separated into two categories: those that result from fluctuations in measurements, and those associated with the theoretical description of our result. For example, if we measure the length of a rectangular table along one edge, we know that any uncertainties, aside from systematic errors, are associated with the fluctuations of our measurements from trial to trial. With an infinite number of measurements we might be able to estimate the length very precisely, but with a finite number of trials there will be a finite uncertainty. If we were to measure the length of the table at equally spaced positions across the table, the measurements would show additional fluctuations corresponding to irregularities in the table itself, and our result could be expressed as the mean length. If, however, we were to describe the shape of an oval table, we would be faced with uncertainties both in the measurement of position of the edge of the table at various points and in the form of the equation to be used to describe the shape, whether it be circular, elliptical, or whatever. Thus, we shall be concerned in the following chapters with a comparison of the distribution of measured data points with the distribution predicted on the basis of a theoretical model. This comparison will help to indicate whether our method of extracting the results is valid or needs modification.

The term *error* suggests a deviation of the result from some "true" value. Usually we cannot know what the true value is, and can only estimate the errors inherent in the experiment. If we repeat an experiment, the results may well differ from those of the first attempt. We express this difference as a *discrepancy* between the two results. Discrepancies arise because we can determine a result only with a given *uncertainty.* For example, when we compare different measurements of a standard physical constant, or compare our result with the accepted value, we should refer to the differences as *discrepancies*, not errors or uncertainties.

Because, in general, we shall not be able to quote the actual error in a result, we must develop a consistent method for determining and quoting the estimated

error. A study of the distribution of the results of repeated measurements of the same quantity can lead to an understanding of these errors so that the quoted error is a measure of the spread of the distribution. However, for some experiments it may not be feasible to repeat the measurements and experimenters must therefore attempt to estimate the errors based on an understanding of the apparatus and their own skill in using it. For example, if the student of Example 1.1 could make only a single measurement of the length of the table, he should examine his meter stick and the table, and try to estimate how well he could determine the length. His estimate should be consistent with the result expected from a study of repeated measurements; that is, to quote an estimate for the *standard error*, he should try to estimate a range into which he would expect repeated measurements to fall about seven times out of ten. Thus, he might conclude that with a fine steel meter stick and a well-defined table edge, he could measure to about ± 1 mm or ± 0.001 m. He should resist the temptation to increase this error estimate, "just to be sure."

We must also realize that the model from which we calculate theoretical parameters to describe the results of our experiment may not be the correct model. In the following chapters we shall discuss hypothetical parameters and probable distributions of errors pertaining to the "true" states of affairs, and we shall discuss methods of making experimental estimates of these parameters and the uncertainties associated with these determinations.

Minimizing Uncertainties and Best Results

Our preoccupation with error analysis is not confined just to the determination of the precision of our results. In general, we shall be interested in obtaining the maximum amount of useful information from the data on hand without being able either to repeat the experiment with better equipment or to reduce the statistical uncertainties by making more measurements. We shall be concerned, therefore, with the problem of extracting from the data the best estimates of theoretical parameters and of the random errors, and we shall want to understand the effect of these errors on our results, so that we can determine what confidence we can place in our final results. It is reasonable to expect that the most reliable results we can calculate from a given set of data will be those for which the estimated errors are the smallest. Thus, our development of techniques of error analysis will help to determine the optimum estimates of parameters to describe the data.

It must be noted, however, that even our best efforts will yield only *estimates* of the quantities investigated.

1.3 PARENT AND SAMPLE DISTRIBUTIONS

If we make a measurement x_1 of a quantity x, we expect our observation to approximate the quantity, but we do not expect the experimental data point to be exactly equal to the quantity. If we make another measurement, we expect to observe a discrepancy between the two measurements because of random errors, and we do not expect either determination to be exactly correct, that is, equal to x. As we make more and more measurements, a pattern will emerge from the data. Some of the measurements will be too large, some will be too small. On the average, however,

we expect them to be distributed around the correct value, assuming we can neglect or correct for systematic errors.

If we could make an infinite number of measurements, then we could describe exactly the distribution of the data points. This is not possible in practice, but we can hypothesize the existence of such a distribution that determines the probability of getting any particular observation in a single measurement. This distribution is called the *parent distribution.* Similarly, we can hypothesize that the measurements we have made are samples from the parent distribution and they form the *sample distribution.* In the limit of an infinite number of measurements, the sample distribution becomes the parent distribution.

EXAMPLE 1.2 In a physics laboratory experiment, students drop a ball 50 times and record the time it takes for the ball to fall 2.00 m. One set of observations, corrected for systematic errors, ranges from about 0.59 s to 0.70 s, and some of the observations are identical. Figure 1.2 shows a histogram or frequency plot of these measurements. The height of a data bar represents the number of measurements that fall between the two values indicated by the upper and lower limits of the bar on the abscissa of the plot. (See Appendix D.)

If the distribution results from random errors in measurement, then it is very likely that it can be described in terms of the *Gaussian* or *normal error distribution*, the familiar bell-shaped curve of statistical analysis, which we shall discuss in Chapter 2. A Gaussian curve, based on the mean and standard deviation of these measurements, is plotted as the solid line in Figure 1.2. This curve summarizes the data of the sample distribution in terms of the Gaussian model and provides an estimate of the parent distribution.

The measured data and the curve derived from them clearly do not agree exactly. The coarseness of the experimental histogram distinguishes it at once from the smooth theoretical Gaussian curve. We might imagine that, if the students were to make a great many measurements or combine several sets of measurements so that they could plot the histogram in finer and finer bins, under ideal circumstances the histogram would eventually approach a smooth Gaussian curve. If they were to calculate the parameters from such a large sample, they could determine the parent distribution represented by the dotted curve in Figure 1.2.

It is convenient to think in terms of a *probability density function* $p(x)$, normalized to unit area (*i.e.*, so that the integral of the entire curve is equal to 1) and defined such that in the limit of a very large number N of observations, the number ΔN of observations of the variable x between x and $x + \Delta x$ is given by $\Delta N = Np(x)\Delta x$. The solid and dashed curves in Figure 1.2 have been scaled in this way so that the ordinate values correspond directly to the numbers of observations expected in any range Δx from a 50-event sample and the area under each curve corresponds to the total area of the histogram.

Notation

A number of parameters of the parent distribution have been defined by convention. We use Greek letters to denote them, and Latin letters to denote experimental estimates of them.

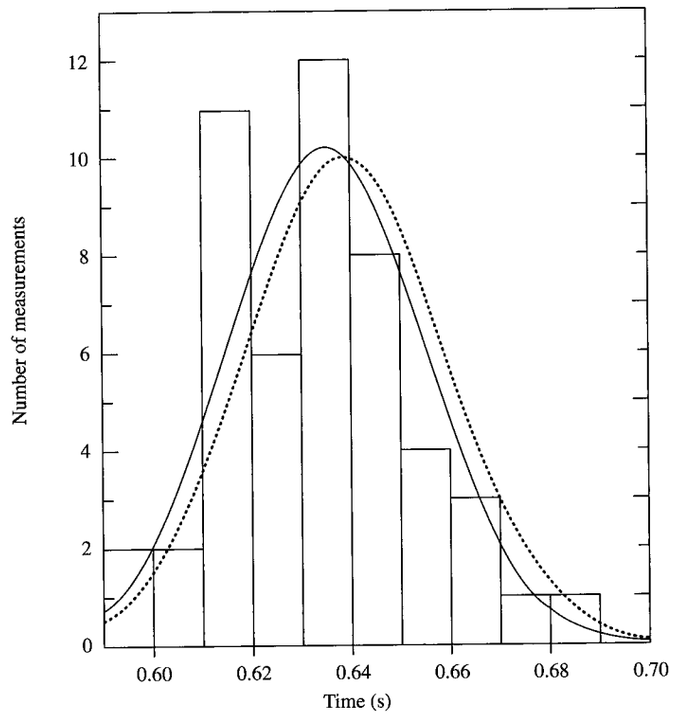


FIGURE 1.2
Histogram of measurements of the time for a ball to fall 2.00 m. The solid Gaussian curve was calculated from the mean ($\bar{T} = 0.635$ s) and standard deviation ($s = 0.020$ s) estimated from these measurements. The dashed curve was calculated from the parent distribution with mean $\mu = 0.639$ s and standard deviation $\sigma = 0.020$ s.

In order to determine the parameters of the parent distribution, we assume that the results of experiments asymptotically approach the parent quantities as the number of measurements approaches infinity; that is, the parameters of the experimental distribution equal the parameters of the parent distribution *in the limit of an infinite number of measurements.* If we specify that there are N observations in a given experiment, then we can denote this by

$$(\text{parent parameter}) = \lim_{N\to\infty} (\text{experimental parameter})$$

If we make N measurements and label them x_1, x_2, x_3, and so forth, up to a final measurement x_N, then we can identify the sum of all these measurements as

$$\sum_{i=1}^{N} x_i \equiv x_1 + x_2 + x_3 + \cdots + x_N$$

where the left-hand side is interpreted as the sum of the observations x_i over the index i from $i = 1$ to $i = N$ inclusive. Because we shall be making frequent use of the

sum over N measurements of various quantities, we simplify the notation by omitting the index whenever we are considering a sum where the index i runs from 1 to N;

$$\sum x_i \equiv \sum_{i=1}^{N} x_i$$

Mean, Median, and Mode

With the preceding definitions, the *mean* $\bar{x}$ of the experimental distribution is given as the sum of N determinations x_i of the quantity x divided by the number of determinations

$$\bar{x} \equiv \frac{1}{N}\sum x_i \tag{1.1}$$

and the mean μ of the parent population is defined as the limit

$$\mu \equiv \lim_{N\to\infty}\left(\frac{1}{N}\sum x_i\right) \tag{1.2}$$

The mean is therefore equivalent to the centroid or *average* value of the quantity x.

The *median* of the parent population $\mu_{1/2}$ is defined as that value for which, in the limit of an infinite number of determinations x_i, half the observations will be less than the median and half will be greater. In terms of the parent distribution, this means that the probability is 50% that any measurement x_i will be larger or smaller than the median

$$P(x_i < \mu_{1/2}) = P(x_i > \mu_{1/2}) = 1/2 \tag{1.3}$$

so that the median line cuts the area of the probability density distribution in half. Because of inconvenience in computation, the median is not often used as a statistical parameter.

The *mode*, or *most probable value* $\mu_{\max}$, of the parent population is that value for which the parent distribution has the greatest value. In any given experimental measurement, this value is the one that is most likely to be observed. In the limit of a large number of observations, this value will probably occur most often

$$P(\mu_{\max}) \geq P(x \neq \mu_{\max}) \tag{1.4}$$

The relationship of the mean, median, and most probable value to one another is illustrated in Figure 1.3. For a symmetrical distribution these parameters would all be equal by the symmetry of their definitions. For an asymmetric distribution such as that of Figure 1.3, the median generally falls between the most probable value and the mean. The most probable value corresponds to the peak of the distribution, and the areas on either side of the median are equal.

Deviations

The *deviation* d_i of any measurement x_i from the mean μ of the parent distribution is defined as the difference between x_i and μ:

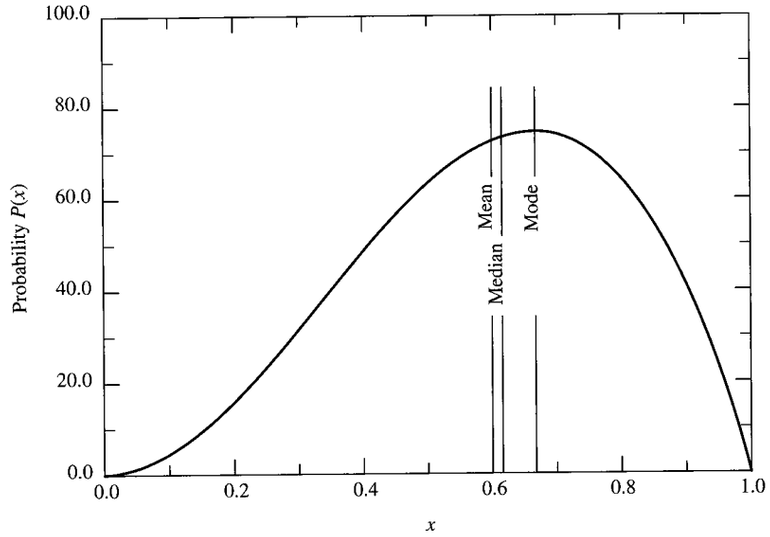


FIGURE 1.3
Asymmetric distribution illustrating the positions of the mean, median, and mode of the variable.

$$d_i \equiv x_i - \mu \tag{1.5}$$

For computational purposes, deviations are generally defined with respect to the mean, rather than the median or most probable value. If μ is the true value of the quantity, d_i is also the true error in x_i.

The average of the deviations $\bar{d}$ must vanish by virtue of the definition of the mean in Equation (1.2):

$$\lim_{N\to\infty} \bar{d} = \lim_{N\to\infty}\left[\frac{1}{N}\sum (x_i - \mu)\right] = \lim_{N\to\infty}\left(\frac{1}{N}\sum x_i\right) - \mu = 0 \tag{1.6}$$

The *average deviation* α, therefore, is defined as the average of the absolute values of the deviations:

$$\alpha \equiv \lim_{N\to\infty}\left[\frac{1}{N}\sum |x_i - \mu|\right] \tag{1.7}$$

The average deviation is a measure of the *dispersion* of the expected observations about the mean. The presence of the absolute value sign makes its use inconvenient for statistical analysis.

A parameter that is easier to use analytically and that can be justified fairly well on theoretical grounds to be a more appropriate measure of the dispersion of the observations is the *standard deviation* σ. The *variance* σ^2 is defined as the limit of the average of the squares of the deviations from the mean μ:

$$\sigma^2 \equiv \lim_{N\to\infty}\left[\frac{1}{N}\sum (x_i - \mu)^2\right] = \lim_{N\to\infty}\left(\frac{1}{N}\sum x_i^2\right) - \mu^2 \tag{1.8}$$

and the standard deviation σ is the square root of the variance. Note that the second form of Equation (1.8) is often described as "the average of the squares minus the square of the average." The standard deviation is the root mean square of the deviations, and is associated with the *second moment* of the data about the mean. The corresponding expression for the variance s^2 of the sample population is given by

$$s^2 \equiv \frac{1}{N-1}\sum (x_i - \bar{x})^2 \tag{1.9}$$

where the factor $N - 1$, rather than N, is required in the denominator to account for the fact that the parameter $\bar{x}$ has been determined from the data and not independently. We note that the symbol σ (instead of s) is often used to represent the best estimate of the standard deviation of the parent distribution determined from a sample distribution.

Significance

The mean μ and the standard deviation, as well as the median, the most probable value, and the average deviation, are all parameters that characterize the information we are seeking when we perform an experiment. Often we wish to describe our distribution in terms of just the mean and standard deviation. The mean may not be exactly equal to the datum in question if the parent distribution is not symmetrical about the mean, but it should have the same characteristics. If a more detailed description is desired, it may be useful to compute higher moments about the mean.

In general, the best we can say about the mean is that it is one of the parameters that specifies the probability distribution: It has the same units as the "true" value and, in accordance with convention, we shall consider it to be the best estimate of the "true" value under the prevailing experimental conditions.

The variance s^2 and the standard deviation s characterize the uncertainties associated with our experimental attempts to determine the "true" values. For a given number of observations, the uncertainty in determining the mean of the parent distribution is proportional to the standard deviation of that distribution. The standard deviation s is, therefore, an appropriate measure of the uncertainty due to fluctuations in the observations in our attempt to determine the "true" value.

Although, in general, the distribution resulting from purely statistical errors can be described well by the two parameters, the mean and the standard deviation, we should be aware that, at distances of a few standard deviations from the mean of an experimental distribution, nonstatistical errors may dominate. In especially severe cases, it may be preferable to describe the spread of the distribution in terms of the average deviation, rather than the standard deviation, because the latter tends to deemphasize measurements that are far from the mean. There are also distributions for which the variance does not exist. The average deviation or some other quantity must be used as a parameter to indicate the spread of the distribution in such cases.

In the following sections, however, we shall be concerned mainly with distributions that result from statistical errors and for which the variance exists.

1.4 MEAN AND STANDARD DEVIATION OF DISTRIBUTIONS

We can define the mean μ and the standard deviation σ in terms of the distribution $p(x)$ of the parent population. The probability density $p(x)$ is defined such that in the limit of a very large number of observations, the fraction dN of observations of the variable x that yield values between x and $x + dx$ is given by $dN = Np(x)\,dx$.

The mean μ is the *expectation value* $\langle x \rangle$ of x, and the variance σ^2 is the expectation value $\langle (x - \mu)^2 \rangle$ of the square of deviations of x from μ. The expectation value $\langle f(x) \rangle$ of any function of x is defined as the weighted average of $f(x)$, over all possible values of the variable x, with each value of $f(x)$ weighted by the probability density distribution $p(x)$.

Discrete Distributions

If the probability function is a discrete function $P(x)$ of the observed value x, we replace the sum over the individual observations Σx_i in Equation (1.2) by a sum over the values of the possible observations multiplied by the number of times these observations are expected to occur. If there are n possible different observable values of the quantity x, which we denote by x_j (where the index j runs from 1 to n with no two values of x_j equal), we should expect from a total of N observations to obtain each observable $NP(x_j)$ times. The mean can then be expressed as

$$\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x_i = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} [x_j NP(x_j)]$$

$$= \lim_{N\to\infty} \sum [x_j P(x_j)] \tag{1.10}$$

Similarly, the variance σ in Equation (1.8) can be expressed in terms of the probability function $P(x)$:

$$\sigma^2 = \lim_{N\to\infty} \sum_{j=1}^{n} [(x_j - \mu)^2 P(x_j)] = \lim_{N\to\infty} \sum_{j=1}^{n} [x_j^2 P(x_j)] - \mu^2 \tag{1.11}$$

In general, the expectation value of any function of $f(x)$ is given by

$$\langle f(x) \rangle = \sum_{j=1}^{n} [f(x_j) P(x_j)] \tag{1.12}$$

Continuous Distributions

If the probability density function is a continuous smoothly varying function $p(x)$ of the observed value x, we replace the sum over the individual observations by an integral over all values of x multiplied by the probability $p(x)$. The mean μ becomes the first moment of the parent distribution

$$\mu = \int_{-\infty}^{\infty} xp(x)\,dx \tag{1.13}$$

and the variance σ^2 becomes the second central product moment

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 p(x)\,dx = \int_{-\infty}^{\infty} x^2 p(x)\,dx - \mu^2 \tag{1.14}$$

The expectation value of any function of x is

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)p(x)\,dx \tag{1.15}$$

What is the connection between the probability distribution of the parent population and an experimental sample we obtain? We have already seen that the uncertainties of the experimental conditions preclude a determination of the "true" values themselves. As a matter of fact, there are three levels of abstraction between the data and the information we seek:

1. From our experimental data points we can determine a sample frequency distribution that describes the way in which these particular data points are distributed over the range of possible data points. We use $\bar{x}$ to denote the mean of the data and s^2 to denote the sample variance. The shape and magnitude of the sample distribution vary from sample to sample.
2. From the parameters of the sample probability distribution we can estimate the parameters of the probability distribution of the parent population of possible observations. Our best estimate for the mean μ is the mean of the sample distribution $\bar{x}$, and the best estimate for the variance σ^2 is the sample variance s^2. Even the shape of this parent distribution must be estimated or assumed.
3. From the estimated parameters of the parent distribution we estimate the results sought. In general, we shall assume that the estimated parameters of the parent distribution are equivalent to the "true" values, but the estimated parent distribution is a function of the experimental conditions as well as the "true" values, and these may not necessarily be separable.

Let us refer again to Figure 1.2, which shows a histogram of time interval measurements and two Gaussian curves, a solid curve based on the parameters $\overline{T} = 0.635$ s and $s = 0.020$ s, which were determined experimentally from the data displayed in the histogram, and a dotted curve based on the parameters $\mu = 0.639$ s and $\sigma = 0.020$ s of the parent distribution. (Although, in general we don't know the properties of the parent distribution, they could have been estimated to high precision in another experiment involving many more measurements.) Comparing the two curves, we observe a slight difference between the experimental mean $\overline{T}$ and the "true" mean μ, and between s and σ.

By considering the data to be a *sample* from the parent population with the values of the observations distributed according to the parent population, we can estimate the shape and dispersion of the parent distribution to obtain useful information on the precision and reliability of our results. Thus, we consider the sample

mean $\overline{T}$ to be our best estimate from the data of the mean μ, and we consider the sample variance s^2 to be our best estimate from the data of the variance σ^2, from which we can estimate the uncertainty in our estimate of μ.

SUMMARY

Errors: Difference between measured and "true" values. Generally applied to the uncertainty in a measurement. Not *blunders* or *mistakes.*
Systematic error: Reproducible inaccuracy introduced by faulty equipment, calibration, or technique.
Random error: Indefiniteness of result introduced by finite precision of measurement or statistical variations. Measure of fluctuation after repeated experimentation.
Uncertainty: Magnitude of error that is estimated to have been made in determination of results.
Accuracy: Measure of how close the result of an experiment comes to the "true" value.
Precision: Measure of how carefully the result is determined without reference to any "true" value.
Significant figures:

1. The leftmost nonzero digit is the most significant digit.
2. If there is no decimal point, the rightmost nonzero digit is the least significant digit.
3. If there is a decimal point, the rightmost digit is the least significant digit, even if it is zero.
4. All digits between the least and most significant digits are counted as significant digits.

Roundoff: Truncate the number to the specified number of significant digits and treat the excess digits as a decimal fraction.

1. If the fraction is greater than ½, increment the new least significant digit.
2. If the fraction is less than ½, do not increment.
3. If the fraction equals ½, increment the least significant digit only if it is odd.

Parent population: Hypothetical infinite set of data points of which the experimental data points are assumed to be a random sample.
Parent distribution: Probability distribution of the parent population from which the sample data are chosen.
Expectation value $f(x)$: Weighted average of a function $f(x)$ over all values of x:

$$\langle f(x)\rangle = \lim_{N\to\infty}\left[\frac{1}{N}\sum f(x_i)\right] = \sum_{i=1}^{n}[f(x)P(x_j)] = \int_{-\infty}^{\infty} f(x)P(x)\,dx$$

Median $\mu_{1/2}$: $P(x_i < \mu_{1/2}) = P(x > \mu_{1/2}) = \frac{1}{2}$

Most probable value μ_{max}: $P(\mu_{max}) \geq P(x \neq \mu_{max})$

Mean: $\mu \equiv \langle x\rangle$

Average deviation: $\alpha \equiv \langle |x_i - \mu|\rangle$

Variance: $\sigma^2 \equiv \langle (x_i - \mu)^2\rangle = \langle x^2\rangle - \mu^2$

Standard deviation: $\sigma = \sqrt{\sigma^2}$

Sample mean: $\bar{x} = (1/N)\Sigma x_i$

Sample variance: $s^2 = \dfrac{1}{(N-1)}\Sigma(x_i - \bar{x})^2$

EXERCISES

1.1. How many significant features are there in the following numbers?
(*a*) 976.45 (*b*) 84,000 (*c*) 0.0094 (*d*) 301.07
(*e*) 4.000 (*f*) 10 (*g*) 5280 (*h*) 400.
(*i*) 4.00×10^2 (*j*) 3.010×10^4

1.2. What is the most significant figure in each of the numbers in Exercise 1.1? What is the least significant?

1.3. Round off each of the numbers in Exercise 1.1 to two significant digits.

1.4. Find the mean, median, and most probable value of x for the following data (from rolling dice).

i	x_i	i	x_i	i	x_i	i	x_i	i	x_i
1	3	6	8	11	12	16	6	21	5
2	7	7	9	12	8	17	7	22	10
3	3	8	7	13	6	18	8	23	8
4	7	9	5	14	6	19	9	24	8
5	12	10	7	15	7	20	8	25	8

1.5. Find the mean, median, and most probable grade from the following set of grades. Group them to find the most probable value.

i	x_i	i	x_i	i	x_i	i	x_i
1	73	11	73	21	69	31	56
2	91	12	46	22	70	32	94
3	72	13	64	23	82	33	51
4	81	14	61	24	90	34	79
5	82	15	50	25	63	35	63
6	46	16	89	26	70	36	87
7	89	17	91	27	94	37	54
8	75	18	82	28	44	38	100
9	62	19	71	29	100	39	72
10	58	20	76	30	88	40	81

1.6. Calculate the standard deviation of the data of Exercise 1.4.

1.7. Calculate the standard deviation of the data of Exercise 1.5.
1.8. Justify the second equality in Equations (1.8) and (1.14).
1.9. Carefully measure in centimeters the length of the cover of this book along the bound edge. Estimate the uncertainty in your measurement. Quote your answer with its uncertainty in decimal form and in scientific notation.

CHAPTER 2

PROBABILITY DISTRIBUTIONS

Of the many probability distributions that are involved in the analysis of experimental data, three play a fundamental role: the *binomial distribution*, the *Poisson distribution*, and the *Gaussian distribution*. Of these, the Gaussian, or normal error, distribution is undoubtedly the most important in statistical analysis of data. Practically, it is useful because it seems to describe the distribution of random observations for many experiments, as well as describing the distributions obtained when we try to estimate the parameters of most other probability distributions.

The Poisson distribution is generally appropriate for counting experiments where the data represent the number of items or events observed per unit interval. It is important in the study of random processes such as those associated with the radioactive decay of elementary particles or nuclear states, and is also applied to data that have been sorted into ranges to form a frequency table or a histogram.

The binomial distribution is generally applied to experiments in which the result is one of a small number of possible final states, such as the number of "heads" or "tails" in a series of coin tosses, or the number of particles scattered forward or backward relative to the direction of the incident particle in a particle physics experiment. Because both the Poisson and the Gaussian distributions can be considered as limiting cases of the binomial distribution, we shall devote some attention to the derivation of the binomial distribution from basic considerations.

2.1 BINOMIAL DISTRIBUTION

Suppose we toss a coin in the air and let it land. There is a 50% probability that it will land heads up and a 50% probability that it will land tails up. By this we mean that if we continue tossing a coin repeatedly, the fraction of times that it lands with heads up will asymptotically approach ½, indicating that there was a probability of

½ of doing so. For any given toss, the probability cannot determine whether or not it will land heads up; it can only describe how we should expect a large number of tosses to be divided into two possibilities.

Suppose we toss two coins at a time. There are now four different possible permutations of the way in which they can land: both heads up, both tails up, and two mixtures of heads and tails depending on which one is heads up. Because each of these permutations is equally probable, the probability for any choice of them is ¼ or 25%. To find the probability for obtaining a particular mixture of heads and tails, without differentiating between the two kinds of mixtures, we must add the probabilities corresponding to each possible kind. Thus, the total probability of finding either head up and the other tail up is ½. Note that the sum of the probabilities for all possibilities (¼ + ¼ + ¼ + ¼) is always equal to 1 because *something* is bound to happen.

Let us extrapolate these ideas to the general case. Suppose we toss n coins into the air, where n is some integer. Alternatively, suppose that we toss one coin n times. What is the probability that exactly x of these coins will land heads up, without distinguishing which of the coins actually belongs to which group? We can consider the probability $P(x; n)$ to be a function of the number n of coins tossed and of the number x of coins that land heads up. For a given experiment in which n coins are tossed, this probability $P(x; n)$ will vary as a function of x. Of course, x must be an integer for any physical experiment, but we can consider the probability to be smoothly varying with x as a continuous variable for mathematical purposes.

Permutations and Combinations

If n coins are tossed, there are 2^n different possible ways in which they can land. This follows from the fact that the first coin has two possible orientations, for each of these the second coin also has two such orientations, for each of these the third coin also has two, and so on. Because each of these possibilities is equally probable, the probability for any one of these possibilities to occur at any toss of n coins is $1/2^n$.

How many of these possibilities will contribute to our observations of x coins with heads up? Imagine two boxes, one labeled "heads" and divided into x slots, and the other labeled "tails." We shall consider first the question of how many permutations of the coins result in the proper separation of x in one box and $n - x$ in the other; then we shall consider the question of how many combinations of these permutations should be considered to be different from each other.

In order to enumerate the number of *permutations* $Pm(n, x)$, let us pick up the coins one at a time from the collection of n coins and put x of them into the "heads" box. We have a choice of n coins for the first one we pick up. For our second selection we can choose from the remaining $n - 1$ coins. The range of choice is diminished until the last selection of the xth coin can be made from only $n - x + 1$ remaining coins. The total number of choices for coins to fill the x slots in the "heads" box is the product of the numbers of individual choices:

$$Pm(n, x) = n(n-1)(n-2)\cdots(n-x+2)(n-x+1) \tag{2.1}$$

This expansion can be expressed more easily in terms of factorials

$$Pm(n, x) = \frac{n!}{(n-x)!} \tag{2.2}$$

So far we have calculated the number of permutations $Pm(n, x)$ that will yield x coins in the "heads" box and $n - x$ coins in the "tails" box, with the provision that we have identified which coin was placed in the "heads" box first, which was placed in second, and so on. That is, we have *ordered* the x coins in the "heads" box. In our computation of 2^n different possible permutations of the n coins, we are only interested in which coins landed heads up or heads down, not which landed first. Therefore, we must consider contributions different only if there are different coins in the two boxes, not if the x coins within the "heads" box are permuted into different time orderings.

The number of different *combinations* $C(n, x)$ of the permutations in the preceding enumeration results from combining the $x!$ different ways in which x coins in the "heads" box can be permuted within the box. For every $x!$ permutations, there will be only one new combination. Thus, the number of different combinations $C(n, x)$ is the number of permutations $Pm(n, x)$ divided by the degeneracy factor $x!$ of the permutations:

$$C(n, x) = \frac{Pm(n, x)}{x!} = \frac{n!}{x!(n-x)!} = \binom{n}{x} \tag{2.3}$$

This is the number of different possible combinations of n items taken x at a time, commonly referred to as $\binom{n}{x}$ or "n over x."

Probability

The probability $P(x; n)$ that we should observe x coins with heads up and $n - x$ with tails up is the product of the number of different combinations $C(n, x)$ that contribute to that set of observations multiplied by the probability for each of the combinations to occur, which we have found to be $(½)^n$.

Actually, we should separate the probability for each combination into two parts: one part is the probability $p^x = (½)^x$ for x coins to be heads up; the other part is the probability $q^{n-x} = (1 - ½)^{n-x} = (½)^{n-x}$ for the other $n - x$ coins to be tails up. For symmetrical coins, the product of these two parts $p^x q^{n-x} = (½)^n$ is the probability of the combination with x coins heads up and $n - x$ coins tails up. In the general case, the probability p of success for each item is not equal in magnitude to the probability $q = 1 - p$ for failure. For example, when tossing a die, the probability that a particular number will show is $p = 1/6$, while, the probability of its not showing is $q = 1 - 1/6 = 5/6$ so that $p^x q^{n-x} = (1/6)^x \times (5/6)^n$.

With these definitions of p and q, the probability $P_B(x; n, p)$ for observing x of the n items to be in the state with probability p is given by the *binomial distribution*

$$P_B(x; n, p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \tag{2.4}$$

where $q = 1 - p$. The name for the binomial distribution comes from the fact that the coefficients $P_B(x; n, p)$ are closely related to the binomial theorem for the expansion of a power of a sum. According to the binomial theorem,

$$(p + q)^n = \sum_{x=0}^{n} \left[\binom{n}{x} p^x q^{n-x} \right] \tag{2.5}$$

The $(j + 1)$th term, corresponding to $x = j$, of this expansion, therefore, is equal to the probability $P_B(j; n, p)$. We can use this result to show that the binomial distribution coefficients $P_B(x; n, p)$ are normalized to a sum of 1. The right-hand side of Equation (2.5) is the sum of probabilities over all possible values of x from 0 to n and the left-hand side is just $1^n = 1$.

Mean and Standard Deviation

The mean of the binomial distribution is evaluated by combining the definition of μ in Equation (1.10) with the formula for the probability function of Equation (2.4):

$$\mu = \sum_{x=0}^{n} \left[x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right] = np \tag{2.6}$$

We interpret this to mean that if we perform an experiment with n items and observe the number x of successes, after a large number of repeated experiments the average $\bar{x}$ of the number of successes will approach a mean value μ given by the probability for success of each item p times the number of items n. In the case of coin tossing where $p = \frac{1}{2}$, we should expect on the average to observe half the coins land heads up, which seems eminently reasonable.

The variance σ^2 of a binomial distribution is similarly evaluated by combining Equations (1.11) and (2.4):

$$\sigma^2 = \sum_{x=0}^{n} \left[(x-\mu)^2 \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right] = np(1-p) \tag{2.7}$$

The evaluation of these sums is left as an exercise. We are mainly interested in the results, which are remarkably simple.

If the probability for a single success p is equal to the probability for failure $p = q = \frac{1}{2}$, then the distribution is symmetric about the mean μ, and the median $\mu_{1/2}$ and the most probable value are both equal to the mean. In this case, the variance σ^2 is equal to half the mean: $\sigma^2 = \mu/2$. If p and q are not equal, the distribution is asymmetric with a smaller variance.

Example 2.1. Suppose we toss 10 coins into the air a total of 100 times. With each coin toss we observe the number of coins that land heads up and denote that number by x_i, where i is the number of the toss; i ranges from 1 to 100 and x_i can be any integer from 0 to 10. The probability function governing the distribution of the observed values of x is given by the binomial distribution $P_B(x; n, p)$ with $n = 10$ and $p = \frac{1}{2}$. This is the parent distribution and is not affected by the number N of repeated procedures in the experiment.

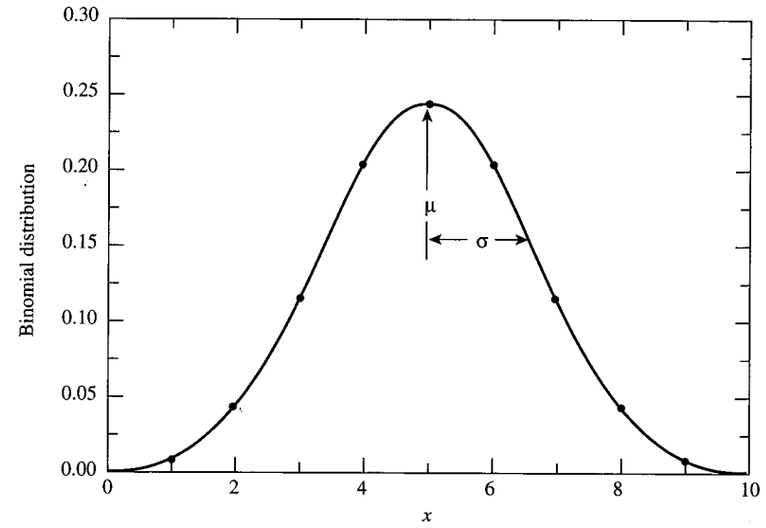


FIGURE 2.1
Binomial distribution for $\mu = 5.0$ and $p = \frac{1}{2}$ shown as a continuous curve although the function is only defined at the discrete points indicated by the round dots.

The parent distribution $P_B(x; 10, \frac{1}{2})$ is shown in Figure 2.1 as a smooth curve drawn through discrete points. The mean μ is given by Equation (2.6):

$$\mu = np = 10(\tfrac{1}{2}) = 5$$

the standard deviation σ is given by Equation (2.7):

$$\sigma = \sqrt{np(1-p)} = \sqrt{10(\tfrac{1}{2})(\tfrac{1}{2})} = \sqrt{2.5} \simeq 1.58$$

The curve is symmetric about its peak at the mean so that approximately 25% of the throws yield five heads and five tails, about 20% yield four heads and six tails, and the same fraction yields six heads and four tails. The magnitudes of the points are such that the sum of the probabilities over all ten points is equal to 1.

Example 2.2. Suppose we roll ten dice. What is the probability that x of these dice will land with the 1 up? If we throw one die, the probability of its landing with 1 up is $p = \frac{1}{6}$. If we throw ten dice, the probability for x of them to land with 1 up is given by the binomial distribution $P_B(x; n, p)$ with $n = 10$ and $p = \frac{1}{6}$:

$$P_B\left(x; 10, \frac{1}{6}\right) = \frac{10!}{x!(10-x)!} \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{10-x}$$

This distribution is illustrated in Figure 2.2 as a smooth curve drawn through discrete points. The mean and standard deviation are

$$\mu = 10/6 \simeq 1.67$$

and

$$\sigma = \sqrt{10(1/6)(5/6)} \simeq 1.18$$

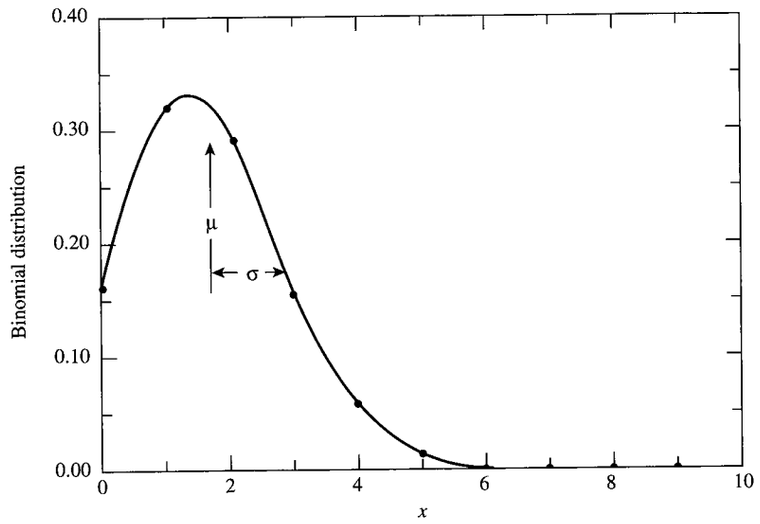


FIGURE 2.2
Binomial distribution for $\mu = 10/6$ and $p = 1/6$ shown as a continuous curve.

The distribution is not symmetric about the mean or about any other point. The most probable value is $x = 1$, but the peak of the smooth curve occurs for a slightly larger value of x.

Example 2.3 A particle physicist makes some preliminary measurements of the angular distribution of K mesons scattered from a liquid hydrogen target. She knows that there should be equal numbers of particles scattered forward and backward in the center-of-mass system of the particles. She measures 1000 interactions and finds that 472 scatter forward and 528 backward. What uncertainty should she quote in these numbers?

The uncertainty is given by the standard deviation from Equation (2.7),

$$\sigma = \sqrt{np(1-p)} = \sqrt{1000(\tfrac{1}{2})(\tfrac{1}{2})} = \sqrt{250} \simeq 15.8$$

Thus, she could quote

$$f_F = (472 \pm 15.8)/1000 = 0.472 \pm 0.15$$

for the fraction of particles scattered in the forward direction and

$$f_B = (528 \pm 15.8)/1000 = 0.528 \pm 0.15$$

for the fraction scattered backward.

Note that the uncertainties in the numbers scattering forward and backward must be the same because losses from one group must be made up in the other.

If the experimenter did not know the a priori probabilities of scattering forward and backward, she would have to estimate p and q from her measurements; that is,

$$p \simeq 472/1000 = 0.472$$

and

$$q \simeq 528/1000 = 0.528$$

She would then calculate

$$s \simeq \sqrt{1000\,(0.472)(0.528)} = \sqrt{249.2} = 15.8$$

For probability p near 50%, the standard deviation is relatively insensitive to uncertainties in the experimental determination of p.

2.2 POISSON DISTRIBUTION

The Poisson distribution represents an approximation to the binomial distribution for the special case where the average number of successes is much smaller than the possible number; that is, when $\mu \ll n$ because $p \ll 1$. For such experiments the binomial distribution correctly describes the probability $P_B(x; n, p)$ of observing x events per time interval out of n possible events, each of which has a probability p of occurring, but the large number n of possible events makes exact evaluation from the binomial distribution impractical. Furthermore, neither the number n of possible events nor the probability p for each is usually known. What may be known instead is the average number of events μ expected in each time interval or its estimate $\bar{x}$. The Poisson distribution provides an analytical form appropriate to such investigations that describes the probability distribution in terms of just the variable x and the parameter μ.

Let us consider the binomial distribution in the limiting case of $p \ll 1$. We are interested in its behavior as n becomes infinitely large while the mean $\mu = np$ remains constant. Equation (2.4) for the probability function of the binomial distribution can be written as

$$P_B(x; n, p) = \frac{1}{x!}\frac{n!}{(n-x)!}p^x(1-p)^{-x}(1-p)^n \tag{2.8}$$

If we expand the second factor

$$\frac{n!}{(n-x)!} = n(n-1)(n-2)\cdots(n-x-2)(n-x-1) \tag{2.9}$$

we can consider it to be the product of x individual factors, each of which is very nearly equal to n because $x \ll$ n in the region of interest. The second factor in Equation (2.8) thus asymptotically approaches n^x. The product of the second and third factors then becomes $(np)^x = \mu^x$. The fourth factor is approximately equal to $1 + px$, which tends to 1 as p tends to 0.

The last factor can be rearranged by substituting μ/p for n and expanding the expression to show that it asymptotically approaches $e^{-\mu}$:

$$\lim_{p\to 0}(1-p)^x = \lim_{p\to 0}[(1-p)^{1/p}]^\mu = \left(\frac{1}{e}\right)^\mu = e^{-\mu} \tag{2.10}$$

Combining these approximations, we find that the binomial distribution probability function $P_B(x;\ n,\ p)$ asymptotically approaches the *Poisson distribution* $P_P(x;\ \mu)$ as p approaches 0:

$$\lim_{p\to 0} P_B(x; n, p) = P_P(x; \mu) \equiv \frac{\mu^x}{x!} e^{-\mu} \tag{2.11}$$

Because this distribution is an approximation to the binomial distribution for $p \ll 1$, the distribution is asymmetric about its mean μ and will resemble that of Figure 2.2. Note that $P_p(x; \mu)$ does not become 0 for $x = 0$ and is not defined for negative values of x. This restriction is not troublesome for counting experiments because the number of counts per unit time interval can never be negative.

Derivation

The Poisson distribution can also be derived for the case where the number of events observed is small compared to the total possible number of events.[1] Assume that the average rate at which events of interest occur is constant over a given interval of time and that event occurrences are randomly distributed over that interval. Then, the probability dP of observing no events in a time interval dt is given by

$$dP(0; t, \tau) = -P(0; t, \tau)\frac{dt}{\tau} \tag{2.12}$$

where $P(x; t, \tau)$ is the probability of observing x events in the time interval dt, τ is a constant proportionality factor that is associated with the mean time between events, and the minus sign accounts for the fact that increasing the differential time interval dt decreases the probability proportionally. Integrating this equation yields the probability of observing no events within a time t to be

$$P(0; t, \tau) = P_0 e^{-t/\tau} \tag{2.13}$$

where P_0, the constant of integration, is equal to 1 because $P(0; t, \tau) = 1$ at $t = 0$.

The probability $P(x; t, \tau)$ for observing x events in the time interval τ can be evaluated by integrating the differential probability

$$d^x P(x; t, \tau) = \frac{e^{-t/\tau}}{x!} \prod_{i=1}^{x} \frac{dt_i}{\tau} \tag{2.14}$$

which is the product of the probabilities of observing each event in a different interval dt_i and the probability $e^{-t/\tau}$ of not observing any other events in the remaining time. The factor of $x!$ in the denominator compensates for the ordering implicit in the probabilities $dP_i(1, t, \tau)$ as discussed in the preceding section on permutations and combinations.

Thus, the probability of observing x events in the time interval t is obtained by integration

$$P_P(x; \mu) = P(x; t, \tau) = \frac{e^{-t/\tau}}{x!}\left(\frac{t}{\tau}\right)^x \tag{2.15}$$

[1]This derivation follows that of Orear (1958), pages 21–22.

or

$$P_P(x; \mu) = \frac{\mu^x}{x!} e^{-\mu} \tag{2.16}$$

which is the expression for the Poisson distribution, where $\mu = t/\tau$ is the average number of events observed in the time interval t. Equation (2.16) represents a normalized probability function; that is, the sum of the function evaluated at each of the allowed values of the variable x is unity:

$$\sum_{x=0}^{\infty} P_P(x, \mu) = \sum_{x=0}^{\infty} \frac{\mu^x}{x!} e^{-\mu} = e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} = e^{-\mu} e^{\mu} = 1 \tag{2.17}$$

Mean and Standard Deviation

The Poisson distribution, like the binomial distribution, is a *discrete* distribution. That is, it is defined only at integral values of the variable x, although the parameter μ is a positive, real number. The mean of the Poisson distribution is actually the parameter μ that appears in the probability function $P_P(x; \mu)$ of Equation (2.16). To verify this, we can evaluate the expectation value $\langle x \rangle$ of x:

$$\langle x \rangle = \sum_{x=0}^{\infty} \left(x \frac{\mu^x}{x!} e^{-\mu} \right) = \mu e^{-\mu} \sum_{x=1}^{\infty} \frac{\mu^{x-1}}{(x-1)!} = \mu e^{-\mu} \sum_{y=0}^{\infty} \frac{\mu^y}{y!} = \mu \tag{2.18}$$

To find the standard deviation σ, the expectation value of the square of the deviations can be evaluated:

$$\sigma^2 = \langle (x - \mu)^2 \rangle = \sum_{x=0}^{\infty} \left[(x - \mu)^2 \frac{\mu^x}{x!} e^{-\mu} \right] = \mu \tag{2.19}$$

Thus, the standard deviation σ is equal to the square root of the mean μ and the Poisson distribution has only a single parameter, μ.

Computation of the Poisson distribution by Equation (2.16) can be limited by the factorial function in the denominator. The problem can be avoided by using logarithms or by using the recursion relations

$$P(0; \mu) = e^{-\mu} \qquad P(x; \mu) = \frac{\mu}{x} P(x - 1; \mu) \tag{2.20}$$

This form has the disadvantage that, in order to calculate the function for particular values of x and μ, the function must be calculated at all lower values of x as well. However, if the function is to be summed from $x = 0$ to some upper limit to obtain the summed probability or to generate the distribution for a Monte Carlo calculation (Chapter 5), the function must be calculated at all lower values of x anyway.

Example 2.4 As part of an experiment to determine the mean life of radioactive isotopes of silver, students detected background counts from cosmic rays. (See Example 8.1.) They recorded the number of counts in their detector for a series of 100 2-s intervals, and found that the mean number of counts was 1.69 per interval. From the mean they estimated the standard deviation to be $\sigma = \sqrt{1.69} = 1.30$, compared to $s = 1.29$ from a direct calculation with Equation (1.9).

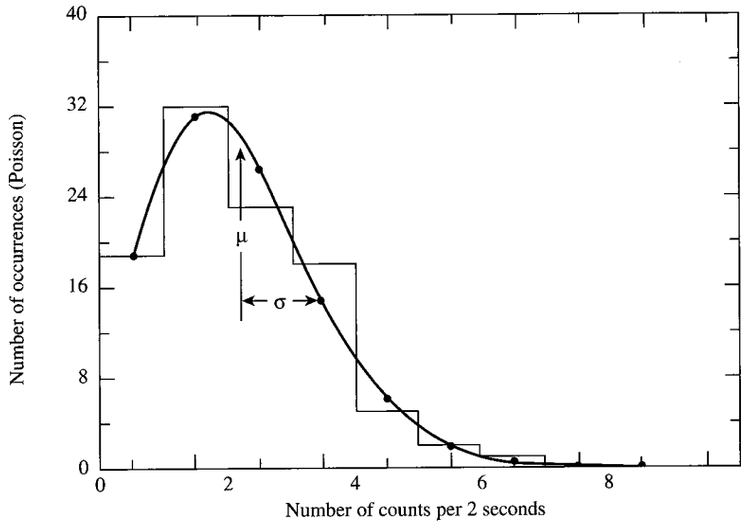


FIGURE 2.3
Histogram of counts in a cosmic ray detector. The Poisson distribution is an estimate of the parent distribution based on the measured mean $\bar{x} = 1.69$. It is shown as a continuous curve although the function is only defined at the discrete points indicated by the round dots.

The students then repeated the exercise, this time recording the number of counts in 15-s intervals for 60 intervals, obtaining a mean of 11.48 counts per interval, with standard deviations $\sigma = \sqrt{11.48} = 3.17$ and $s = 3.39$.

Histograms of the two sets of data are shown in Figures 2.3, and. 2.4. The calculated mean in each case was used as an estimate of the mean of the parent distribution to calculate a Poisson distribution for each data set. The distributions are shown as continuous curves, although only the points at integral values of the abscissa are physically significant.

The asymmetry of the distribution in Figure 2.3 is obvious, as is the fact that the mean μ does not coincide with the most probable value of x at the peak of the curve. The curve of Figure 2.4, on the other hand, is almost symmetric about its mean and the data are consistent with the curve. As μ increases, the symmetry of the Poisson distribution increases and the distribution becomes indistinguishable from the Gaussian distribution.

Summed Probability

We may want to know the probability of obtaining a sample value of x between limits x_1 and x_2 from a Poisson distribution with mean μ. This probability is obtained by summing the values of the function calculated at the integral values of x between the two integral limits x_1 and x_2,

$$S_P(x_1, x_2; \mu) = \sum_{x_1}^{x_2} P_P(x; \mu) \tag{2.21}$$

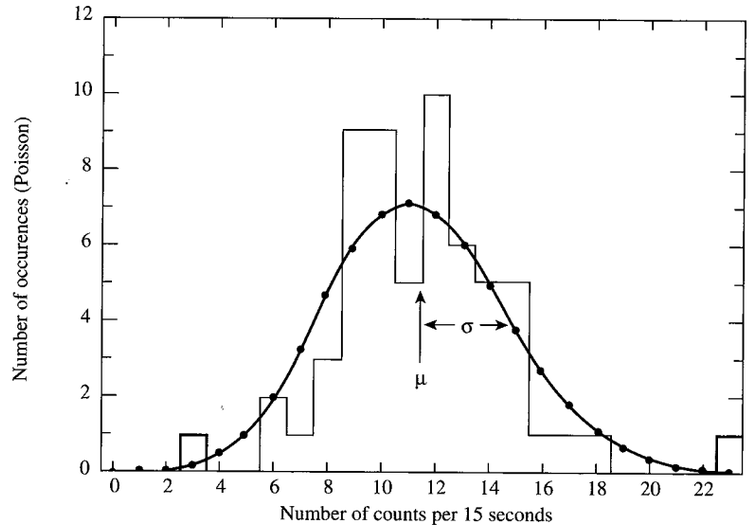


FIGURE 2.4
Histogram of counts in a cosmic ray detector. The Poisson distribution, shown as a continuous curve, is an estimate of the parent distribution based on the measured mean $\bar{x} = 11.48$. Only the calculated points indicated by the round dots are defined.

More likely, we may want to find the probability of recording n or more events in a given interval when the mean number of events is μ. This is just the sum

$$S_P(n, \infty; \mu) = \sum_{x=n}^{\infty} P_P(x; \mu) = 1 - \sum_{x=0}^{n-1} P_P(x; \mu) = 1 - e^{-\mu} \sum_{x=0}^{n-1} \frac{\mu^x}{x!} \tag{2.22}$$

In Example 2.4, the mean number of counts recorded in a 15-s time interval was $\bar{x} = 11.48$. In one of the intervals, 23 counts were recorded. From Equation (2.22), the probability of collecting 23 or more events in a single 15-s time interval is ~ 0.0018, and the probability of this occurring in any one of 60 15-s time intervals is just the complement of the joint probability that 23 or more counts *not* be observed in any of the 60 time intervals, or $p \simeq 1 - (1 - 0.0018)^{60} \simeq 0.10$, or about 10%.

For large values of μ, the probability sum of Equation (2.22) may be approximated by an integral of the Gaussian function.

2.3 GAUSSIAN OR NORMAL ERROR DISTRIBUTION

The Gaussian distribution is an approximation to the binomial distribution for the special limiting case where the number of possible different observations n becomes infinitely large and the probability of success for each is finitely large so $np \gg 1$. It is also, as we observed, the limiting case for the Poisson distribution as μ becomes large.

There are several derivations of the Gaussian distribution from first principles, none of them as convincing as the fact that the distribution is reasonable, that it has a fairly simple analytic form, and that it is accepted by convention and experimentation to be the most likely distribution for most experiments. In addition, it has the satisfying characteristic that the most probable estimate of the mean μ from a random sample of observations x is the average of those observations $\bar{x}$.

Characteristics

The Gaussian probability density is defined as

$$p_G = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \tag{2.23}$$

This is a continuous function describing the probability of obtaining the value x in a random observation from a parent distribution with parameters μ and σ, corresponding to the mean and standard deviation, respectively. Because the distribution is continuous, we must define an interval in which the value of the observation x will fall. The probability density function is properly defined such that the probability $dP_G(x; \mu, \sigma)$ that the value of a random observation will fall within an interval dx around x is given by

$$dP_G(x; \mu, \sigma) = p_G(x; \mu, \sigma)dx \tag{2.24}$$

considering dx to be an infinitesimal differential, and the probability density function to be normalized, so that

$$\int_{x=-\infty}^{x=\infty} dP_G(x; \mu, \sigma) = \int_{x=-\infty}^{x=\infty} p_G(x; \mu, \sigma)dx \tag{2.25}$$

The width of the curve is determined by the value of σ, such that for $x = \mu + \sigma$, the height of the curve is reduced to $e^{-1/2}$ of its value at the peak:

$$p_G(x; \mu \pm \sigma, \sigma) = e^{-1/2} p_G(\mu; \mu, \sigma) \tag{2.26}$$

The shape of the Gaussian distribution is shown in Figure 2.5. The curve displays the characteristic bell shape and symmetry about the mean μ.

We can characterize a distribution by its *full-width at half maximum* Γ, often referred to as the *half-width*, defined as the range of x between values at which the probability $p_G(x; \mu, \sigma)$ is half its maximum value:

$$p_G(\mu \pm \tfrac{1}{2}\Gamma, \mu, \sigma) = \tfrac{1}{2} p_G(\mu; \mu, \sigma) \tag{2.27}$$

With this definition, we can determine from Equation (2.23) that

$$\Gamma = 2.354\,\sigma \tag{2.28}$$

As illustrated in Figure 2.5, tangents drawn along a portion of steepest descent of the curve intersect the curve at the $e^{-1/2}$ points $x = \mu \pm \sigma$ and intersect the x axis at the points $x = \mu \pm 2\sigma$.

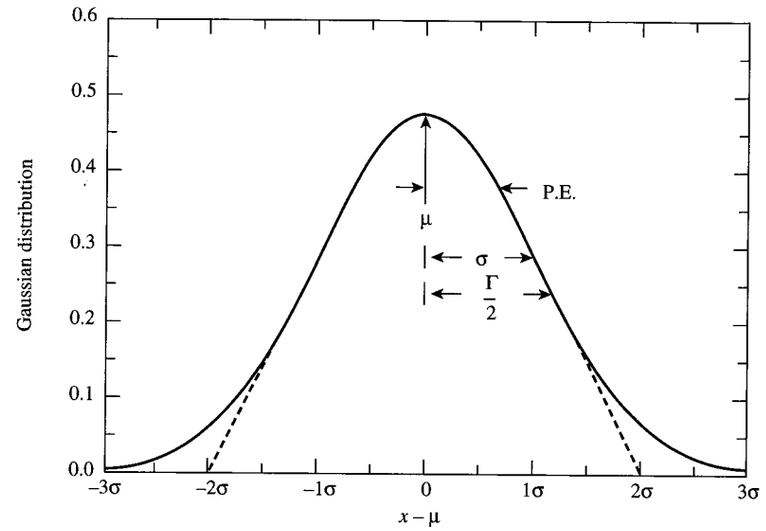


FIGURE 2.5
Gaussian probability distribution illustrating the relation of μ, σ, Γ, and P.E. to the curve. The curve has unit area.

Standard Gaussian Distribution

It is generally convenient to use a standard form of the Gaussian equation obtained by defining the dimensionless variable $z = (x - \mu)/\sigma$, because with this change of variable, we can write

$$p_G(z)\,dz = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \tag{2.29}$$

Thus, from a single computer routine or a table of values of $p_G(z)$, we can find the Gaussian probability function $p_G(x; \mu, \sigma)$ for all values of the parameters μ and σ by changing the variable and scaling the function by $1/\sigma$ to preserve the normalization.

Mean and Standard Deviation

The parameters μ and σ in Equation (2.23) for the Gaussian probability density distribution correspond to the mean and standard deviation of the function. This equivalence can be verified by calculating μ and σ with Equations (1.13) and (1.14) as the expectation values for the Gaussian function of x and $(x - \mu)^2$, respectively.

For a finite data sample, which is expected to follow the Gaussian probability density distribution, the mean and standard deviation can be calculated directly from Equations (1.1) and (1.9). The resulting values of $\bar{x}$ and s will be estimates of the mean μ and standard deviation σ. Values of $\bar{x}$ and s, obtained in this way from the original 50 time measurements in Example 1.2, were used as estimates of μ and

σ in Equation (2.23) to calculate the solid Gaussian curve in Figure 1.2. The curve was scaled to have the same area as the histogram. The curve represents our estimate of the parent distribution based on our measurements of the sample.

Integral Probability

We are often interested in knowing the probability that a measurement will deviate from the mean by a specified amount Δx or greater. The answer can be determined by evaluating numerically the integral

$$P_G(\Delta x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu-\Delta x}^{\mu+\Delta x} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx \tag{2.30}$$

which gives the probability that any random value of x will deviate from the mean by less than $\pm\Delta x$. Because the probability function $P_G(x; \mu, \sigma)$ is normalized to unity, the probability that a measurement will deviate from the mean by *more* than Δx is just $1 - P_G(\Delta x; \mu, \sigma)$. Of particular interest are the probabilities associated with deviations of σ, 2σ, and so forth from the mean, corresponding to 1, 2, and so on standard deviations. We may also be interested in the probable error (σ_{pe}), defined to be the absolute value of the deviation $|x - \mu|$ such that the probability for the deviation of any random observation $|x_i - \mu|$ is less than ½. That is, half the observations of an experiment would be expected to fall within the boundaries denoted by $\mu \pm \sigma_{pe}$.

If we use the standard form of the Gaussian distribution of Equation (2.29), we can calculate the integrated probability $P_G(z)$ in terms of the dimensionless variable $z = (x - \mu)/\sigma$,

$$P_G(z) = \frac{1}{\sqrt{2\pi}} \int_{-\Delta z}^{\Delta z} e^{-z^2/2}\, dz \tag{2.31}$$

where $\Delta z = \Delta x/\sigma$ measures the deviation from the mean in units of the standard deviation σ.

The integral of Equation (2.31) cannot be evaluated analytically, so in order to obtain the probability $P_G(\Delta x; \mu, \sigma)$ it is necessary either to expand the Gaussian function in a Taylor's series and integrate the series term by term, or to integrate numerically. With modern computers, numerical integration is fast and accurate, and reliable results can be obtained from a simple quadratic integration (Appendix A.3).

Tables and Graphs

The Gaussian probability density function $p_G(z)$ and the integral probability $P_G(z)$ are tabulated and plotted in Tables C.1 and C.2, respectively. From the integral probability Table C.2, we note that the probabilities are about 68% and 95% that a given measurement will fall within 1 and 2 standard deviations of the mean, respectively. Similarly, by considering the 50% probability limit we can see that the probable error is given by $\sigma_{pe} = 0.6745\sigma$.

Comparison of Gaussian and Poisson Distributions

A comparison of the Poisson and Gaussian curves reveals the nature of the Poisson distribution. It is the appropriate distribution for describing experiments in which the possible values of the data are strictly bounded on one side but not on the other. The Poisson curve of Figure 2.3 exhibits the typical Poisson shape. The Poisson curve of Figure 2.4 differs little from the corresponding Gaussian curve of Figure 2.5, indicating that for large values of the mean μ, the Gaussian distribution becomes an acceptable description of the Poisson distribution. Because, in general, the Gaussian distribution is more convenient to calculate than the Poisson distribution, it is often the preferred choice. However, one should remember that the Poisson distribution is only defined at 0 and positive integral values of the variable x, whereas the Gaussian function is defined at all values of x.

2.4 LORENTZIAN DISTRIBUTION

There are many other distributions that appear in scientific research. Some are phenomenological distributions, created to parameterize certain data distributions. Others are well grounded in theory. One such distribution in the latter category is the Lorentzian distribution, similar but unrelated to the binomial distribution. The Lorentzian distribution is an appropriate distribution for describing data corresponding to resonant behavior, such as the variation with energy of the cross section of a nuclear or particle reaction or absorption of radiation in the Mössbauer effect.

The *Lorentzian probability density* function $P_L(x; \mu, \Gamma)$, also called the *Cauchy distribution*, is defined as

$$p_L(x; \mu, \Gamma) = \frac{1}{\pi} \frac{\Gamma/2}{(x-\mu)^2 + (\Gamma/2)^2} \tag{2.32}$$

This distribution is symmetric about its mean μ with a width characterized by its half-width Γ. The most striking difference between it and the Gaussian distribution is that it does not diminish to 0 as rapidly; the behavior for large deviations is proportional to the inverse square of the deviation, rather than exponentially related to the square of the deviation.

As with the Gaussian distribution, the Lorentzian distribution function is a continuous function, and the probability of observing a value x must be related to the interval within which the observation may fall. The probability $dP_L(x; \mu, \Gamma)$ for an observation to fall within an infinitesimal differential interval dx around x is given by the product of the probability density function $p_L(x; \mu, \Gamma)$ and the size of the interval dx:

$$dP_L(x; \mu, \Gamma) = p_L(x; \mu, \Gamma)\, dx \tag{2.33}$$

The normalization of the probability density function $p_L(x; \mu, \Gamma)$ is such that the integral of the probability over all possible values of x is unity:

$$\int_{-\infty}^{\infty} p_L(x;\mu,\Gamma)\,dx = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{1+z^2}\,dz = 1 \tag{2.34}$$

where $z = (x-\mu)/(\Gamma/2)$.

Mean and Half-Width

The mean μ of the Lorentzian distribution is given as one of the parameters in Equation (2.32). It is obvious from the symmetry of the distribution that μ must be equal to the mean as well as to the median and to the most probable value.

The standard deviation is not defined for the Lorentzian distribution as a consequence of its slowly decreasing behavior for large deviations. If we attempt to evaluate the expectation value for the square of the deviations

$$\sigma^2 = \langle (x-\mu)^2 \rangle = \frac{1}{\pi}\frac{\Gamma^2}{4}\int_{-\infty}^{\infty}\frac{z^2}{1+z^2}\,dz \tag{2.35}$$

we find that the integral is unbounded: the integral does not converge for large deviations. Although it is possible to calculate a *sample standard deviation* by evaluating the average value of the square of the deviations from the sample mean, this calculation has no meaning and will not converge to a fixed value as the number of samples increases.

The width of the Lorentzian distribution is instead characterized by the *full-width at half maximum* Γ, generally called the *half-width*. This parameter is defined such that when $x = \mu \pm \Gamma/2$, the probability density function is equal to one-half its maximum value, or $p(\mu \pm \Gamma/2;\mu,\Gamma) = \frac{1}{2}p(\mu;\mu,\Gamma)$. Thus, the half-width Γ is the full width of the curve measured between the levels of half maximum probability. We can verify that this identification of Γ with the full-width at half maximum is correct by substituting $x = \mu \pm \Gamma/2$ into Equation (2.32).

The Lorentzian and Gaussian distributions are shown for comparison in Figure 2.6, for $\mu = 10$ and $\Gamma = 2.354$ (corresponding to $\sigma = 1$ for the Gaussian function). Both distributions are normalized to unit area according to their definitions in Equations (2.23) and (2.32). For both curves, the value of the maximum probability is inversely proportional to the half-width Γ. This results in a peak value of $2/\pi\Gamma \simeq 0.270$ for the Lorentzian distribution and a peak value of $1/\sigma\sqrt{2\pi} \simeq 0.399$ for the Gaussian distribution.

Except for the normalization, the Lorentzian distribution is equivalent to the dispersion relation that is used, for example, in describing the cross section of a nuclear reaction for a Breit-Wigner resonance:

$$\sigma = \pi\lambda^2\frac{\Gamma_1\Gamma_2}{(E-E_0)^2+(\Gamma/2)^2} \tag{2.36}$$

SUMMARY

Binomial distribution: Describes the probability of observing x successes out of n tries when the probability for success in each try is p:

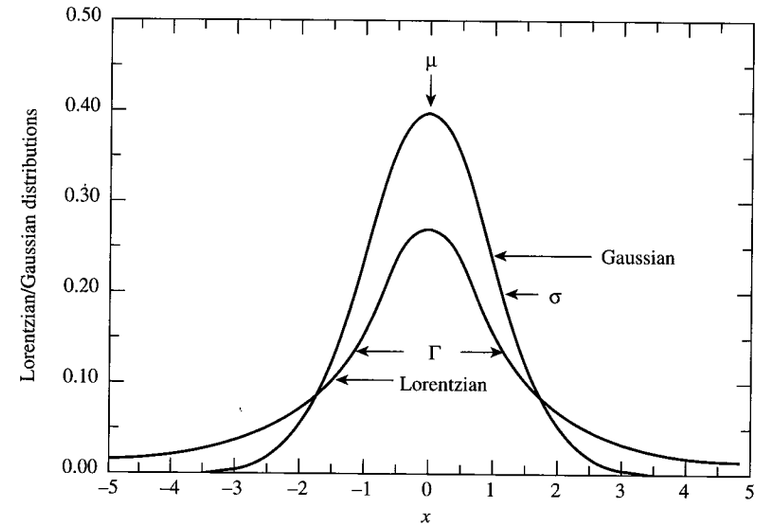


FIGURE 2.6
Comparison of normalized Lorentzian and Gaussian distributions, with $\Gamma = 2.354\sigma$.

$$p_B(x;n,p) = \binom{n}{x}p^x q^{n-x} = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$

$$\mu = np \qquad \sigma^2 = np(1-p)$$

Poisson distribution: Limiting case of the binomial distribution for large n and constant μ; appropriate for describing small samples from large populations.

$$P_P(x;\mu) = \frac{\mu^x}{x!}e^{-\mu}, \qquad \sigma^2 = \mu$$

Gaussian distribution: Limiting case of the binomial distribution for large n and finite p; appropriate for smooth symmetric distributions.

$$p_G(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Half-width $\Gamma = 2.354\sigma$; probable error P.E. $= 0.6745\sigma$.
Standard form:

$$p_G(z)\,dz = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z}{2}\right)^2 dz$$

Lorentzian distribution: Describes resonant behavior

$$p_L(x;\mu,\Gamma) = \frac{1}{\pi}\frac{\Gamma/2}{(x-\mu)^2+(\Gamma/2)^2}$$

EXERCISES

(2.1) Consider five coins labeled a, b, c, d, and e. Let x = *number of heads showing*.
(a) Manually count and tabulate *all* possible permutations for each of the following configurations:
i. $x = 0$
ii. $x = 1$
iii. $x = 2$
iv. $x = 3$
v. $x = 4$
vi. $x = 5$
Compare your results to those given by Equation (2.2).
(b) Manually delete all duplicate permutations from each example of part (a), that is, cross out permutations that repeat a previous combination in a different order. Compare your results to those given by Equation (2.3).

2.2. Evaluate the following:

(a) $\binom{6}{3}$ (b) $\binom{4}{2}$ (c) $\binom{10}{3}$ (d) $\binom{52}{4}$

2.3. Evaluate the binomial distribution $P_B(x; n, p)$ for $n = 6$, $p = \frac{1}{2}$, and $x = 0$ to 6. Sketch the distribution and identify the mean and standard deviation. Repeat for $p = \frac{1}{6}$.

(2.4) The probability distribution of the sum of the points showing on a pair of dice is given by

$$P(x) = \frac{x-1}{36} \qquad 2 \le x \le 7$$
$$= \frac{13-x}{36} \qquad 7 \le x \le 12$$

Find the mean, median, and standard deviation of the distribution.

2.5. Show that the sum in Equation (2.6) reduces to $\mu = np$. *Hint*: Define $y = x - 1$ and $m = n - 1$ and use the fact that

$$\sum_{y=0}^{m}\left[\frac{m!}{y!(m-y)!}p^y(1-p)^{m-y}\right] = \sum_{y=0}^{m} P_B(y; m, p) = 1$$

2.6. On a certain kind of slot machine there are 10 different symbols that can appear in each of three windows. The machine pays off different amounts when either one, two, or three lemons appear. What should be the payoff ratio for each of the three possibilities if the machine is honest and there is no cut for the house?

(2.7) Show that the sum in Equation (2.7) reduces to $\sigma^2 = np(1-p)$. *Hint*: Define $y = x - 1$ and $m = n - 1$ and use the results of Exercise 2.5.

2.8. At rush hour on a typical day, 25.0% of the cars approaching a fork in the street turn left and 75.0% turn right. On a particular day, 283 cars turned left and 752 turned right. Find the predicted uncertainty in these numbers and the probability that these measurements were not made on a "typical day"; that is, find the probability of obtaining a result that is as far or farther from the mean than the result measured on the particular day.

(2.9) In a certain physics course, 7.3% of the students failed and 92.7% passed, averaged over many semesters.
(a) What is the expected number of failures in a particular class of 32 students, drawn from the same population?
(b) What is the probability that five or more students will fail?

2.10. Evaluate and plot the two Poisson distributions of Example 2.4. Plot on each graph the corresponding Gaussian distribution with the same mean and standard deviation.

2.11. Verify that, for the Poisson distribution, if μ is an integer, the probability for $x = \mu$ is equal to the probability for $x = \mu - 1$, $P_P(\mu, \mu) = P_P(\mu - 1; \mu)$.

2.12. Show that the sum in Equation (2.19) reduces to $\sigma^2 = \mu$. *Hint*: Use Equation (2.18) to simplify the expression. Define $y = x - 1$ and show that the sum reduces to $\mu\langle y + 1\rangle = \mu^2$.

2.13. Members of a large collaboration that operated a giant proton-decay detector in a salt mine near Cleveland, Ohio, detected a burst of 8 neutrinos in their apparatus coincident with the optical observation of the explosion of the Supernova 1987A.
(a) If the average number of neutrinos detected in the apparatus is 2 per day, what is the probability of detecting a fluctuation of 8 or more in one day?
(b) In fact, the 8 neutrinos were all detected within a 10-min period. What is the probability of detecting a fluctuation of 8 or more neutrinos in a 10-min period if the average rate is 2 per 24 hours?

2.14. In a scattering experiment to measure the polarization of an elementary particle, a total of $N = 1000$ particles was scattered from a target. Of these, 670 were observed to be scattered to the right and 330 to the left. Assume that there is no uncertainty in $N = N_R + N_L$.
(a) Based on the experimental estimate of the probability, what is the uncertainty in N_R? In N_L?
(b) The asymmetry parameter is defined as $A = (N_R - N_L)/(N_R + N_L)$. Calculate the experimental asymmetry and its uncertainty.
(c) Assume that the asymmetry has been predicted to be $A = 0.400$ and recalculate the uncertainties in (a) and (b) using the predicted probability.

2.15. A problem arises when recording data with electronic counters in that the system may saturate when rates are very high, leading to a "dead time." For example, after a particle has passed through a detector, the equipment will be "dead" while the detector recovers and the electronics stores away the results. If a second particle passes through the detector in this time period, it will not be counted.
(a) Assume that a counter has a dead time of 200 ns (200×10^{-9} s) and is exposed to a beam of 1×10^6 particles per second so that the mean number of particles hitting the counter in the 200-ns time slot is $\mu = 0.2$. From the Poisson probability for this process, find the efficiency of the counter, that is, the ratio of the average number of particles counted to the average number that pass through the counter in the 200-ns time period.
(b) Repeat the calculation for beam rates of 2, 4, 6, 8, and 10×10^6 particles per second, and plot a graph of counter efficiency as a function of beam rate.

2.16. Show by numerical calculation that, for the Gaussian probability distribution, the full-width at half maximum Γ is related to the standard deviation by $\Gamma = 2.354\sigma$ [Equation (2.28)].

(2.17) The probability that an electron is at a distance r from the center of the nucleus of a hydrogen atom is given by

$$dP(r) = Cr^2e^{-r/R}dr$$

Find the mean radius $\bar{r}$ and the standard deviation. Find the value of the constant C.

(2.18) Show that a tangent to the Gaussian function is steepest at $x = \mu \pm \sigma$, and therefore intersects the curve at the $e^{-1/2}$ points. Show also that these tangents intersect the x axis at $x = \mu \pm 2\sigma$.

CHAPTER 3

ERROR ANALYSIS

In Chapter 1 we discussed methods for extracting from a set of data points estimates of the mean and standard deviation that describe, respectively, the desired result and the uncertainties in the results. In this chapter we shall further consider how to estimate uncertainties in our measurements, the sources of the uncertainties, and how to combine uncertainties in separate measurements to find the error in a result calculated from those measurements.

3.1 INSTRUMENTAL AND STATISTICAL UNCERTAINTIES

Instrumental Uncertainties

If the quantity x has been measured with a physical instrument, the uncertainty in the measurement generally comes from fluctuations in readings of the instrumental scale, either because the settings are not exactly reproducible due to imperfections in the equipment, or because of human imprecision in observing settings, or a combination of both. Such uncertainties are called *instrumental* because they arise from a lack of perfect precision in the measuring instruments (including the observer). We can include in this category experiments that deal with measurements of such characteristics as length, mass, voltage, current, and so forth. These uncertainties are often independent of the actual value of the quantity being measured.

Instrumental uncertainties are generally determined by examining the instruments and considering the measuring procedure to estimate the reliability of the measurements. In general, one should attempt to make readings to a fraction of the smallest scale division on the instrument. For example, with a good mercury thermometer, it is often easy to estimate the level of the mercury to a least count of one-half of the

smallest scale division and possibly even to one-fifth of a division. The measurement is generally quoted to plus or minus one-half of the least count, and this number represents an estimate of the standard deviation of a single measurement. Recalling that, for a Gaussian distribution, there is a 68% probability that a random measurement will lie within 1 standard deviation of the mean, we observe that our object in estimating errors is not to place outer limits on the range of the measurement, which is impossible, but to set a particular *confidence level* that a repeated measurement of the quantity will fall this close to the mean or closer. Often we choose the standard deviation, the 68% confidence level, but other levels are used as well. We shall discuss the concept of confidence levels in Chapter 11.

Digital instruments require special consideration. Generally, manufacturers specify a *tolerance;* for example, the tolerance of a digital multimeter may be given as $\pm 1\%$. At any rate, the precision cannot be better than half the last digit on the display. The manufacturer's quoted tolerances may require interpretation as to whether the uncertainty must be treated as a systematic effect or a statistical effect. For example, if a student uses a resistor with a stated 1% tolerance in an experiment, he can expect the stated uncertainty in the resistance to make a systematic contribution to all experiments with that resistor. On the other hand, when he combines his results with those of the other students in the class, each of whom used a different resistor, the uncertainties in the individual resistances contribute in a statistical manner to the variation of the combined sample.

If it is possible to make repeated measurements, then an estimate of the standard deviation can be calculated from the spread of these measurements as discussed in Chapter 1. The resulting estimate of the standard deviation corresponds to the expected uncertainty in a single measurement. In principle, this *internal* method of determining the uncertainty should agree with that obtained by the *external* method of considering the equipment and the experiment itself, and in fact, any significant discrepancy between the two suggests a problem, such as a misunderstanding of some aspect of the experimental procedure. However, when reasonable agreement is achieved, then the standard deviation calculated internally from the data generally provides the better estimate of the uncertainties.

Statistical Uncertainties

If the measured quantity x represents the number of counts in a detector per unit time interval for a random process, then the uncertainties are called *statistical* because they arise not from a lack of precision in the measuring instruments but from overall statistical fluctuations in the collections of finite numbers of counts over finite intervals of time. For statistical fluctuations, we can estimate analytically the standard deviation for each observation, without having to determine it experimentally. If we were to make the same measurement repeatedly, we should find that the observed values were distributed about their mean in a Poisson distribution (as discussed in Section 2.2) instead of a Gaussian distribution. We can justify the use of this distribution intuitively by considering that we should expect a distribution that is related to the binomial distribution, but that is consistent with our boundary conditions that we can collect any positive number of counts, but no fewer than zero counts, in any time interval.

The Poisson distribution and statistical uncertainties do not apply solely to experiment where counts are recorded in unit time intervals. In any experiment in which data are grouped in bins according to some criterion to form a histogram or frequency plot, the number of events in each individual bin will obey Poisson statistics and fluctuate with statistical uncertainties.

One immediate advantage of the Poisson distribution is that the standard deviation is automatically determined:

$$\sigma = \sqrt{\mu} \tag{3.1}$$

The relative uncertainty, the ratio of the standard deviation to the average rate, $\sigma/\mu = 1/\sqrt{\mu}$, decreases as the number of counts received per interval increases. Thus relative uncertainties are smaller when counting rates are higher.

The value for μ to be used in Equation (3.1) for determining the standard deviation σ is, of course, the value of the mean counting rate from the parent population, of which each measurement x is only an approximate sample. In the limit of an infinite number of determinations, the average of all the measurements would very closely approximate the parent value, but often we cannot make more than one measurement of each value of x, much less an infinite number. Thus, we are forced to use $\sqrt{x}$ as an estimate of the standard deviation of a single measurement.

Example 3.1. Consider an experiment in which we count gamma rays emitted by a strong radioactive source. We cannot determine the counting rate instantaneously because no counts will be detected in an infinitesimal time interval. But we can determine the number of counts x detected over a time interval Δt, and this should be representative of the average counting rate over that interval. Assume that we have recorded 5212 counts in a 1-s time interval. The distribution of counts is random in time and follows the Poisson probability function, so our estimate of the standard deviation of the distribution is $\sigma = \sqrt{5212}$. Thus, we should record our result for the number of counts x in the time interval Δt as 5212 ± 72 and the relative error is

$$\frac{\sigma_x}{x} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}} \simeq \frac{1}{72} = 0.014 = 1.4\%$$

There may also be instrumental uncertainties contributing to the overall uncertainties. For example, we can determine the time intervals with only finite precision. However, we may have some control over these uncertainties and can often organize our experiment so that the statistical errors are dominant. Suppose that the major instrumental error in our example is the uncertainty $\sigma_t = 0.01$ s in the time interval $\Delta t = 1.00$ s. The relative uncertainty in the time interval is thus

$$\frac{\sigma_t}{\Delta t} = \frac{0.01}{1.00} = 0.01 = 1.\%$$

This relative instrumental error in the time interval will produce a 1.% relative error in the number of counts x. Because the instrumental uncertainty is comparable to the statistical uncertainty, it might be wise to attempt a more precise measurement of the interval or to increase its length. If we increase the counting time interval from 1 s to 4 s, the number of counts x will increase by about a factor of 4 and the relative statistical error will therefore decrease by a factor of 2 to about 0.7%, whereas the relative instrumental uncertainty will decrease by a factor of 4 to 0.25%, as long as the instrumental uncertainty σ_t remains constant at 0.01 s.

3.2 PROPAGATION OF ERRORS

We often want to determine a dependent variable x that is a function of one or more different measured variables. We must know how to propagate or carry over the uncertainties in the measured variables to determine the uncertainty in the dependent variable.

Example 3.2. Suppose we wish to find the volume V of a box of length L, width W, and height H. We can measure each of the three dimensions to be L_0, width W_0, and height H_0 and combine these measurements to yield a value for the volume:

$$V_0 = L_0 W_0 H_0 \tag{3.2}$$

How do the uncertainties in the estimates L_0, W_0, and H_0, affect the resulting uncertainties in the final result V_0?

If we knew the actual errors, $\Delta L = L - L_0$ and so forth, in each dimension, we could obtain an estimate of the error in the final result V_0 by expanding V about the point (L_0, W_0, H_0) in a Taylor series. The first term in the Taylor expansion gives

$$V \simeq V_0 + \Delta L\left(\frac{\partial V}{\partial L}\right)_{W_0 H_0} + \Delta W\left(\frac{\partial V}{\partial W}\right)_{L_0 H_0} + \Delta H\left(\frac{\partial V}{\partial H}\right)_{L_0 W_0} \tag{3.3}$$

from which we can find $\Delta V = V - V_0$. The terms in parentheses are the partial derivatives of V, with respect to each of the dimensions, L, W, and H, evaluated at the point L_0, W_0, H_0. They are the proportionality constants between changes in V and infinitesimally small changes in the corresponding dimensions. The partial derivative of V with respect to L, for example, is evaluated with the other variables W and H held fixed at the values W_0 and H_0 as indicated by the subscript. This approximation neglects higher-order terms in the Taylor expansion, which is equivalent to neglecting the fact that the partial derivatives are not constant over the ranges of L, W, and H given by their errors. If the errors are large, we must include in this definition at least second partial derivatives ($\partial V^2/\partial L^2$, etc.) and partial cross derivatives ($\partial^2 V/\partial L\, \partial W$, etc.), but we shall omit these from the discussion that follows.

For our example of $V = LWH$, Equation (3.3) gives

$$\Delta V \simeq W_0 H_0 \Delta L + L_0 H_0 \Delta W + L_0 W_0 \Delta H \tag{3.4}$$

which we could evaluate if we knew the uncertainties ΔL, ΔW, and ΔH.

Uncertainties

In general, however, we do not know the actual errors in the determination of the dependent variables (or if we do, we should make the necessary corrections). Instead, we may be able to estimate the error in each measured quantity, or to estimate some characteristic, such as the standard deviation σ, of the probability distribution

of the measured qualities, How can we combine the standard deviation of the individual measurements to estimate the uncertainty in the result?

Suppose we want to determine a quantity x that is a function of at least two measured variables, u and v. We shall determine the characteristics of x from those of u and v and from the fundamental dependence

$$x = f(u, v, \dots) \tag{3.5}$$

Although it may not always be exact, we shall assume that the most probable value for x is given by

$$\bar{x} = f(\bar{u}, \bar{v}, \dots) \tag{3.6}$$

The uncertainty in the resulting value for x can be found by considering the spread of the values of x resulting from combining the individual measurements u_i, v_i, ... into individual results x_i:

$$x_i = f(u_i, v_i, \dots) \tag{3.7}$$

In the limit of an infinite number of measurements, the mean of the distribution will coincide with the average $\bar{x}$ given in Equation (3.6) and we can use the definition of Equation (1.8) to find the variance σ_x^2 (which is the square of the standard deviation σ_x):

$$\sigma_x^2 = \lim_{N\to\infty}\left[\frac{1}{N}\sum (x_i - \bar{x})^2\right] \tag{3.8}$$

Just as we expressed the deviation of V in Equation (3.4) as a function of the deviations in the dimensions L, W, and H, so we can express the deviations $x_i - \bar{x}$ in terms of the deviations $u_i - \bar{u}$, $v_i - \bar{v}$, ... of the observed parameters

$$x_i - \bar{x} \simeq (u_i - \bar{u})\left(\frac{\partial x}{\partial u}\right) + (v - \bar{v})\left(\frac{\partial x}{\partial v}\right) + \cdots \tag{3.9}$$

where we have omitted specific notation of the fact that each of the partial derivatives is evaluated with all the other variables fixed at their mean values.

Variance and Covariance

Combining Equations (3.8) and (3.9) we can express the variance σ_x^2 for x in terms of the variances σ_u^2, σ_v^2, ... for the variables u, v, ..., which were actually measured:

$$\begin{aligned}\sigma_x^2 &\simeq \lim_{N\to\infty}\frac{1}{N}\sum\left[(u_i - \bar{u})\left(\frac{\partial x}{\partial u}\right) + (v_i - \bar{v})\left(\frac{\partial x}{\partial v}\right) + \cdots\right]^2 \\ &\simeq \lim_{N\to\infty}\frac{1}{N}\sum\left[(u_i - \bar{u})^2\left(\frac{\partial x}{\partial u}\right)^2 + (v_i - \bar{v})^2\left(\frac{\partial x}{\partial v}\right)^2\right. \\ &\qquad \left. + 2(u_i - \bar{u})(v_i - \bar{v})\left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial x}{\partial v}\right) + \cdots\right]\end{aligned} \tag{3.10}$$

The first two terms of Equation (3.10) can be expressed in terms of the variances σ_u^2 and σ_v^2 given by Equation (1.8):

$$\sigma_u^2 = \lim_{N\to\infty}\left[\frac{1}{N}\sum (u_i - \bar{u}_i)^2\right] \qquad \sigma_v^2 = \lim_{N\to\infty}\left[\frac{1}{N}\sum (v_i - \bar{v}_i)^2\right] \tag{3.11}$$

In order to express the third term of Equation (3.10) in a similar form, we introduce the *covariances* σ_{uv}^2 between the variables u and v defined analogous to the variances of Equation (3.11):

$$\sigma_{uv}^2 \equiv \lim_{N\to\infty}\left[\frac{1}{N}\sum [(u_i - \bar{u})(v_i - \bar{v})]\right] \tag{3.12}$$

With these definitions, the approximation for the variance σ_x^2 for x given in Equation (3.10) becomes

$$\sigma_x^2 \simeq \sigma_u^2\left(\frac{\partial x}{\partial u}\right)^2 + \sigma_v^2\left(\frac{\partial x}{\partial v}\right)^2 + \cdots + 2\sigma_{uv}^2\left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial x}{\partial v}\right) + \cdots \tag{3.13}$$

Equation (3.13) is known as the *error propagation equation.*

The first two terms in the equation are averages of squares of deviations weighted by the squares of the partial derivatives, and may be considered to be the averages of the squares of the deviations in x produced by the uncertainties in u and in v, respectively. In general, these terms dominate the uncertainties. If there are additional variables besides u and v in the determination of x, their contributions to the variance of x will have similar terms.

The third term is the average of the cross terms involving products of deviations in u and v weighted by the product of the partial derivatives. If the fluctuations in the measured quantities u and v, ... are uncorrelated, then, on the average, we should expect to find equal distributions of positive and negative values for this term, and we should expect the term to vanish in the limit of a large random selection of observations. This is often a reasonable approximation and Equation (3.13) then reduces to

$$\sigma_x^2 \simeq \sigma_u^2\left(\frac{\partial x}{\partial u}\right)^2 + \sigma_v^2\left(\frac{\partial x}{\partial v}\right)^2 + \cdots \tag{3.14}$$

with similar terms for additional variables. In general, we use Equation (3.14) for determining the effects of measuring uncertainties on the final result and neglect the covariant terms. However, as we shall see in Chapter 7, the covariant terms often make important contributions to the uncertainties in parameters determined by fitting curves to data by the least-squares method.

3.3 SPECIFIC ERROR FORMULAS

The expressions of Equations (3.13) and (3.14) were derived for the general relationship of Equation (3.5) giving x as an arbitrary function of u and v, ... In the following specific cases of functions $f(u, v, \dots)$, the parameters a and b are defined as constants and u and v are variables.

Simple Sums and Differences

If the dependent variable x is related to a measured quantity u by the relation

$$x = u + a \tag{3.15}$$

then the partial derivative $\partial x/\partial u = 1$ and the uncertainty in x is just

$$\sigma_x = \sigma_u \tag{3.16}$$

and the relative uncertainty is given by

$$\frac{\sigma_x}{x} = \frac{\sigma_u}{x} = \frac{\sigma_u}{u + a} \tag{3.17}$$

Note that if we are dealing with a small difference between u and a, the uncertainty in x might be greater than the magnitude of x, even for a small relative uncertainty in u.

Example 3.3. In an experiment to count particles emitted by a decaying radioactive source, we measure $N_1 = 723$ counts in a 15-s time interval at the beginning of the experiment and $N_2 = 19$ counts in a 15-s time interval later in the experiment. The events are random and obey Poisson statistics so that we know that the uncertainties in N_1 and N_2 are just their square roots. Assume that we have made a very careful measurement of the background counting rate in the absence of the radioactive source and obtained a value $B = 14.2$ counts with negligible error for the same time interval Δt. Because we have averaged over a long time period, the mean number of background counts in the 15-s interval is not an integral number.

For the first time interval, the corrected number of counts is

$$x_1 = N_1 - B = 723 - 14.2 = 708.8 \text{ counts}$$

The uncertainty in x_1 is given by

$$\sigma_{x_1} = \sigma_{N_1} = \sqrt{723} \simeq 26.9 \text{ counts}$$

and the relative uncertainty is

$$\frac{\sigma_x}{x} = \frac{26.9}{708} = 0.038 \simeq 3.8\%$$

For the second time interval, the corrected number of events is

$$x_2 = N_2 - B = 19 - 14.2 \simeq 4.8 \text{ counts}$$

The uncertainty in x is given by

$$\sigma_{x_2} = \sigma_{N_2} = \sqrt{19} \simeq 4.4 \text{ counts}$$

and the relative uncertainty in x is

$$\frac{\sigma_x}{x} \simeq \frac{4.4}{4.8} = 0.91$$

Weighted Sums and Differences

If x is the weighted sum of u and v,

$$x = au + bv \tag{3.18}$$

the partial derivatives are simply the constants

$$\left(\frac{\partial x}{\partial u}\right) = a \qquad \left(\frac{\partial x}{\partial v}\right) = b \tag{3.19}$$

and we obtain

$$\sigma_x^2 = a^2\sigma_u^2 + b^2\sigma_v^2 + 2ab\sigma_{uv}^2 \tag{3.20}$$

Note the possibility that the variance σ_x^2 might vanish if the covariance σ_{uv}^2 has the proper magnitude and sign. This could happen in the unlikely event that the fluctuations were completely correlated so that each erroneous observation of u was exactly compensated for by a corresponding erroneous observation of v.

Example 3.4. Suppose that, in the previous example, the background counts B were not averaged over a long time period but were simply measured for 15 s to give $B = 14$ with standard deviation $\sigma_B = \sqrt{14} \simeq 3.7$ counts. Then the uncertainty in x would be given by

$$\sigma_x^2 = \sigma_N^2 + (-\sigma_B)^2 = N + B$$

because the uncertainties in N and B are equal to their square roots.

For the first time interval, we would calculate

$$x_1 = (723 - 14) \pm \sqrt{723 + 14} = 709 \pm 27.1 \text{ counts}$$

and the relative uncertainty would be

$$\frac{\sigma_x}{x} = \frac{27.1}{709} \simeq 0.038$$

For the second time interval, we would calculate

$$x_2 = (19 - 14) \pm \sqrt{19 + 14} = 5 \pm 5.7 \text{ counts}$$

and the relative uncertainty would be

$$\frac{\sigma_x}{x} = \frac{5.7}{5} \simeq 1.1$$

Multiplication and Division

If x is the weighted product of u and v,

$$x = auv \tag{3.21}$$

the partial derivatives of each variable are functions of the other variable,

$$\left(\frac{\partial x}{\partial u}\right) = av \qquad \left(\frac{\partial x}{\partial v}\right) = au \tag{3.22}$$

and the variance of x becomes

$$\sigma_x^2 = (av\sigma_u)^2 + (au\sigma_v)^2 + 2a^2uv\sigma_{uv}^2 \tag{3.23}$$

which can be expressed more symmetrically as

$$\frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} + 2\frac{\sigma_{uv}^2}{uv} \tag{3.24}$$

Similarly, if x is obtained through division,

$$x = \frac{au}{v} \tag{3.25}$$

the relative variance for x is given by

$$\frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} - 2\frac{\sigma_{uv}^2}{uv} \tag{3.26}$$

Example 3.5. The area of a triangle is equal to half the product of the base times the height $A = bh/2$. If the base and height have values $b = 5.0 \pm 0.1$ cm and $h = 10.0 \pm 0.3$ cm, the area is $A = 25.0$ cm^2 and the relative uncertainty in the area is given by

$$\frac{\sigma_A^2}{A^2} = \frac{\sigma_b^2}{b^2} + \frac{\sigma_h^2}{h^2} \tag{3.27}$$

or

$$\begin{aligned}\sigma_A^2 &= A^2\left(\frac{\sigma_b^2}{b^2} + \frac{\sigma_h^2}{h^2}\right)\\ &= 25^2(\text{cm})^4\left(\frac{0.1^2}{5^2} + \frac{0.3^2}{10^2}\right)(\text{cm}^2/\text{cm}^2)\\ &\simeq 0.81\ \text{cm}^4\end{aligned}$$

Although the absolute uncertainty in the height is 3 times the absolute uncertainty in the base, the relative uncertainty is only 1½ times as large and its contribution to the variance of the area is only $(1½)^2$ as large.

Powers

If x is obtained by raising the variable u to a power

$$x = au^b \tag{3.28}$$

the derivative of x with respect to u is

$$\left(\frac{\partial x}{\partial u}\right) = abu^{b-1} = \frac{bx}{u} \tag{3.29}$$

and relative error in x becomes

$$\frac{\sigma_x}{x} = b\frac{\sigma_u}{u} \tag{3.30}$$

For the special cases of $b = +1$, we have

$$x = au \qquad \sigma_x = a\sigma_\mu$$

so

$$\frac{\sigma_x}{x} = \frac{\sigma_u}{u} \tag{3.31}$$

For $b = -1$, we have

$$x = \frac{a}{u} \qquad \sigma_x = -\frac{a\sigma_u}{u^2}$$

so

$$\frac{\sigma_x}{x} = -\frac{\sigma_u}{u} \tag{3.32}$$

The negative sign indicates that, in division, a positive error in u will produce a corresponding negative error in x.

Example 3.6. The area of a circle is proportional to the square of the radius $A = \pi r^2$. If the radius is determined to be $r = 10.0 \pm 0.3$ cm, the area is $A = 100.\pi$ cm^2 with an uncertainty given by

$$\frac{\sigma_A}{A} = 2\,\frac{\sigma_r}{r}$$

or

$$\sigma_A = 2A\,\frac{\sigma_r}{r} = 2\pi(10.0\ \text{cm})^2(0.3\ \text{cm})/(10.0\ \text{cm}) = 6\pi\ \text{cm}^2$$

Exponentials

If x is obtained by raising the natural base to a power proportional to u,

$$x = ae^{bu} \tag{3.33}$$

the derivative of x with respect to u is

$$\frac{\partial x}{\partial u} = abe^{bu} = bx \tag{3.34}$$

and the relative uncertainty becomes

$$\frac{\sigma_x}{x} = b\sigma_u \tag{3.35}$$

If the constant that is raised to the power is not equal to e, the expression can be rewritten as

$$\begin{aligned}x &= a^{bu} \\ &= (e^{\ln a})^{bu} = e^{(b\ln a)u}\\ &= e^{cu} \text{ with } c = b\ln a\end{aligned} \tag{3.36}$$

where ln indicates the natural logarithm. Solving in the same manner as before we obtain

$$\frac{\sigma_x}{x} = c\sigma_u = (b\ln a)\sigma_u \tag{3.37}$$

Logarithms

If x is obtained by taking the logarithm of u,

$$x = a \ln(bu) \tag{3.38}$$

the derivative with respect to u is

$$\frac{\partial x}{\partial u} = \frac{ab}{u} \tag{3.39}$$

$$\sigma_x = ab\frac{\sigma_u}{u} \tag{3.40}$$

Angle Functions

If x is determined as a function of u, such as

$$x = a \cos(bu) \tag{3.41}$$

The derivative of x with respect to u is

$$\frac{dx}{du} = -ab \sin(bu) \tag{3.42}$$

so

$$\sigma_x = -\sigma_u ab \sin(bu) \tag{3.43}$$

Similarly, if

$$x = a \sin(bu) \tag{3.44}$$

then

$$\frac{dx}{du} = ab \cos(bu) \tag{3.45}$$

so

$$\sigma_x = \sigma_u ab \cos(bu) \tag{3.46}$$

Note that σ_u is the uncertainty in an angle and therefore must be expressed in radians.

These relations can be useful for making quick estimates of the uncertainty in a calculated quantity caused by the uncertainty in a measured variable. For a simple product or quotient of the measured variable u with a constant, a 1% error in u causes a 1% error in x. If u is raised to a power b, the resulting error in x becomes b% for a 1% uncertainty in u. Even if the complete expression for x involves other measured variables, $x = f(u, v, \ldots)$ and is considerably more complicated than these simple examples, it is often possible to use these relations to make approximate estimates of uncertainties.

3.4 APPLICATION OF ERROR EQUATIONS

Even for relatively simple calculations, such as those encountered in undergraduate laboratory experiments, blind application of the general error propagation expression

[Equation (3.14)] can lead to very lengthy and discouraging equations, especially if the final results depend on several different measured quantities. Often the error equations can be simplified by neglecting terms that make negligible contributions to the final uncertainty, but this requires a certain amount of practice.

Approximations

Students should practice making quick, approximate estimates of the various contributions to the uncertainty in the final result by considering separately the terms in Equation (3.14). A convenient rule of thumb is to neglect terms that make final contributions that are less than 10% of the largest contribution. (Like all rules of this sort, one should be wary of special cases. Several smaller contributions to the final uncertainty can sum to be as important as one larger uncertainty.)

Example 3.7. Suppose that the area of a rectangle $A = LW$ is to be determined from the following measurements of the lengths of two sides:

$$L = 22.1 \pm 0.1 \text{ cm} \qquad W = 7.3 \pm 0.1 \text{ cm}$$

The relative contribution of σ_L to the error in L will be

$$\frac{\sigma_{A_L}}{A} = \frac{\sigma_L}{L} = \frac{0.1}{22.1} = 0.005$$

and the corresponding contribution of σ_W will be

$$\frac{\sigma_{A_W}}{A} = \frac{\sigma_W}{W} = \frac{0.1}{7.3} \simeq 0.014$$

The contribution from σ_L is thus about one-third of that from σ_W. However, when the contributions are combined, we obtain

$$\sigma_A = A\sqrt{0.014^2 + 0.005^2}$$

which can be expanded to give

$$\sigma_A \simeq 0.014A\left(1 + \frac{1}{2}\left(\frac{0.005}{0.014}\right)^2\right) \simeq 0.014A(1 + 0.06) = 0.015A$$

Thus, the effective contribution from σ_L is only about 6% of the effective contribution from σ_W and could safely be neglected in this calculation.

Computer Calculation of Uncertainties

Finding analytic forms for the partial derivatives is sometimes quite difficult. One should always break Equation (3.14) into separate components and not attempt to find one complete equation that incorporates all error terms. In fact, if the analysis is being done by computer, it may not even be necessary to find the derivatives explicitly. The computer can find numerically the variations in the dependent variable caused by variations in each independent, or measured, variable.

Suppose that we have a particularly complicated equation, or set of equations, relating our final result x to the individually measured variables u, v, and so forth. Let us assume that the actual equations are programmed as a computer function CALCULATE, which returns the single variable x when called with arguments corresponding to the measured parameters

```
X = CALCULATE(U, V, W . . . )
```

We shall further assume that correlations are small so that the covariances may be ignored. Then, to find the variations of x with the measured quantities u, v, and so forth, we can make successive calls to the function of the form

```
DXU = CALCULATE(U + DU, V, W, . . . ) -X,
DXV = CALCULATE(U, V + DV, W, . . . ) - X,
DXW = CALCULATE(U, V, W + DW, . . . ) - X,
ETC.
```

where DU, DV, DW, and so forth are the standard deviations σ_u, σ_v, σ_w, and so on. The resulting contributions to the uncertainty in x are combined in quadrature as

```
DX = SQRT(SQR(DXU) + SQR(DXV) + SQR(DXW) + . . . )
```

Note that it would not be correct to incorporate all the variations into one equation such as

```
DX = CALCULATE(U + DU, V + DV, W + DW, . . . ) - X
```

because this would imply that the errors DU, DV, and so on were actually known quantities, rather than independent, estimated variations of the measured quantities, corresponding to estimates of the widths of the distributions of the measured variables.

SUMMARY

Covariance: $\sigma_{uv}^2 = \langle (u - \bar{u})(v - \bar{v}) \rangle$.

Propagation of errors: Assume $x = f(u, v)$:

$$\sigma_x^2 = \sigma_u^2 \left(\frac{\partial x}{\partial u}\right)^2 + \sigma_v^2 \left(\frac{\partial x}{\partial v}\right)^2 + 2\sigma_{uv}^2 \left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial x}{\partial v}\right)$$

For u and v uncorrelated, $\sigma_{uv}^2 = 0$.

Specific formulas:

$x = au + bv$	$\sigma_x^2 = a^2\sigma_u^2 + b^2\sigma_v^2 + 2ab\sigma_{uv}^2$
$x = auv$	$\frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} + 2\frac{\sigma_{uv}^2}{uv}$
$x = \frac{au}{v}$	$\frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} - 2\frac{\sigma_{uv}^2}{uv}$
$x = au^b$	$\frac{\sigma_x}{x} = b\frac{\sigma_u}{u}$
$x = ae^{bu}$	$\frac{\sigma_x}{x} = b\sigma_u$
$x = a^{bu}$	$\frac{\sigma_x}{x} = (b \ln a)\sigma_u$
$x = a \ln(bu)$	$\sigma_x = ab\frac{\sigma_u}{u}$
$x = a \cos(bu)$	$\sigma_x = -\sigma_u ab \sin(bu)$
$x = a \sin(bu)$	$\sigma_x = \sigma_u ab \cos(bu)$

EXERCISES

3.1 Find the uncertainty σ_x in x as a function of the uncertainties σ_u and σ_v in u and v for the following functions:
(*a*) $x = 1/2(u + v)$ (*b*) $x = 1/2(u - v)$ (*c*) $x = 1/u^2$
(*d*) $x = uv^2$ (*e*) $x = u^2 + v^2$

3.2. If the diameter of a round table is determined to within 1%, how well is its area known? Would it be better to determine its radius to within 1%?

3.3. The resistance R of a cylindrical conductor is proportional to its length L and inversely proportional to its cross-sectional area $A = \pi r^2$. Which should be determined with higher precision, r or L, to optimize the determination of R? How much higher?

3.4. The initial activity N_0 and the mean life τ of a radioactive source are known with uncertainties of 1% each. The activity follows the exponential distribution $N = N_0 e^{-t/\tau}$. The uncertainty in the initial activity N_0 dominates at small t; the uncertainty in the mean life τ dominates at large t ($t \gg \tau$). For what value of t/τ do the uncertainties in N_0 and τ contribute equally to the uncertainty in N? What is the resulting uncertainty in N?

3.5. Snell's law relates the angle of refraction θ_2 of a light ray traveling in a medium of index of refraction n_2 to the angle of incidence θ_1 of a ray traveling in a medium of index n_1 through the equation $n_2 \sin\theta_2 = n_1 \sin\theta_1$. Find n_2 and its uncertainty from the following measurements:

$$\theta_1 = (22.03 \pm 0.2)^\circ \qquad \theta_2 = (14.45 \pm 0.2)^\circ \qquad n_1 = 1.0000$$

3.6 The change in frequency produced by the Doppler shift when a sound source of frequency f is moving with velocity v toward a fixed observer is given by $\Delta f = fv/(u - v)$, where u is the velocity of sound. From the following values of u, f, and v and their uncertainties, calculate Δf and its uncertainty. Which, if any, of the uncertainties make a negligible contribution to the uncertainty in Δf?

$$u_1 = (332 \pm 8) \text{ m/s}; \quad f = (1000 \pm 1)\text{Hz}; \quad \text{and } v = (0.123 \pm 0.003) \text{ m/s}.$$

3.7. The radius of a circle can be calculated from measurements of the length L of a chord and the distance h from the chord to the circumference of the circle from the equation $R = L^2/2h + h/2$. Calculate the radius and its uncertainty from the following values of L and h.
(*a*) $L = (125.0 \pm 5.0)$ cm, $h = (0.51 \pm 0.22)$ cm
(*b*) $L = (125.0 \pm 5.0)$ cm, $h = (57.4 \pm 1.2)$ cm
Was it necessary to use the second term to calculate R in both (*a*) and (*b*)? Explain.

3.8. Students measure the speed of sound in the laboratory by creating a sound pulse that travels down a 1-m tube and reflects back so that both the initial and reflected pulses

are detected by the same microphone. The signals are recorded by computer and the pulse amplitudes versus time are displayed on the monitor. The students measure the time intervals for ten such pairs of pulses on the monitor and record the following transit times in milliseconds:

Trial	1	2	3	4	5	6	7	8	9	10
Transit times	5.77	5.78	5.74	5.80	5.78	5.83	5.76	5.78	5.76	5.78

(*a*) Examine the data and try to estimate the spread of the data, that is, their standard deviation.
(*b*) Calculate the mean transit time, the standard deviation of the sample, and the standard error (error in the mean).
(*c*) One of the transit time measurements differs from the mean by more than 2 standard deviations. In a ten-event sample, how many measurements are predicted by Gaussian statistics to differ from the mean by 2 or more standard deviations? Refer to Table C.2.
(*d*) Calculate the speed of sound and its uncertainty from the data.

3.9 Students in the undergraduate laboratory recorded the following counts in 1-min intervals from a radioactive source. The nominal mean decay rate from the source is 3.7 decays per minute.

Decays per minute	0	1	2	3	4	5	6	7	8	9	10
Frequencey of occurrence	1	9	20	24	19	11	11	0	3	1	1

(*a*) Find the mean decay rate and its standard deviation. Compare the standard deviation to the value expected from the Poisson distribution for the mean value that you obtained.
(*b*) Plot a histogram of the data and show Poisson curves of both the parent and observed distributions.

3.10. Find by numerical integration the probability of observing a value from the Gaussian distribution that is:
(*a*) More than 1 standard deviation (σ) from the mean.
(*b*) More than 2 standard deviations from the mean.
(*c*) More than 3 standard deviations from the mean.

3.11. Find by numerical integration the probability of observing a value from the Lorentzian distribution that is:
(*a*) More than 1 half-width ($\Gamma/2$) from the mean.
(*b*) More than 2 half-widths from the mean.
(*c*) More than 3 half-widths from the mean.

CHAPTER 4

ESTIMATES OF MEAN AND ERRORS

4.1 METHOD OF LEAST SQUARES

In Chapter 2 we defined the mean μ of the parent distribution and noted that the most probable estimate of the mean μ of a random set of observations is the average $\bar{x}$ of the observations. The justification for that statement is based on the assumption that the measurements are distributed according to the Gaussian distribution. In general, we expect the distribution of measurements to be either Gaussian or Poisson, but because these distributions are indistinguishable for most physical situations we can assume the Gaussian distribution is obeyed.

Method of Maximum Likelihood

Assume that, in an experiment, we have observed a set of N data points that are randomly selected from the infinite set of the parent population, distributed according to the parent distribution. If the parent distribution is Gaussian with mean μ and standard deviation σ, the probability dP_i for making any single observation x_i within an interval dx is given by

$$dP_i = p_i dx \tag{4.1}$$

with probability function $p_i = p_G(x_i; \mu, \sigma)$ [see Equation(2.23)]. For simplicity, we shall denote the probability P_i for making an observation x_i by

$$P_i = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right] \tag{4.2}$$

Because, in general, we do not know the mean μ of the distribution for a physical experiment, we must estimate it from some experimentally derived parameter. Let us call the estimate μ'. What formula for deriving μ' from the data will yield the maximum likelihood that the parent distribution had a mean equal to μ?

If we hypothesize a trial distribution with a mean μ' and standard deviation $\sigma' = \sigma$, the probability of observing the value x_i is given by the probability function

$$P_i(\mu') = \frac{1}{\sigma\sqrt{2\Pi}} \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu'}{\sigma}\right)^2\right] \tag{4.3}$$

Considering the entire set of N observations, the probability for observing that particular set is given by the product of the individual probability functions, $P_i(\mu')$,

$$P(\mu') = \prod_{i=1}^{N} P_i(\mu') \tag{4.4}$$

where the symbol Π denotes the product of the N probabilities $P_i(\mu')$.

The product of the constants multiplying the exponential in Equation (4.3) is the same as the product to the Nth power, and the product of the exponentials is the same as the exponential of the sum of the arguments. Therefore, Equation (4.4) reduces to

$$P(\mu') = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \exp\left[-\frac{1}{2}\sum\left(\frac{x_i - \mu'}{\sigma}\right)^2\right] \tag{4.5}$$

According to the *method of maximum likelihood,* if we compare the probabilities $P(\mu')$ of obtaining our set of observations from various parent populations with different means μ' but with the same standard deviation $\sigma' = \sigma$, the probability is greatest that the data were derived from a population with $\mu' = \mu$; that is, the most likely population from which such a set of data might have come is assumed to be the correct one.

Calculation of the Mean

The method of maximum likelihood states that the most probable value for μ' is the one that gives the maximum value for the probability $P(\mu')$ of Equation (4.5). Because this probability is the product of a constant times an exponential to a negative argument, maximizing the probability $P(\mu')$ is equivalent to minimizing the argument X of the exponential,

$$X = -\frac{1}{2}\sum\left(\frac{x_i - \mu'}{\sigma}\right)^2 \tag{4.6}$$

To find the minimum value of a function X we set the derivative of the function to 0,

$$\frac{dX}{d\mu'} = -\frac{d}{d\mu'}\frac{1}{2}\sum\left(\frac{x_i - \mu'}{\sigma}\right)^2 = 0 \tag{4.7}$$

and obtain

$$\frac{dX}{d\mu'} = -\frac{1}{2}\sum\frac{d}{d\mu'}\left(\frac{x_i - \mu'}{\sigma}\right)^2 = \sum\left(\frac{x_i - \mu'}{\sigma^2}\right) = 0 \tag{4.8}$$

which, because σ is a constant, gives

$$\mu' = \bar{x} \equiv \frac{1}{N}\sum x_i \tag{4.9}$$

Thus, the maximum likelihood method for estimating the mean by maximizing the probability $P(\mu')$ of Equation (4.5) shows that the most probable value of the mean is just the average $\bar{x}$ as defined in Equation (1.1).

Estimated Error in the Mean

What uncertainty σ is associated with our determination of the mean μ' in Equation (4.9)? We have assumed that all data points x_i were drawn from the same parent distribution and were thus obtained with an uncertainty characterized by the same standard deviation σ. Each of these data points contributes to the determination of the mean μ' and therefore each data point contributes some uncertainty to the determination of the final results. A histogram of our data points would follow the Gaussian shape, peaking at the value μ' and exhibiting a width corresponding to the standard deviation σ. Clearly we are able to determine the mean to much better than $\pm\sigma$, and our determination will improve as we increase the number of measured points N and are thus able to improve the agreement between our experimental histogram and the smooth Gaussian curve.

In Chapter 3 we developed the error propagation equation [see Equation (3.13)] for finding the contribution of the uncertainties in several terms contributing to a single result. Applying this relation to Equation (4.9) to find the variance σ_μ^2 of the mean μ', we obtain

$$\sigma_\mu^2 = \sum\left[\sigma_i^2\left(\frac{\partial\mu'}{\partial x_i}\right)^2\right] \tag{4.10}$$

where the variance σ_i^2 in each measured data point x_i is weighted by the square of the effect $\partial\mu'/\partial x_i$, that that data point has on the result. This approximation neglects correlations between the measurements x_i as well as second- and higher-order terms in the expansion of the variance σ_μ^2, but it should be a reasonable approximation as long as none of the data points contributes a major portion of the final result.

If the uncertainties of the data points are all equal $\sigma_i = \sigma$, the partial derivatives in Equation (4.10) are simply

$$\frac{\partial\mu'}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\frac{1}{N}\sum x_i\right) = \frac{1}{N} \tag{4.11}$$

and combining Equations (4.10) and (4.11), we obtain

$$\sigma_\mu^2 = \sum \left[\sigma_i^2 \left(\frac{1}{N} \right)^2 \right] = \frac{\sigma^2}{N} \tag{4.12}$$

for the estimated error in the mean σ_μ. Thus, the standard deviation of our determination of the mean μ' and, therefore, the precision of our estimate of the quantity μ, improves as the square root of the number of measurements.

The standard deviation σ of the parent population can be estimated from a consideration of the measuring equipment and conditions, or internally from the data, according to Equation (1.8):

$$\sigma \simeq s = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2} \tag{4.13}$$

which gives for the uncertainty σ_μ in the determination of the mean

$$\sigma_\mu = \frac{\sigma}{\sqrt{N}} \simeq \frac{s}{\sqrt{N}} \tag{4.14}$$

where σ_μ is referred to as the standard deviation of the mean, or the *standard error.* In principle, the value of σ obtained from Equation (4.13) should be consistent with the estimate made from the experimental equipment.

Example 4.1 We return to the student's measurement of the dropped ball (Example 1.2). Let us assume that the time for the ball to fall 2.00 m had been established previously by careful measurements to be $T_{est} = 0.639$ s. The student drops the ball 50 times and concludes, from a consideration of the electronic timer and the experimental arrangement that the uncertainty in each of his individual measurements is ± 0.020 s, consistent with the standard deviation determined from the data. This finite precision of the apparatus results in a spread of observations grouped around the established time as illustrated by the histogram of the data in Figure 1.2.

Because the uncertainties in all the data points are equal ($s_i = s$), the student calculates from his measurements and Equation (4.9) that his estimate of the mean time is $\mu \simeq \bar{T} = 0.635$s, with a standard deviation from Equation (4.13) of $\sigma \simeq s = 0.020$ s. From Equation (4.14), he estimates the uncertainty in his determination of the mean to be $\sigma_\mu \simeq s/\sqrt{N} = 0.020/\sqrt{50}$ or $\sigma_\mu \simeq 0.0028$. He quotes his experimental result as $T_{exp} = (0.635 \pm 0.003)$ s.

To compare his experimental value T_{exp} to the established value T_{est}, the student calculates the number of standard deviations by which the two differ, $n = |T_{exp} - T_{est}|/\sigma_\mu = 1.4$. From the integral of the Gaussian probability equation in Table C.2, we observe that we might expect a measurement to be within 1.4 standard deviations in about 83.8% of repeated experiments, or to exceed 1.4 standard deviations in about 16.2% of the cases.

It is important to realize that the standard deviation of the *data* does not decrease with repeated measurement; it just becomes better determined. On the other hand, the standard deviation of the *mean* decreases as the square root of the number of measurements, indicating the improvement in our ability to estimate the mean of the distribution. Graphically we could illustrate this improvement by plotting a

histogram of the data and noting that our ability to determine the peak of the distribution improves as the number of measurements increases and the distribution becomes smoother.

A Warning About Statistics

Equation (4.12) might suggest that the error in the mean of a set of measurements x_i can be reduced indefinitely by repeated measurements of x_i. We should be aware of the limitations of this equation before assuming that an experimental result can be improved to any desired degree of accuracy if we are willing to do enough work. There are three main limitations to consider: those of available time and resources, those imposed by systematic errors, and those imposed by nonstatistical fluctuations.

The first of these limitations is a very practical one. It may not be possible to take enough repeated measurements to make a significant improvement in the standard deviation of the result. The student of Example 1.2 may be able to make 50 measurements of the time, but might not have the patience to make four times as many measurements to cut the uncertainty by a factor of 2. Similarly, an experiment at a particle accelerator may be assigned 1000 hours of beam time. It may not be possible to increase the allocation to 16,000 hours to improve the precision of the result by a factor of 4.

All experiments are subject to systematic errors at some level. Even after every possible effort has been made to understand the experimental equipment and correct for all known defects and errors of calibration, there comes a point at which further knowledge is unobtainable. For instance, any error in the placement of the detectors that measure times at the beginning and ending of the ball's fall in Example 1.2 will lead to a systematic uncertainty in the time (or in the distance through which the ball fell) and thus in the final result of the experiment.

The phrase "nonstatistical fluctuations" can hide a multitude of sins, or at least problems, in our experiments. It is a rare experiment that follows the Gaussian distribution beyond 3 or 4 standard deviations. More likely, some unexplained data points, or *outliers,* may appear in our data sample, far from the mean. Such points may imply the existence of other contaminating points within the central probability region, masked by the large body of good points. A thorough study of background effects and sources of possible contaminating is obviously required, but at some level, these effects are bound to limit the accuracy of the experiment.

What are we to make of those unexpected points that appear in our data plots well beyond their level of probability? Some may arise from a chance careless measurement. Did our attention wander at the instant when we should have recorded the data point? Did we accidentally interchange two digits in writing down our measurement? Perhaps we can understand and make corrections for some of these effects. Other anomalies in the data may be caused by equipment malfunction. Did our electronic detector respond to a particularly striking clash of metal from the local all-powerful rock radio station? Did our trusty computer decide to check e-mail rather than respond to an urgent data interrupt? And was the distribution that we chose to represent our data the correct one for this experiment?

We may be able to make corrections for these problems, once we are aware of their existence, but there are always others. At some level, things will happen that we cannot understand, and for which we cannot make corrections, and these "things" will cause data to appear where statistically no data should exist, and data points to vanish that should have been there. The moral is, be aware and do not trust statistics in the tails of the distributions.

Elimination of Data Points

There will be occasions when we feel justified in eliminating or correcting outlying data points. For example, suppose that among the time measurements in Example 1.2, the student had recorded one as 0.86s. The student would likely conclude that he had meant to write 0.68s and either ignore or correct the point. What if one measurement had been recorded as 0.72s? Should any action be taken? The point is about 4 standard deviations away from the mean of all the data points, and referring to Table C.2 we see that there is about a 0.06% probability of obtaining in a single measurement a value that is that far from the mean. Thus, in a sample of 50 such measurements we should expect to collect about $50 \times 0.00006 = 0.003$ such events.

The established condition for discarding data in such circumstances is known as Chauvenet's criterion, which states that we should discard a data point if we expect less than half an event to be farther from the mean than the suspect point. If our sample point satisfies this requirement and, as long as we are convinced that our data do indeed follow the Gaussian distribution, we may discard the point with reasonable confidence and recalculate the mean and standard deviation. Thus, for the two examples cited in the preceding paragraph, it would be permissible under Chauvenet's criterion to discard both the 0.86s and the 0.72s data points.

Removing an outlying point has a greater effect on the standard deviation than on the mean of a data sample, because the standard deviation depends on the squares of the deviations from the mean. Deleting one such point will lead to a smaller standard deviation and perhaps another point or two will now become candidates for rejection. We should be very cautious about changing data unless we are confident that we understand the source of the problem we are seeking to correct, and repeated point deletion is generally not recommended. The importance of keeping good records of any changes to the data sample must also be emphasized.

Weighting the Data–Nonuniform Uncertainties

In developing the probability $P(\mu')$ of Equation (4.5) from the individual probabilities $P_i(\mu')$ of Equation (4.3), we assumed that the data points were all extracted from the same parent population. In some circumstances, however, there will be data points that have been measured with better or worse precision than others. We can express this quantitatively by assuming parent distributions with the same mean μ but with different standard deviations σ_i.

If we assign to each data point x_i its own standard deviation σ_i representing the precision with which that particular data point was measured, Equation (4.5) for the probability $P(\mu')$ that the observed set of N data points come from parent distributions with means $\mu_i = \mu'$ and standard deviations σ_i becomes

$$P(\mu') = \prod_{i=1}^{n} \left(\frac{1}{\sigma_i\sqrt{2\pi}}\right) \exp\left[-\frac{1}{2}\sum\left(\frac{x_i - \mu'}{\sigma_i}\right)^2\right] \tag{4.15}$$

Using the method of maximum likelihood, we must maximize this probability, which is equivalent to minimizing the argument in the exponential. Setting the first derivative of the argument to 0, we obtain

$$-\frac{1}{2}\frac{d}{d\mu'}\sum\left(\frac{x_i - \mu'}{\sigma_i}\right)^2 = \sum\left(\frac{x_i - \mu'}{\sigma_i^2}\right) = 0 \tag{4.16}$$

The most probable value is therefore the weighted average of the data points

$$\mu' = \frac{\Sigma(x_i/\sigma_i^2)}{\Sigma(1/\sigma_i^2)} \tag{4.17}$$

where each data point x_i in the sum is weighted inversely by its own variance σ_i^2.

Error in the Weighted Mean

If the uncertainties of the data points are not equal, we evaluate $\partial\mu'/\partial x_i$ from the expression of Equation (4.17) for the mean μ':

$$\frac{\partial\mu'}{\partial x_i} = \frac{\partial}{\partial x_i}\frac{\Sigma(x_i/\sigma_i^2)}{\Sigma(1/\sigma_i^2)} = \frac{1/\sigma_i^2}{\Sigma(1/\sigma_i^2)} \tag{4.18}$$

Substituting this result into Equation (4.10) yields a general formula for the uncertainty of the mean σ:

$$\sigma_\mu^2 = \sum \frac{1/\sigma_i^2}{[\Sigma(1/\sigma_i^2)]^2} = \frac{1}{\Sigma(1/\sigma_i^2)} \tag{4.19}$$

Relative Uncertainties

It may be that the relative values of σ_i are known, but the absolute magnitudes are not. For example, if one set of data is acquired with one scale range and another set with a different scale range, the σ_i may be equal within each set but differ by a known factor between the two sets, as would be the case if σ_i were proportional to the scale range. In such a case, the *relative* values of the σ_i should be included as weighting factors in the determination of the mean μ and its uncertainty, and the *absolute* magnitudes of the σ_i can be estimated from the dispersion of the data points around the mean.

Let us define weighting factors w_i such that

$$kw_i = 1/\sigma_i^2 \tag{4.20}$$

where k is an unknown scaling constant and the σ_i are the standard deviations associated with each measurement. We assume that the weights w_i are known but that the absolute values of the standard deviations σ_i are not. Then, Equation (4.17) can be written

$$\mu' = \frac{\Sigma(x_i/\sigma_i^2)}{\Sigma(1/\sigma_i^2)} = \frac{\Sigma kw_i x_i}{\Sigma kw_i} = \frac{\Sigma w_i x_i}{\Sigma w_i} \tag{4.21}$$

and the result depends only on the relative weights and not on the absolute magnitudes of the σ_i.

To find the error in the estimate μ' of the mean we must calculate a weighted *average variance* of the data:

$$\sigma^2 = \frac{\Sigma w_i (x_i - \mu')^2}{\Sigma w_i} \times \frac{N}{(N-1)} = \left(\frac{\Sigma w_i x_i^2}{\Sigma w_i} - \mu'^2 \right) \times \frac{N}{(N-1)} \tag{4.22}$$

where the last factor corrects for the fact that the mean μ' was itself determined from the data. We may recognize the expression in brackets as the difference between the weighted average of the squares of our measurements x_i and the square of the weighted average. The variance of the mean can then be determined by substituting the expression for σ^2 from Equation (4.22) into Equation (4.14):

$$\sigma_\mu^2 = \frac{\sigma^2}{N} \tag{4.23}$$

If they are required, the value of the scaling constant k and of the values of the separate variances σ_i can be estimated by equating the two expressions for σ_μ of Equations (4.14) and (4.19) and replacing $1/\sigma_i^2$ by kw_i to give

$$\frac{\sigma^2}{N} = \frac{1}{\Sigma(1/\sigma_i^2)} = \frac{1}{k\Sigma w_i} \tag{4.24}$$

so

$$k = \frac{N}{\sigma^2} \frac{1}{\Sigma w_i} \tag{4.25}$$

and therefore

$$\sigma_i^2 = \frac{1}{kw_i} = \frac{\sigma^2 \Sigma w_i}{N w_i} \tag{4.26}$$

Example 4.2. A student performs an experiment to determine the voltage of a standard cell. The student makes 40 measurements with the apparatus and finds a result $\bar{x}_1 = 1.022$ V with a spread $s_1 = 0.01$ V in the observations. After looking over her data she realizes that she could improve the equipment to decrease the uncertainty by a factor of 2.5 ($s_2 = 0.004$ V) so she makes 10 more measurements that yield a result $\bar{x}_2 = 1.018$ V.

The mean of all these observations is given by Equation (4.17):

$$\mu \simeq \bar{x} = \frac{\dfrac{40(1.022)}{0.01^2} + \dfrac{10(1.018)}{0.004^2}}{\dfrac{40}{0.01^2} + \dfrac{10}{0.004^2}} \text{ V}$$

$$= \frac{4.00(1.022) + 6.25(1.018)}{4.00 + 6.25} \text{ V}$$

$$= 1.0196 \text{ V}$$

The uncertainty σ_μ in the mean is given by Equation (4.19):

$$\sigma_\mu \simeq s = \left(\frac{40}{0.01^2} + \frac{10}{0.004^2} \right)^{-1/2} = 0.00099 \text{ V}$$

The result should be quoted as $\mu = (1.0196 \pm 0.0010)$V although $\mu = (1.020 \pm 0.001)$V would also be acceptable. Carrying the fourth place (which is completely undefined) after the decimal point just eliminates any possible rounding errors if these data should later be merged with data from other experiments.

The precision of the final result in Example 4.2 is better than that for either part of the experiment. The uncertainties in the estimates of the means μ_1 and μ_2 determined from the two sets of data independently are given by Equation (4.14):

$$s_2 = \frac{0.01}{\sqrt{40}} \text{ V} = 0.0016 \text{ V} \qquad s_2 = \frac{0.004}{\sqrt{10}} \text{ V} = 0.0013 \text{ V}$$

A comparison of these values illustrates the fact that taking more measurements decreases the resulting uncertainty only as the square root of the number of observations, which for this case is not so important as decreasing σ_i.

What if the student did not know the absolute uncertainties in her measurements, but only that the uncertainties had been improved by a factor of 2.5? She could obtain the estimate of the mean directly from Equation (4.21) by replacing $1/\sigma_1^2$ by the weight $w_i = 1$, and $1/\sigma_2^2$ by the weight $w_i = 2.5^2$, to give

$$\mu \simeq \frac{40(1)(1.022)\text{ V} + 10(2.5^2)(1.018)\text{ V}}{40(1) + 10(2.5)^2} = 1.0196 \text{ V}$$

To find the error in the mean the student could calculate σ from her data by Equation (4.22) and use Equation (4.23) to estimate σ_μ.

Discarding Data

Even though the student in Example 4.2 made four times as many observations at the lower precision (higher uncertainty), the high-precision contribution is over 1.5 times as effective as the low-precision data in determining the mean. The student should probably consider ignoring the low-precision data entirely and using only the high-precision data. Why should we ever throw away data that are not known to be bad? Additionally, because in this case the earlier data are weighted so as to be rather unimportant to the result, what is the point in neglecting them and thereby wasting all the effort that went into collecting those first 40 data points?

These are questions that arise again and again in experimental science as one works to find the elusive parameters of the parent distribution. The answer lies in the fact that experiments tend to be improved over time and often the earliest data-taking period is best considered a training period for the experimenters and a "shakedown" period for the equipment. Why risk contaminating the sample with data of uncertain results when they contribute so little to the final result? The relative standard deviations of the two data sets can serve as a guide. If the spread of the

later distribution shows marked improvement over that of the earlier data, then we should seriously consider throwing away the earlier data unless we are certain of their reliability. There is no hard and fast rule that defines when a group of data should be ignored—common sense must be applied. However, we should make an effort to overcome the natural bias toward using all data simply to recover our investment of time and effort. Greater reliability may be gained by using the cleaner sample alone.

4.2 STATISTICAL FLUCTUATIONS

For some experiments the standard deviations σ_i can be determined more accurately from a knowledge of the estimated parent distribution than from the data or from other experiments. If the observations are known to follow the Gaussian distribution, the standard deviation σ is a free parameter and must be determined experimentally. If, however, the observations are known to be distributed according to the Poisson distribution, the standard deviation is equal to the square root of the mean.

As discussed in Chapter 2, Poisson probability is appropriate for describing the distribution of the data points in counting experiments where the observations are the numbers of events detected per unit time interval. In such experiments, there are fluctuations in the counting rate from observation to observation that result solely from the intrinsically random nature of the process and are independent of any imprecision in measuring the time interval or of any inexactness in counting the number of events occurring in the interval. Because the fluctuations in the observations result from the statistical nature of the process, they are classified as *statistical fluctuations,* and the resulting errors in the final determinations are classified as *statistical errors.*

In any given time interval there is a finite chance of observing *any* positive (or zero) integral number of events. The probability for observing any specific number of counts is given by the Poisson probability function, with mean μ_t, where the subscript t indicates that these are average values for the time interval of length Δt. Thus, if we make N measurements of the number of counts in time intervals of fixed length Δt, we expect that a histogram of the number of counts x_i recorded in each time interval would follow the Poisson distribution for mean μ_t.

Mean and Standard Deviation

For values of the mean μ_t greater than about ten, the Gaussian distribution closely approximates the shape of the Poisson distribution. Therefore, we can use the formula of Equation (4.9) for estimating the mean with the assumption that all data points were extracted from the same parent population and thus have the same uncertainties:

$$\mu_t \simeq \bar{x}_t = \frac{1}{N}\sum x_i \tag{4.27}$$

Here the x_i are the numbers of events detected in the N time intervals Δt, and the assumption that the data were all drawn from the same parent population is equivalent to assuming that the lengths of the time intervals were the same for all measurements.

According to Equation (2.19), the variance σ^2 for a Poisson distribution is equal to the mean μ:

$$\sigma_i^2 = \mu_t \simeq \bar{x}_i \tag{4.28}$$

The uncertainty in the mean σ_{t_μ} is obtained by combining Equations (4.12) and (4.28):

$$\sigma_{t_\mu} = \frac{\sigma_t}{\sqrt{N}} = \sqrt{\frac{\mu_t}{N}} \simeq \sqrt{\frac{\bar{x}_t}{N}} \tag{4.29}$$

We usually wish to find the mean number of counts per unit time, which is just

$$\mu = \frac{\mu_t}{\Delta t} \quad \text{with} \quad \sigma_\mu = \frac{\sigma_{t_\mu}}{\Delta t} = \sqrt{\frac{\mu}{N\Delta t}} \tag{4.30}$$

As we might expect, the uncertainty in the mean number of counts per unit time σ_μ is inversely proportional to the square roots of both the time interval Δt and the number of measurements N.

In some experiments, as in Example 4.2, data may be obtained with varying uncertainties. For purely statistical fluctuations, this implies that counts were recorded in varying time intervals Δt_i. If we wish to find the mean number of counts μ per unit time from such data, there are two possible ways to proceed. If we have the raw data counts (the x_i) and we know they are all independent, then we can simply add all the x_i and divide the sum by the sum of the time intervals:

$$\mu = \frac{\Sigma x_i}{\Sigma \Delta t_i} \quad \text{and} \quad \sigma^2 = \mu$$

The more likely situation is that we know only the means μ_j and corresponding standard deviations σ_j of the means, obtained from the experiments. For example, when dealing with published experimental data, we should assume that the errors incorporate instrumental as well as statistical uncertainties. With such data, the safest procedure is to apply Equations (4.17) and (4.19) to evaluate the weighted mean μ of the individual means μ_i and the standard deviation σ_μ of the mean:

$$\mu \simeq \frac{\Sigma(\mu_j/\sigma_j^2)}{\Sigma(1/\sigma_j^2)} \quad \text{and} \quad \sigma_\mu^2 = \frac{1}{\Sigma(1/\sigma_j^2)} \tag{4.31}$$

Example 4.3. The activity of a radioactive source is measured $N = 10$ times with a time interval $\Delta t = 1$ min. The data are given in Table 4.1. The average of these data points is $\bar{x} = 15.1$ counts per minute. The spread of the data points is characterized by $\sigma = 3.9$ counts per minute calculated from the mean according to Equation (4.27). The uncertainty in the mean is calculated according to Equation (4.29) to be $\sigma_{\bar{x}} \simeq 1.2$ counts per minute.

TABLE 4.1
Experimental data for the activity of a radioactive source from the experiment of Example 4.3

Interval Δt_i (min)	Counts x_i	
1	19	
1	11	
1	24	$\bar{x} = \frac{1}{N}\Sigma x_i = 151$ counts per 10 minutes
1	16	$= 15.1$ counts per minute
1	11	
1	15	
1	22	$\sigma \simeq \sqrt{\bar{x}} = 3.9$ counts per minute
1	9	
1	9	$\sigma_{\bar{x}} \simeq \frac{\sigma}{\sqrt{N}} = 1.2$ counts per minute
1	15	
	Sum = 151	
10	147	$\sigma_{10} = \sqrt{147}$ counts per 10 minutes
		$= 1.2$ counts per minute
Total 20	298	$\bar{x}_{20} = (151 + 147)/(10 + 10)$
		$= 298/20 = 14.9$ counts per minute
		$\sigma_{20} = \sqrt{298}$ counts per 20 minutes
		$= 0.9$ counts per minute

Note: The data tabulated are the number of counts x_i detected in each time interval Δt_i.

If we were to combine the data into one observation $x' = \Sigma x_i$ from one 10-min interval, we would obtain the same result. The activity is $x' = 151$ counts per 10 minutes $= 15.1$ counts per minute as before. The uncertainty in the result is given by the standard deviation of the single data point $\sigma_{x'} = \sqrt{151} = 12.3$ counts per 10 minutes $= 1.2$ counts per minute.

Suppose that we made an additional measurement for a 10-min period and obtained $x'' \cong 147$ counts. We could combine x' and x'' exactly as before to obtain a total

$$\bar{x}_T = x' + x'' = (151 + 147)/(10 + 10) = 14.9 \text{ counts per minute}$$

with an uncertainty

$$\sigma_{\bar{x}_T} = \sqrt{298}/20 = 0.87 \text{ counts per minute}$$

which is smaller than $\sigma_{\bar{x}}$ by a factor of $\sqrt{2}$. Alternatively, we could combine the original data points according to Equation (4.17) and calculate the uncertainty in the final result σ_T by combining the uncertainties of the individual data points according to Equation (4.19).

Note that, although we could have simplified matters by recording all the data as one experimental point, $x = 298$ counts per 20 minutes, by so doing, we would lose all independent information about the shape of the distribution that could be used as a partial check on the validity of the experiment.

4.3 PROBABILITY TESTS

The object of our analysis is to obtain the best estimates, $\bar{x}$ and s_μ, of the mean μ and its uncertainty σ_μ, and to interpret the probability associated with the uncertainty as a measure of our success in determining the parent parameters. Regardless of the method used to make the measurements and analyze the data, we must always estimate the uncertainty in our results to indicate numerically our confidence in them.

Generally, we relate the uncertainty to a Gaussian probability. We have noted that approximately 68% of the measurements in a Gaussian distribution fall within ± 1 standard deviation of the mean μ. Thus, when we find the average of a large number of individual measurements, we expect the distribution of means to be Gaussian, centered on $\bar{x} \simeq \mu$ with width $s \simeq \sigma$, so that approximately 68% of our measurements of x would fall within the range $(\bar{x} - s) < x < (\bar{x} + s)$. Similarly, if we were to repeat the entire experiment many times, we should expect our individual determinations of $\bar{x}$ to form a Gaussian distribution about the mean μ, with width $s_\mu = s/\sqrt{N} \simeq \sigma/\sqrt{N}$. Again, we should expect that approximately 68% of our determinations of $\bar{x}$ should fall within the range $(\mu - s_\mu) < \bar{x} < (\mu + s_\mu)$. If we are convinced that we have made careful and unbiased measurements, we make a slight logical leap to state that there is approximately 68% probability that the true value of the mean μ lies in the range $(\bar{x} - s_\mu) < \mu < (\bar{x} + s_\mu)$ or that the specified range is the 68% *confidence interval.*

Rather than state confidence intervals in terms of 1 standard deviation, we may prefer to state a range that refers to a specific probability level. For example, we may wish to state that our result lies between two values, x_1 and x_2 with a 90% level of confidence, which would correspond to $x_1 = \bar{x} - 1.64\, s_\mu$ and $x_2 = \bar{x} + 1.64\, s_\mu$. Thus, in Example 4.1, the student may report 90% probability that the mean time is within the interval $0.635 \pm (1.64 \times 0.0028)$ s, or $\bar{T} = (0.635 \pm 0.005)$ s at a 90% confidence level. In science, it is customary to report 1 standard deviation uncertainties unless we state otherwise. In other fields, for example political polling, it is customary to report a 95% confidence level, corresponding to approximately 2 standard deviations. American polls are generally accompanied by a statement like "Poll of 1000 adults; margin of error plus or minus 3 percentage points." Canadian media would report "Poll results are likely to be accurate within 3 percentage points 19 times out of 20." If you assume a binomial distribution, you should realize that both statements have almost the same content.

Student's *t* Distribution

We should be aware that Gaussian probability may not apply to our particular data set, and even an experimental distribution that nominally follows Gaussian statistics is apt to deviate in the tails. When the data set is small, there is another consideration. Not only the mean, but also our estimate s_μ of the standard error σ_μ may be

poorly determined. The probabilities that we calculate from the Gaussian distribution take no account of the latter problem.

In such cases, a better estimate of the probability can be obtained from *Student's t* distribution,[1] which describes the distribution of the parameter $t = |x - \bar{x}|/s_\mu$, where t is the number of standard deviations of the sample distribution s_μ by which x differs from $\bar{x}$.

$$p_t(t, \nu) = \frac{1}{\sqrt{(\nu\pi)}} \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

where the gamma function $\Gamma(n)$ is equivalent to the factorial function n! extended to nonintegral arguments. (See Equation 11.7).

Unlike the Gaussian distribution, Student's t distribution depends upon the number of degrees of freedom ν. If $\bar{x}$ represents the mean of N numbers and x is not derived from the data, then $\nu = N - 1$. If both x and $\bar{x}$ are means, s_μ must be the joint standard deviation of x and $\bar{x}$, and ν must be the total number of degrees of freedom. In the limit of large ν, Student's t and Gaussian probability distributions agree. As with the Gaussian distribution, we are usually interested in integrated values that relate to the probability of obtaining a result within a specific range $\pm t$ standard deviations. For example, we might wish to report our estimate of the probability that the true value of μ lies within the range $(\bar{x} - ts_\mu) < \mu < (\bar{x} + ts_\mu)$ with $t = |\bar{x} - \mu|/s_\mu$.

Table C.8 lists probabilities obtained by integrating the Student's t distribution from $x = \bar{x} - ts_\mu$ to $x = \bar{x} + ts_\mu$ for specified values of t and the number of degrees of freedom ν. The corresponding values for Gaussian probability (which are independent of ν) are listed in the last column.

Consider again Example 4.1 in which the student made 50 time measurements and found that the mean of his measurements deviated by $1.4s_\mu$ from the established value. From Gaussian probability we observed that approximately 84% of experiments should yield a result that is within 1.4 standard deviations of the expected result. From Student's t distribution (Table C.8.), we observe that the probability is lower by about 0.6%. However, suppose the student made only six measurements using a more precise measuring system and again obtained a result that differed from the mean by $t = 1.4s_\mu$ (see Exercise 4.12). Small numbers of measurements are common in undergraduate laboratory experiments, where time may be short and the measurements may be tedious. What probability is implied for 5 degrees of freedom by a difference of $t = 1.4s_\mu$? The Gaussian probability is unchanged at ~84%; Student's t predicts ~78%. Thus, for experiments with only a few degrees of freedom, Gaussian probability overestimates the confidence level associated with a given range t. Another way of looking at this is to note that, for the same confidence level, Student's t probability requires a larger uncertainty estimate than does Gaussian probability.

Generally, a result is considered to be significant only at confidence levels of 95% or better. In Gaussian probability, this corresponds to a range of approximately $\pm 2\sigma$. We can observe from Table C.8 that for a sample of only three data points

[1]"Review of Particle Physics," *The European Physical Journal C,* vol. 15, p. 193 (2000)

($\nu = 2$), the Student's t probability for 95% confidence corresponds to a range of more than $\pm 4\sigma$.

4.4 CHI-SQUARE TESTS OF A DISTRIBUTION

Once we have calculated the mean and standard deviation from our data, we may be in a position to say even more about the parent population. If we can be fairly confident of the type of parent distribution that describes the spread of the data points (e.g., Gaussian or Poisson distribution), then we can describe the parent distribution in detail and predict the outcome of future experiments from a statistical point of view.

Because we are concerned with the behavior of the probability density function $p(x_i)$ as a function of the observed values of x_i, a complete discussion will be postponed until Chapter 11 following the development of procedures for comparing data with complex functions. Let us for now use the results of Chapter 11 without derivation. The test that we shall describe here is the χ^2 (chi-square) test for goodness of fit.

Probability Distribution

If N measurements x_i are made of the quantity x, we can truncate the data to a common least count and group the observations into frequencies of identical observations to make a histogram. Let us assume that j runs from 1 to n so there are n possible different values of x_j, and let us call the frequency of observations, or number of counts in each histogram bin, $h(x_j)$ for each different measured value of x_j. If the probability for observing the value x_j in any random measurement is denoted by $P(x_j)$, then the expected number of such observations is $y(x_j) = NP(x_j)$, where N is the total number of measurements. Figures 4.1 and 4.2 show the same six-bin histogram, drawn from a Gaussian parent distribution with mean $\mu = 5.0$ and standard deviation $\sigma = 1$, corresponding to 100 total measurements. The parent distribution, $y(x_j) = NP(x_j)$, is illustrated by the solid Gaussian curve on each histogram.

For each measured value x_j, there is a standard deviation $\sigma_j(\mathrm{h})$ associated with the uncertainty in the observed frequency $h(x_j)$. This is not the same as the uncertainty σ_i associated with the spread of the individual measurements x_i about their mean μ, but rather describes the spread of the measurements of each of the frequencies $h(x_j)$ about its mean μ_j. If we were to repeat the experiment many times to determine the distribution of frequency measurements at each value of x_j, we should find each parent distribution to be Poisson with mean $\mu_j = y(x_j)$ and variance $\sigma_j^2(y) = y(x_j)$. Thus, for each value of x_j, there is a distribution curve, $P_j(y_k)$, that describes the probability of obtaining the value of the frequency $h_k(x_j)$ in the kth trial experiment when the expected value is $y(x_j)$. It is the spread of these measurements for each value of j that is characterized by $\sigma_j(h)$. These distributions are illustrated in Figures 4.1 and 4.2 as dotted Poisson curves at each value of x_j. In Figure 4.1 the Poisson curves are centered at the observed frequencies $h(x_j)$ with standard deviations $\sigma_j(h) = \sqrt{h(x_j)}$. In principle, we should center the Poisson curves at the

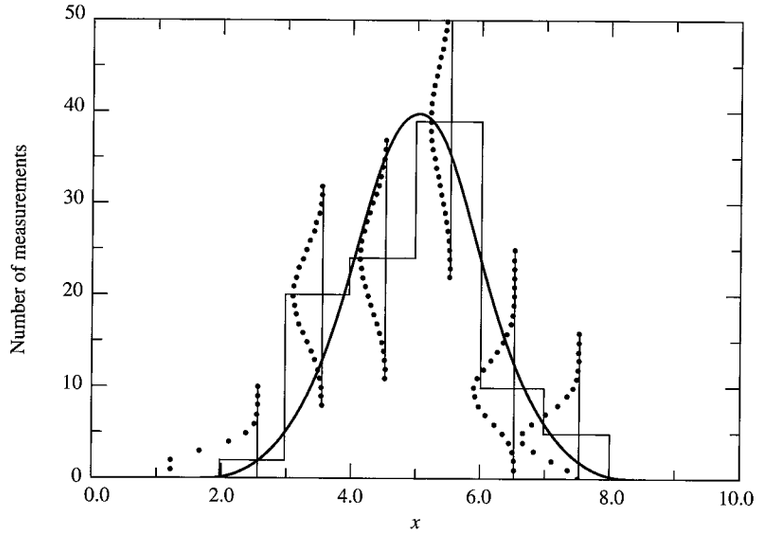


FIGURE 4.1
Histogram, drawn from a Gaussian distribution mean $\mu = 5.0$ and standard deviation $\sigma = 1$, corresponding to 100 total measurements. The parent distribution $y(x_j) = NP(x_j)$ is illustrated by the large Gaussian curve. The smaller dotted curves represent the Poisson distribution of events in each bin, based on the sample data.

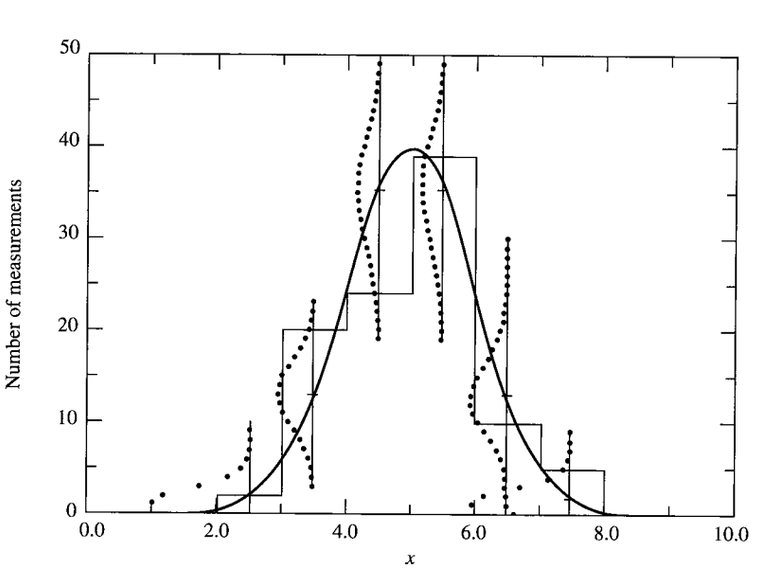


FIGURE 4.2
The same histogram as shown in Figure 4.1 with dotted curves representing the Poisson distribution of events in each bin, based on the parent distribution.

frequencies $\mu_j = y(x_j)$ with standard deviation $\sigma_j(h) = \sqrt{\mu_j}$ of the parent population as illustrated in Figure 4.2. However, in an actual experiment, we generally would not know these parameters.

Definition of χ^2

With the preceding definitions for n, N, x_j, $h(x_j)$, $P(x_j)$, and $\sigma_j(h)$, the definition of χ^2 from Chapter 11 is

$$\chi^2 \equiv \sum_{j=1}^{n} \frac{[h(x_j) - NP(x_j)]^2}{\sigma_j(h)^2} \tag{4.32}$$

In most experiments, however, we do not know the values of $\sigma_j(h)$ because we make only one set of measurements $f(x_j)$. Fortunately, these uncertainties can be estimated from the data directly without measuring them explicitly.

If we consider the data of Figure 4.2, we observe that for each value of x_j, we have extracted a proportionate random sample of the parent population for that value. The fluctuations in the observed frequencies $h(x_j)$ come from the statistical probabilities of making random selections of finite numbers of items and are distributed according to the Poisson distribution with $y(x_j)$ as mean. Although the distribution of frequencies $y(x_j)$ in Figure 4.2 is Gaussian, the probability functions for the spreads of the measurements of each frequency are Poisson distributions.

For the Poisson distribution, the variance $\sigma_j(h)^2$ is equal to the mean $y(x_j)$ of the distribution, and thus we can estimate $\sigma_j(h)$ from the data to be $\sigma_j(h) = \sqrt{NP(x_j)} \simeq \sqrt{h(x_j)}$. Equation (4.32) simplifies to

$$\chi^2 \equiv \sum_{j=1}^{n} \frac{[h(x_j) - NP(x_j)]^2}{NP(x_j)} \simeq \sum_{j=1}^{n} \frac{[h(x_j) - NP(x_j)]^2}{h(x_j)} \tag{4.33}$$

Test of χ^2

As defined in Equations (4.32) and (4.33), χ^2 is a statistic that characterizes the dispersion of the observed frequencies from the expected frequencies. If the observed frequencies were to agree exactly with the predicted frequencies $h(x_j) = NP(x_j)$, then we should find $\chi^2 = 0$. From our understanding of probability, we realize that this is not a very likely outcome of an experiment. The numerator of Equation (4.32) is a measure of the spread of the observations; the denominator is a measure of the expected spread. We might imagine that for good agreement, the average spread of the data would correspond to the expected spread, and thus we should get a contribution of about one from each frequency, or $\chi^2 \simeq n$ for the entire distribution. This is almost correct. In fact, the true expectation value for χ^2 is

$$\langle \chi^2 \rangle = \nu = n - n_c \tag{4.34}$$

where ν is the number of degrees of freedom and is equal to the number n of sample frequencies minus the number n_c of constraints or parameters that have been calculated from the data to describe the probability function $NP(x_j)$. For our example,

even if $NP(x_j)$ is chosen completely independently of the distribution $h(x_j)$, there is still the normalizing factor N corresponding to the total number of events in the distribution, so that the expectation value of χ^2 must at best be $\langle \chi^2 \rangle = n - 1$.

In order to estimate the probability that our calculated values of χ^2 are consistent with our expected distribution of the data, we must know how χ^2 is distributed. If our value of χ^2 corresponds to a reasonable high probability, then we can have confidence in our assumed distribution.

It is convenient to define the *reduced chi-square* as $\chi^2_\nu \equiv \chi^2/\nu$, with expectation value $\langle \chi^2_\nu \rangle = 1$. Values of χ^2_ν much larger than 1 result from large deviations from the assumed distribution and may indicate poor measurements, incorrect assignment of uncertainties, or an incorrect choice of probability function. Very small values of χ^2_ν are equally unacceptable and may imply some misunderstanding of the experiment. Rather than consider the probability of obtaining any particular value of χ^2 or χ^2_ν (which is infinitesimally small), we shall use an integral test to determine the probability of observing a value of χ^2_ν equal to or greater than the one we calculated. This is similar to our consideration of the probability that a measurement of a variable deviates by more than a certain amount from the mean.

Table C.4 gives the probability that a random sample of data points drawn from the assumed probability distribution would yield a value of χ^2 as large as or larger than the observed value in a given experiment with ν degrees of freedom.

If the probability is reasonably close to 1, then the assumed distribution describes the spread of the data points well. If the probability is small, either the assumed distribution is not a good estimate of the parent distribution or the data sample is not representative of the parent distribution. There is no yes-or-no answer to the test; in fact, we should expect to find a probability of about 0.5 with $\chi^2_\nu \simeq 1$, because statistically the observed values of χ^2 should exceed the norm half the time. But in most cases, the probability is either reasonably large or unreasonably small, and the test is fairly conclusive. A further discussion of the statistical significance of the χ^2 probability function will be given in Chapter 11.

Let us consider again the data of Example 1.2 (and 4.1), which are summarized as a histogram in Figure 1.2 with the frequencies listed in Table 4.2. To test the agreement between the data and the predicted distribution, we have calculated the function $y(x_j) = NP(x_j)$ at each value of x_j from the mean and standard deviation of the parent distribution (column 3 of Table 4.2), and from the mean and standard deviation of the data, that is, from the sample distribution (column 6). The uncertainties σ_j calculated as the square roots of the values predicted by the parent distribution and by the sample distribution are listed in columns 4 and 7 respectively. The individual contributions (before squaring) to the values of χ^2, $[h(x_j) - NP(x_j)]/\sigma_j$, are listed in columns 5 and 8. The calculated values of χ^2 from the comparison between the data and each distribution are the sums of the squares of these last quantities.

For the comparison of the 11 data points with the parent distribution we have one constraint, the normalization constant N determined from the data, and therefore the expectation value of χ^2 is $\nu = 11 - 1 = 10$. We obtained $\chi^2 = 13.03$ and thus, $\chi^2_\nu = 1.30$. Interpolating in Table C.4, we observe that the corresponding probability of obtaining a value $\chi^2_\nu \geq 1.30$ with 10 degrees of freedom is ~23%. For a similar comparison with an *estimate* of the parent distribution based on the mean

and standard deviation of the data, we have two additional constraints, the mean and standard deviation. Thus, for this comparison, the expectation value of χ^2 is $\nu = 11 - 3 = 8$. We obtained $\chi^2 = 7.85$ and, thus, $\chi^2_\nu = 0.98$. The corresponding probability for obtaining a value $\chi^2_\nu \geq 0.98$ with 8 degrees of freedom is ~45%.

Generalizations of the χ^2 Test

In the preceding example we knew the parent distributions and were therefore able to determine the uncertainties $\sigma_j(h)$ from the predicted probability. In most cases, where the actual parameters of the probability function are being determined in the calculation, we must use an estimate of the parent population based on these parameters and must estimate the uncertainties in the $y(x_j)$ from the data themselves. To do this we must replace the uncertainties in columns 4 and 7 of Table 4.2 with the square roots of the observed frequencies in column 2.

Furthermore, although our example was clearly based on a simple probability function, the χ^2 test is often generalized to compare data obtained in any type of experiment to the prediction of a model. The uncertainties in the measurements may be instrumental or statistical or a combination of both, and the uncertainty $\sigma_j(h)^2$ in the denominator of Equation (4.32) may represent a Gaussian error distribution rather than the Poisson distribution. In fact, several of the histogram bins in our example contained small numbers of counts, and thus, the statistical application of the test was not strictly correct, because we assume Gaussian statistics in the χ^2 calculation. However, the test still provides us with a reproducible method of evaluating

TABLE 4.2
χ^2 analysis of the data of Example 4.1

	Observed frequency	From parent distribution			From sample distribution		
Time	h_j	y_j	σ_j	$\frac{y_j - h_j}{\sigma_j}$	y_j	σ_j	$\frac{y_j - h_j}{\sigma_j}$
0.595	2	0.89	0.94	−1.18	1.35	1.16	0.56
0.605	2	2.35	1.53	0.23	3.24	1.80	−0.69
0.615	11	4.85	2.20	−2.79	6.05	2.46	2.01
0.625	6	7.81	2.79	0.65	8.80	2.97	−0.94
0.635	12	9.78	3.13	−0.71	9.97	3.16	0.64
0.645	8	9.53	3.09	0.50	8.80	2.97	−0.27
0.655	4	7.24	2.69	1.20	6.05	2.46	−0.83
0.665	3	4.28	2.07	0.62	3.24	1.80	−0.13
0.675	1	1.97	1.40	0.69	1.35	1.16	−0.30
0.685	1	0.71	0.84	−0.35	0.44	0.66	0.85
0.695	0	0.20	0.44	0.44	0.11	0.33	−0.33
		$\chi^2_\nu = 13.03/10 = 1.30$			$\chi^2_\nu = 7.85/8 = 0.98$		

Note: Parameters of the parent Gaussian distribution are $\mu = 0.639$ and $\sigma = 0.020$ s; parameters estimated from the sample distribution are $\mu = 0.635$ s and $\sigma = 0.020$ s.

the quality of our data, and if we are concerned with statistical accuracy, we can merge the low-count bins to satisfy the Gaussian statistics requirement.

Another application of the chi-squared test is in comparing two sets of data to attempt to decide whether or not they were drawn from the same parent population. Suppose that we have measured two distributions, $g(x_j)$ and $h(x_j)$, and wish to determine the probability that the two sets were not drawn from the same parent probability distribution $P(x_j)$. Clearly, we could apply the χ^2 test separately to the two sets of data and determine separately χ^2 probabilities that each set was not associated with the supposed parent population $P(x_j)$. However, we can also make a direct test, independent of the parent population, by writing

$$\chi^2 = \sum_{j=1}^{n} \frac{[g(x_j) - h(x_j)]^2}{\sigma^2(g) + \sigma^2(h)} \tag{4.35}$$

The denominator $\sigma^2(g) + \sigma^2(h)$ is just the variance of the difference $g(x_j) - h(x_j)$. As in the previous examples, the expectation value of χ^2 depends on the relation between the two parts of the numerator, $g(x_j)$ and $h(x_j)$. If the two parts, corresponding to the distributions of the two data sets, were obtained completely independently of one another, then the number of degrees of freedom equals n and $\langle\chi^2\rangle = n$. If one of the distributions $g(x_j)$ or $h(x_j)$ has been normalized to the other, then the number of degrees of freedom is reduced by 1 and $\langle\chi^2\rangle = n - 1$. Again, we interpret the χ^2 probability in a negative sense. If the value of χ^2/ν is large, and therefore the probability given in Table C.4 is low, we may conclude that the two sets of data were drawn from different distributions. However, for a low value of χ^2 and therefore high probability, we cannot draw the opposite conclusion that the two data sets $g(x_j)$ and $h(x_j)$ were drawn from the same distribution. There is always the possibility that there are indeed two different but closely similar distributions and that our data are not sufficiently sensitive to detect the difference between the two.

Constraints and Degrees of Freedom

Equation (4.34) defines the number of degrees of freedom, ν, and $<\chi^2>$, the expectation value of χ^2. To clarify the relation between constraints and degrees of freedom in a χ^2 test, consider a data set that is expected to show a linear relation between the measured values x_j and h_j, that is,

$$y_j = A + Bx_j$$

Clearly, two measurements of y at two different values of x are required just to define the two parameters, A and B, of the straight line so there are two constraints ($n_c = 2$) on the system and at least three measurements ($n = 3$) must be made before a test can be applied. Under these circumstances, if we assume that points $j = 1$ and $j = 2$ are used to calculate A and B, Equation (4.32) becomes

$$\chi^2 = (h_3 - y_3)^2/\sigma_3^2(h)$$

and we should expect to find

$$\langle\chi^2\rangle = n - n_c = 3 - 2 = 1$$

Similarly, if we measure $n = 4$ points, there will be two points available for the χ^2 test or 2 degrees of freedom. Of course, in general, we would not use just two points to calculate the two parameters. Rather, we should perform a *least-squares fit* in which all measurements are treated equally (or weighted according to their uncertainties). However, the same principle holds: we impose two constraints on our calculation to define the two parameters of a straight line, leaving 2 degrees of freedom.

SUMMARY

Weighted mean:

$$\bar{x} = \frac{\Sigma(x_i/\sigma_i^2)}{\Sigma(1/\sigma_i^2)} \xrightarrow[\sigma_i=\sigma]{} \frac{1}{N}\sum x_i$$

Variance of mean:

$$\sigma_\mu^2 = \frac{1}{\Sigma(1/\sigma_j^2)} \xrightarrow[\sigma_i=\sigma]{} \frac{\sigma^2}{N}$$

Instrumental uncertainties: Fluctuations in measurements due to finite precision of measuring instruments:

$$\sigma^2 \simeq s^2 = \frac{1}{N-1}\sum (x_i - \bar{x})^2$$

Statistical fluctuations: Fluctuations in observations resulting from statistical probability of taking random samples of finite numbers of items:

$$\sigma^2 = \mu \simeq \bar{x}$$

χ^2 *test:* Comparison of observed frequency distribution $h(x_j)$ of possible observations x_j versus predicted distribution $NP(x_j)$, where N is the number of data points and $P(x_j)$ is the theoretical probability distribution:

$$\chi^2 \equiv \sum_{j=1}^{n} \frac{[h(x_j) - NP(x_j)]^2}{\sigma_j(h)^2}$$

Degrees of freedom ν: Number of data points minus the number of parameters to be determined from the data points.
Reduced χ^2: $\chi_\nu^2 = \chi^2/\nu$. For χ^2 tests, χ_ν^2 should be approximately equal to 1.
Graphs and tables of χ^2: Table C.4 gives the probability that a random sample of data when compared to its *parent distribution* would yield values of χ_ν^2 as large as or larger than the observed value.

EXERCISES

4.1. Calculate the standard deviation and the error in the mean value of x from the data of Exercise 1.4. Are the values reasonable? (See Exercise 2.4.)

4.2. Repeat Exercise 4.1 for the data of Exercise 1.5.

4.3. Read the data of Example 2.4 from Figures 2.3 and 2.4. Recalculate the curves and calculate χ^2 and χ^2_ν for the agreement between the curves and the histograms. Use only bins with five or more counts.

4.4. Work out the intermediate steps in Equation (4.19).

4.5. A student measures the period of a pendulum and obtains the following values.

Trial	1	2	3	4	5	6	7	8
Period	1.35	1.34	1.32	1.36	1.33	1.34	1.37	1.35

(*a*) Find the mean and standard deviation of the measurements and the standard deviation of the mean.

(*b*) Estimate the probability that another single measurement will fall within 0.02 s of the mean.

4.6. (*a*) Find the mean and the standard deviation of the mean of the following numbers under the assumption that they were all drawn from the same parent population.

(*b*) In fact, data points 1 through 20 were measured with uniform uncertainty σ, whereas data points 21 through 30 were measured more carefully so that the uniform uncertainty was only $\sigma/2$. Find the mean and standard deviation of the mean under these conditions.

Trial	$x(\sigma)$	Trial	$x(\sigma)$	Trial	$x(\sigma/2)$
1	2.40	11	1.94	21	2.59
2	2.45	12	1.55	22	2.65
3	2.47	13	2.12	23	2.55
4	3.13	14	2.17	24	2.07
5	2.92	15	3.06	25	2.61
6	2.85	16	1.97	26	2.61
7	2.05	17	2.23	27	2.54
8	2.52	18	3.20	28	2.76
9	2.94	19	2.24	29	2.37
10	1.89	20	2.60	30	2.57

4.7. A counter is set to count gamma rays from a radioactive source. The total number of counts, including background, recorded in each 1-min interval is listed in the accompanying table. An independent measurement of the background in a 5-min interval gave 58 counts. From these data find:

(*a*) The mean background in a 1-min interval and its uncertainty.

(*b*) The corrected counting rate from the source alone and its uncertainty.

Trial	1	2	3	4	5	6	7	8	9	10
Total counts	125	130	105	126	128	119	137	131	115	116

4.8. The *Particle Data Tables* list the following eight experimental measurements of the mean lifetime of the K_s meson with their uncertainties, in units of 10^{-10} s. Find the weighted mean of the data and the uncertainty in the mean.

0.8971 ± 0.0021 0.8941 ± 0.0014 0.8929 ± 0.0016 0.8920 ± 0.0044 0.881 ± 0.009
0.8924 ± 0.0032 0.8937 ± 0.0048 0.8958 ± 0.0045

4.9. Eleven students in an undergraduate laboratory combined their measurements of the mean lifetime of an excited state. Their individual measurements are tabulated.

Student	1	2	3	4	5	6	7	8	9	10	11
$\tau(s)$	34.3	32.2	35.4	33.5	34.7	33.5	27.9	32.0	32.4	31.0	19.8
σ_τ	1.6	1.2	1.5	1.4	1.6	1.5	1.9	1.2	1.4	1.8	1.3

Find the maximum likelihood estimate of the mean and its uncertainty.

4.10. Assume that you have a box of resistors that have a Gaussian distribution of resistances with mean value $\mu = 100\ \Omega$ and standard deviation $\sigma = 20\ \Omega$ (i.e., 20% resistors). Suppose that you wish to form a subgroup of resistors with $\mu = 100\ \Omega$ and standard deviation of 5 Ω (i.e., 5% resistors) by selecting all resistors with resistance between the two limits $r_1 = \mu - a$ and $r_2 = \mu + a$.

(*a*) Find the value of a.

(*b*) What fraction of the resistors should satisfy the condition?

(*c*) Find the standard deviation of the remaining sample.

4.11. Suppose that 1000 adults responded to a poll about a current bill in Congress, and that 622 approved, while 378 disapproved.

(*a*) Assume that there was 50% a priori probability of obtaining either answer and calculate the standard deviation of the result. Find the "margin of error," that is, the uncertainty that corresponds to a 95% confidence interval. (Use Gaussian probability. Justify this.)

(*b*) Assume the probabilities implied by the observed numbers of votes in each category and repeat the calculation. Note the insensitivity of the standard deviation of the binomial distribution to variations in probability near 50%.

(*c*) Refer to the two statements about polling reports in Section 4.3 and show that they are approximately equivalent.

4.12. Six measurements of the length of a wooden block yielded the following values: 20.3, 20.4, 19.8, 20.4, 19.9, 20.7.

(*a*) From these numbers, calculate the mean, standard deviation, and standard error. Assume that the actual mean length has been established by previous measurements to be 20.00 cm and calculate t, the number of standard errors by which the calculated mean differs from the established value.

Refer to the tables in Appendix C to find the limits on the 95% confidence level for both Gaussian and Student's t probabilities.

(*b*) The experiment was repeated to obtain a total of 25 data sets of six measurements each from which the following 25 values of the mean were calculated.

20.25 20.10 20.02 20.12 20.00 19.73 19.73 20.13 20.22 20.22 20.27 19.83 20.00
19.77 20.10 20.28 19.97 19.88 20.32 19.98 20.05 20.23 19.92 19.97 19.77

Find the mean of these "means" and calculate their standard deviation. Compare this standard deviation to the standard error calculated in (*a*).

4.13. The following data represent the frequency distribution of 200 variables drawn from a parent Gaussian population with mean $\mu = 26.00$ and standard deviation $\sigma = 5.00$. The bins are two units wide and the lower edge of the first bin is at $x = 14$.

4 8 11 20 26 31 29 22 26 13 5 2 3

(*a*) Plot a histogram of these data.

(*b*) From the mean μ and standard deviation σ, calculate the Gaussian function that represents the parent distribution, normalized to the area of the histogram. Your first point should be calculated at $x = 15$, the midpoint of the first bin.

(*c*) Calculate χ^2 to test the agreement between the data and the theoretical curve.

(*d*) What is the expectation value of χ^2?

(*e*) Refer to Table C.4 to find the χ^2 probability of the fit, that is, the probability of drawing a random sample from the parent population that will yield a value of χ^2 as large as or larger than your calculated value.

4.14. Plot a histogram in ten-point bins of the course grades listed in Exercise 1.5. Plot a Gaussian curve based on the mean and standard deviation of the data, normalized to the area of the histogram. Apply the χ^2 test and check the associated probability from Table C.4.

CHAPTER 5

MONTE CARLO TECHNIQUES

5.1 INTRODUCTION

We saw in Chapter 4 the importance of probability distributions in the analysis of data samples, and observed that we are usually interested in the integrals or sums of such distributions over specified ranges. Although we have considered only experiments that are described by a single distribution, most experiments involve a combination of many different probability distributions. Consider, for example, a simple scattering experiment to measure the angular distribution of particles scattered from protons in a fixed target. The magnitude and direction of the momentum vector of the incident particles, the probability that a particle will collide with a proton in the target, and the resulting momentum vectors of the scattered particles can all be described in terms of probability distributions. The final experimental result can be treated in terms of a multiple integration over all these distributions.

Analytical evaluation of such an integral is rarely possible, so numerical methods must be used. However, even the simplest first-order numerical integration can become very tedious for a multidimensional integral. A one-dimensional integral of a function can be determined efficiently by evaluating the function N times on a regular grid, where the number of samples N depends on the structure of the function and the required accuracy. (See Appendix A.3.) A two-dimensional integral requires sampling in two dimensions and, for accuracy comparable to that of the corresponding one-dimensional problem, requires something like N^2 samples. A three-dimensional integral requires something like N^3 samples. For integrals with many dimensions, the number of grid points at which the function must be calculated becomes excessively large.

Before we continue with methods of extracting parameters from data, let us look at the Monte Carlo method, a way of evaluating these multiple integrals that depends on random sampling from probability density distributions, rather than regular grid-based sampling techniques. The Monte Carlo method provides the experimental scientist with one of the most powerful tools available for planning experiments and analyzing data. Basically, Monte Carlo is a method of calculating multiple integrals by random sampling. Practically, it provides a method of simulating experiments and creating models of experimental data. With a Monte Carlo calculation, we can test the statistical significance of data with relatively simple calculations that require neither a deep theoretical understanding of statistical analysis nor sophisticated programming techniques.

The name *Monte Carlo* comes from the city on the Mediterranean with its famous casino, and a Monte Carlo calculation implies a statistical method of studying problems based on the use of random numbers, similar to those generated in the casino games of chance. One might reasonably ask whether the study of science can be aided by such associations, but in fact, with Monte Carlo techniques, very complicated scientific and mathematical problems can be solved with considerable ease and precision.

Example 5.1. Suppose that we wish to find the area of a circle of radius r_c but have forgotten the equation. We might inscribe the circle within a square of known area A_s and cover the surface of the square uniformly with small markers, say grains of rice. We find the ratio of the number of grains that lie within the circle to those that cover the square, and determine the area of the circle A_c from the relation

$$A_c = A_s N_c / N_s \tag{5.1}$$

where N_c and N_s are the numbers of grains of rice within the boundaries of the circle and of the square, respectively.

What would be the accuracy of this determination; that is, how close should we expect our answer to agree with the true value for the area of a circle? Clearly it would depend on the number and size of the rice grains relative to the size of the square, and on the uniformity of both the grains and their distribution over the square. What if we decided that instead of attempting to cover the square uniformly, we would be content with a random sampling obtained by tossing the rice grains from a distance so that they landed randomly on the square, with every location equally probable? Then we would obtain an interesting result: Our problem would reduce to a simple binomial calculation as long as we did not overpopulate the square but kept the density of rice grains low so that position of any grain on the square was not influenced by the presence of other grains. We should find that, for a fixed number of grains N_s thrown onto the square, the uncertainty σ in the measurement of the circular area would be given by the standard deviation for the binomial distribution with probability $p = A_c/A_s$,

$$\sigma = \sqrt{N_s p(1-p)} = \sqrt{N_c(1-p)} \tag{5.2}$$

Thus, if we were to increase the number of rice grains N_c by a factor of 4, the relative error in our determination of the area of the circle would decrease by a factor of 2.

Replacing the tossed rice grains by a set of computer generated random numbers is an obvious improvement. Let us inscribe our circle of unit radius in a square of side length 2, and generate $N = 100$ pairs of random numbers between -1 and $+1$ to determine the area. Then the probability of a "hit" is just the ratio of the area of the circle to the area of a square, or $p = \pi/4$, so in 100 tries, the mean number of hits will be $\mu = 100p = 78.5$, and the standard deviation, from Equation (5.2), will be $\sigma = \sqrt{Np(1-p} = \sqrt{100(\pi/4)(1-\pi/4)} = 4.1$. For our measurements of the area of the circle with 100 tries we should expect to obtain from Equation (5.1) $A_c = A_s \times N_c/N_s = (78.5 \pm 4.1) \times 2^2/100 = 3.14 \pm 0.16$.

Figure 5.1 shows a typical distribution of hits from one "toss" of 100 pairs of random numbers. In this example there were 73 hits, so we should estimate the area and its uncertainty from Equations (5.1) and (5.2) to be $A = 2.92 \pm 0.18$. To determine the uncertainty, we assumed that we did not know the a priori probability $p = \pi/4$ and, therefore, we used our experimental estimate $p \simeq 73/100$.

Figure 5.2 shows a histogram of the circle area estimates obtained in 100 independent Monte Carlo runs, each with 100 pairs of random numbers (or a total of 10,000 "tosses"). The Gaussian curve was calculated from the mean, $A = 3.127$, and standard deviation, $\sigma = 0.156$, of the 100 estimated areas.

Obviously, the area determination problem of Example 5.1 is much too simple to require a Monte Carlo calculation. However, for problems involving integrations of many variables and for those with complicated integration limits, the Monte

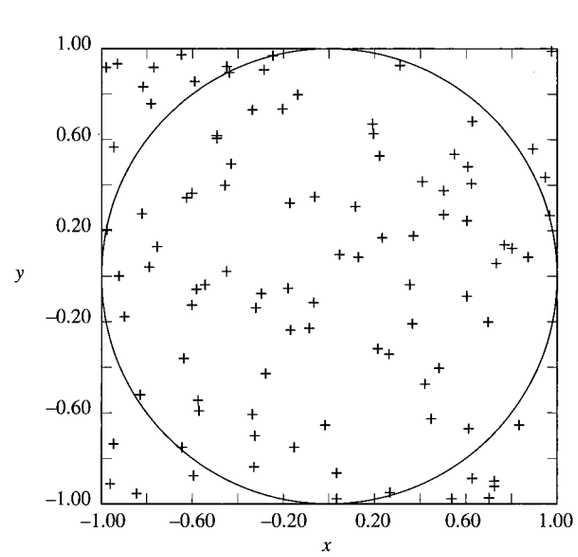


FIGURE 5.1
Estimation of the area of a circle by the Monte Carlo method. The plot illustrates a typical distribution of hits from one "toss" of 100 pairs of random numbers uniformly distributed between -1.00 and $+1.00$.

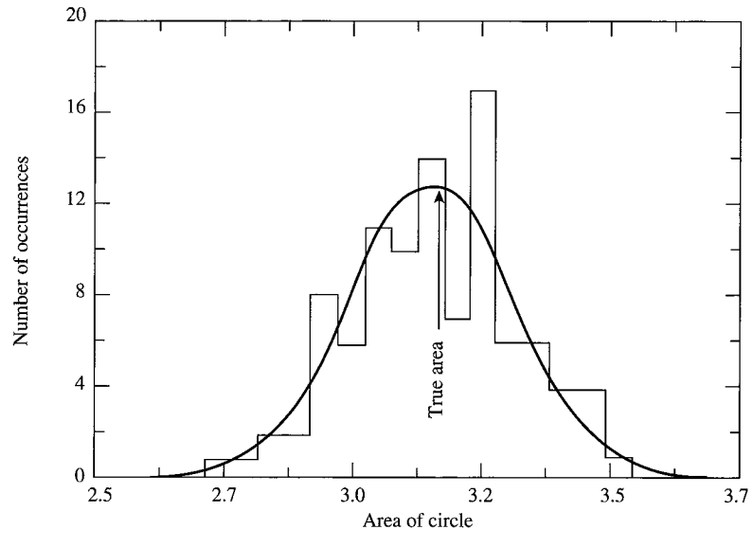


FIGURE 5.2
Histogram of the circle area estimates obtained in 100 independent Monte Carlo runs, each with 100 pairs of random numbers. The Gaussian curve was calculated from the mean $A = 3.127$ and standard deviation $\sigma = 0.156$ of the 100 estimated areas.

Carlo technique is invaluable, with its straightforward sampling and its relatively simple determination of the uncertainties.

5.2 RANDOM NUMBERS

A successful Monte Carlo calculation requires a reliable set of *random numbers*, but truly random numbers for use in calculations are hard to obtain. One might think of a scheme based upon measuring the times between cosmic ray hits in a detector, or on some physical process such as the generation of noise in an electronic circuit. Such numbers would be random in the sense that it would be impossible to predict the value of the next number from previous numbers but they are hardly convenient to use in extended calculations, and some might not have the necessary uniformity required for a Monte Carlo calculation.

In fact, it is generally preferable to use *pseudorandom numbers*, numbers generated by a computer algorithm designed to produce a sequence of apparently uncorrelated numbers that are uniformly distributed over a predefined range. In addition to the convenience of being able to generate these numbers within the Monte Carlo program itself, pseudorandom numbers have another important advantage over truly random numbers for Monte Carlo calculations. A Monte Carlo program may use a great many random numbers, and the path of the calculation through the program will depend on the numbers chosen in each run. With truly random numbers, every run of a Monte Carlo calculation would follow a different path and produce different results. Such a program would be very difficult to debug.

With pseudorandom numbers, we can repeat a calculation with the same sequence of numbers, and search for any particular problems that may be hidden in the code.

There are other advantages too. If we are studying the sensitivity of a calculation to variations in a selected parameter, we can reduce the variance of the *difference* between results calculated with two trial values of the parameter by using the same random number sequence for those parts of the calculation that are independent of the parameter in question. Finally, a pseudorandom number generator can be written to be *portable*; that is, the sequence of numbers produced by the algorithm is independent of computer hardware and language, so that a given program will produce the same results when run on different computers. In view of these advantages and the fact that we rarely, if ever, encounter situations where truly random numbers are required, we shall henceforth use the term *random numbers* to denote *pseudorandom numbers.*

In general, our random number generator must satisfy the following basic criteria:

1. The distribution of the numbers should be uniform within a specified range and should satisfy statistical tests for randomness, such as lack of predictability and of correlations among neighboring numbers.
2. The calculation should produce a large number of unique numbers before repeating the cycle.
3. The calculation should be very fast.

A simple multiplication method is often used to generate random numbers, or *uniform deviates*, as they are often called. An integer starting value or *seed* r_0 and two integer constants are chosen. Successive random numbers are derived from the recursion relation

$$r_{i+1} = (a \times r_i) \bmod m \tag{5.3}$$

where the mod operation corresponds to dividing the product in parentheses by the integer m to obtain the remainder. With appropriate choices of constants a and m, we can obtain a finite sequence of numbers that appear to be randomly selected between 1 and $m - 1$. The length of the sequence is determined by the choice of constants and is limited by the computer word size. For example, if we choose $m = 37$ and $a = 5$, Equation (5.3) gives us the cycle of 36 nicely mixed up numbers, listed in Table 5.1. Random number generators included with computer languages are often based on some variation of this multiplication technique. Careful and thorough statistical studies must be made to be sure that an untested random number generator produces an acceptable sequence of numbers.

Because the numbers generated by Equation (5.3) are not truly random, we might worry that our calculations are affected by hidden correlations in successively generated numbers. We can improve the randomness of our sample by *shuffling* the numbers. We generate two sequences of numbers with different generators a and m; one sequence is stored in an array and a number from the second sequence is used as an index to select numbers from the first sequence. For large programs that employ many random numbers, this method is limited by storage space, although *local shuffling* within a block of random numbers can be used.

TABLE 5.1
Pseudorandom numbers

i	r_i	i	r_i	i	r_i	i	r_i
1	1	10	6	19	36	28	31
2	5	11	30	20	32	29	7
3	25	12	2	21	12	30	35
4	14	13	10	22	23	31	27
5	33	14	13	23	4	32	24
6	17	15	28	24	20	33	9
7	11	16	29	25	26	34	8
8	18	17	34	26	19	35	3
9	16	18	22	27	21	36	15

Note: The generating equation is $r_{i+1} = (a \times r_i) \bmod m$, with $a = 5$ and $m = 37$. The cycle repeats $a_{37} = a_1$, $a_{38} = a_2$, and so forth.

Even a modest Monte Carlo program can require many random numbers and, to assure the statistical significance of results, we must be certain that the calculation does not use more than the maximum number generated by the algorithm before the sequence repeats. The sample generator of Equation (5.3) cannot produce more than $m - 1$ different values of r_i. The actual cycle length may be less than this range, depending on the choice of constants. The cycle length can be increased by employing two or more independent sequences such that the resulting cycle length is proportional to the product of the lengths of the component cycles.

A generator developed by Wichmann and Hill,[1] based on a simple linear combination of numbers from three independent sequences, is said to have a very long cycle ($\sim 7 \times 10^{12}$) and appears to be well tested. Because the algorithm uses three seeds, it is a little longer and slower than one- or two-seed algorithms, but its long repeat cycle, portability, and lack of correlations seem to make it a convenient, worry-free generator for most purposes. The algorithm is listed in Appendix E.

Although the fact that pseudorandom number generators always produce the same sequences of numbers from the same seeds is an advantage in program debugging, it may be a disadvantage in production running. For example, a simulation program developed for use as a science museum display could be very uninteresting if it repeated the same sequence of events every time it was run. If unpredictable seeds are required, they can easily be derived from the least counts of the computer clock. Commercial routines often include such a method of randomizing the starting seeds. On the other hand, if we wish to run a simulation program several times and to combine the results of the several different runs, the safest method to assure the statistical independence of the separate runs is to record the last values of the seeds at the end of each run and use these as starting seeds for the next run.

A thorough discussion of random number generation and of the Monte Carlo technique is given in Knuth (1981).

[1]The authors include a thorough and very useful discussion of the tests applied to a random number sequence, and of the development and testing of the published algorithm.

Warning

If you are using random numbers provided in commercial programs such as spread sheets or even scientific data analysis programs, you should always check the random number distributions for correlations, and make sure that the function behaves as advertised. For example, in early versions of one very popular scientific data analysis program, the choice of seed had no effect on the numbers produced by the random number routine.

5.3 RANDOM NUMBERS FROM PROBABILITY DISTRIBUTIONS

Transformation Method

Most number generators scale their output to provide real numbers uniformly distributed between 0 and 1. In general, however, we require numbers drawn from specific probability distributions. Let us define uniform deviates $p(r)$ drawn from a standard probability density distribution that is uniform between $r = 0$ and $r = 1$:

$$p(r) = \begin{cases} 1 & \text{for } 0 \le r < 1 \\ 0 & \text{otherwise} \end{cases} \tag{5.4}$$

The distribution is *normalized* so that

$$\int_{-\infty}^{\infty} p(r)\, dr = \int_{0}^{1} 1\, dr = 1 \tag{5.5}$$

We shall refer to $p(r)$ as the *uniform distribution.*

Suppose that we require random deviates from a different normalized probability density distribution $P(r)$, which is defined to be uniform between $x = -1$ and 1; that is, the distribution

$$P(x) = \begin{cases} \frac{1}{2} & \text{for } -1 \le x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{5.6}$$

If we choose a random deviate r between 0 and 1 from the uniform distribution of Equation (5.4), it is obvious that we can calculate another random deviate x as a function of r:

$$x = f(r) = 2r - 1 \tag{5.7}$$

which will be uniformly distributed between -1 and $+1$. This is an example of a simple linear transformation.

To pick a random sample x from the distribution Equation (5.6), we started with a random deviate r drawn from the uniform distribution of Equation (5.4) and found a function $f(r)$ that gave the required relation between x and r. Let us find a general relation for obtaining a random deviate x from any probability density distribution $P(x)$, in terms of the random deviate r drawn from the uniform probability distribution $p(r)$.

Conservation of probability requires that the intervals Δr and Δx be related by the following expression

$$|p(r)\,\Delta r| = |P(x)\Delta x| \tag{5.8}$$

and, therefore, we can write

$$\int_{r=-\infty}^{r} p(r)\,dr = \int_{x=-\infty}^{x} P(x)\,dx \quad \text{or} \quad \int_{r=0}^{r} 1\,dr = \int_{x=-\infty}^{x} P(x)\,dx \tag{5.9}$$

which gives the general result

$$r = \int_{x=-\infty}^{x} P(x)\,dx \tag{5.10}$$

Thus, to find x, selected randomly from the probability distribution $P(x)$, we generate a random number r from the uniform distribution and find the value of the limit x that satisfies the integral equation (5.10).

Example 5.2. Consider the distribution described by the equation

$$p(x) = \begin{cases} A(1 + ax^2) & \text{for } -1 \le x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

where $P(x)$ is positive or zero everywhere within the specified range, and the normalizing constant A is chosen so that

$$\int_{-1}^{1} P(x)\,dx = 1 \tag{5.12}$$

We have

$$r = \int_{-\infty}^{x} P(x)\,dx = \int_{-1}^{x} A(1 + ax^2)\,dx \tag{5.13}$$

which gives

$$r = A(x + ax^3/3 + 1 + a/3) \tag{5.14}$$

and therefore, to find x we must solve the third-degree equation (5.14).

The procedure we have described is referred to as the *transformation method* of generating random deviates from probability distributions. In general, neither the integral equation (5.13) nor the solution of the resulting equation (5.14) can be obtained analytically, so numerical calculations are necessary.

The following steps are required to generate random deviates from a specific probability distribution by the transformation method with a numerical integration:

1. Decide on the range of x. Some probability density functions are defined in a finite range, as in Equation (5.6); others, such as the Gaussian function, extend to infinity. For numerical calculations, reasonable finite limits must be set on the range of the variable.

2. Normalize the probability function. If it is necessary to impose limits on the range of the variable x, then the function must be renormalized to assure that the integral is unity over the newly defined range. The normalization integral should be calculated by the same analytical integration or numerical integration routine that is used to find y.
3. Generate a random variable r drawn from the uniform distribution $p(r)$.
4. Integrate the normalized probability function $P(x)$ from negative infinity (or its defined lower limit) to the value $x = x$, where x satisfies Equation (5.10).

Because the Monte Carlo method usually requires the generation of large numbers of individual events, it is essential to have available fast numerical interpolation and integration routines. To reduce computing time, it is often efficient to set up tables of repeatedly used solutions or integrals within the initializing section of a Monte Carlo program. For example, to pick a random deviate x from the distribution of Equation (5.11), we could do the integral of Equation (5.13) numerically at the beginning of our program, and set up a table of values of r versus x. Then, when we require a random number from the distribution, we generate a random number r and search the table for the corresponding value of x. In general, the search should be followed by an interpolation within the table (see Appendix A.1.) to avoid introducing excessive graininess into the resulting distribution. It would be even more convenient, but a little trickier, to produce a table of x versus r, so that the required value of x could be obtained from an index derived from r. In all cases of precalculated tables, it is important to consider the resolution required in the generated variable, because this will determine the intervals at which data must be stored, and therefore the size of the table, and the time required for a search.

Rejection Method

Although the *transformation method* is probably the most useful method for obtaining random deviates drawn from particular distributions, the *rejection method* is often the easiest to use. This is the method that we used in Example 5.1 to find the area of a circle, by generating random numbers uniformly over the surface of the circle and rejecting all except those that fell within the circumference.

Example 5.3. Suppose we wish to obtain random deviates between $x = -1$ and $x = +1$, drawn from the distribution function

$$P(x) = 1 + ax^2 \tag{5.15}$$

which is just the unnormalized distribution of Equation (5.11). To use the rejection method, we begin by generating a random deviate x' uniformly distributed between -1 and $+1$, corresponding to the allowed range of x, and a second random deviate y' uniformly distributed between 0 and $(1 + a)$, corresponding to the allowed range of $P(x)$. We can see that x' and y' must be given by

$$x' = -1 + 2r_i \quad \text{and} \quad y' = (1 + a)r_{i+1} \tag{5.16}$$

where r_i and r_{i+1} are successively generated random values of r drawn from the uniform distribution.

We count an event as a "hit" if the point (x', y') falls between the curve defined by $P(x)$ and the x axis, that is, if $y' < P(x')$, and a "miss" if it falls above the curve. In the limit of a large number of trials, the entire plot, including the area between the curve and the x axis, will be uniformly populated by this operation and our selected samples will be the x coordinates of the "hits," or the values of x', drawn randomly from the distribution $P(x)$. Note that with this method it is not necessary to normalize the distribution to form a true probability function. It is sufficient that the distribution be positive and well behaved within its allowed range.

The advantage of the rejection method over the transformation method is its simplicity. An integration is not required—only the probability function itself must be calculated. A disadvantage of the method is often its low efficiency. In a complex Monte Carlo program only a small fraction of the events may survive the complete calculation to become successful "hits" and the generation and subsequent rejection of so many random numbers may be very time consuming. To reduce this problem, it is advisable to place the strictest possible limits on the random coordinates used to map out the distribution function when using the rejection method.

5.4 SPECIFIC DISTRIBUTIONS

Gaussian Distribution

Almost any Monte Carlo calculation that simulates experimental measurements will require the generation of deviates drawn from a Gaussian distribution, or *Gaussian deviates.* A common application is simulation of measuring uncertainties by *smearing* variables. Fortunately, because of the convenient scaling properties of the Gaussian function, it is only necessary to generate Gaussian deviates from the standard distribution

$$P_G(z)\,dz = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz \tag{5.17}$$

with mean 0 and standard deviation 1, and to scale to different means μ and standard deviations σ by calculating

$$x = \sigma z + \mu \tag{5.18}$$

There are several different ways of obtaining random samples of the variable z from the distribution $P_G(z)$ of Equation (5.17). The two most obvious are the rejection and transformation methods discussed previously. Because the Gaussian function is defined between $-\infty$ and $+\infty$, these methods require that limits be placed on the range of z. For low-statistics calculations in which the Gaussian function is being used to simulate smearing of data caused by measuring errors, a range of $\pm 3\sigma$ should be satisfactory because all but ~0.3% of normally distributed events lie within this range.

Because the Gaussian function cannot be integrated analytically, numerical integrations are required for the transformation method. Decisions must be made on the order of integration and the step size as well as on the limits. A first- or second-order numerical integration (Appendix A.3.) is generally satisfactory, with a linear interpolation to find an approximation to the value of x in Equation (5.10) at the required value of the integral.

An interesting method for generating Gaussian deviates is based on the fact that if we repeatedly calculate the means of groups of numbers drawn randomly from any distribution, the distribution of those means tends to a Gaussian as the number of means increases. Thus, if we calculate many times the sums of N uniform deviates, drawn from the uniform distribution, we should expect the sums to fall into a truncated Gaussian distribution, bounded by 0 and N, with mean value $N/2$. If we generate N values of r from the distribution of Equation (5.4) and calculate

$$r_G = \sum_{i=1}^{N} r_i - N/2 \tag{5.19}$$

the variable r_G will be drawn from an approximately Gaussian distribution with mean $\mu = 0$ and standard deviation $\sigma = \sqrt{N/12}$. We should note that the maximum range of r_G will be limited to $\mu \pm N/2$ or $\mu \pm \sigma\sqrt{3N}$. For $N = 2$, the sum is a triangle function and as N increases, the distribution quickly takes on a Gaussian-like shape. Values of N as small as $N = 4$ are suitable for low statistics calculations. With $N = 4$, we have $\sigma = \sqrt{1/3} \simeq 0.058$ and the range of r_G from -2 to $+2$ corresponds to $\mu \pm \sigma\sqrt{12}$ or $\mu \pm 3.46\sigma$. If a better approximation to the Gaussian function is require and calculation time is not a problem, $N = 12$ is particularly convenient because the resulting variance and standard deviation are unity.

A particularly elegant method for obtaining random numbers drawn from the Gaussian distribution was suggested by Box and Müller (1958). This method makes use of the fact that, although the simple transformation method requires an integration of the Gaussian function, it is possible to find a function that generates the two-dimensional Gaussian distribution,

$$f(z_1, z_2) = \frac{1}{2\pi} \exp\left(-\frac{(z_1^2 + z_2^2)}{2}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_1^2}{2}\right) \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_2^2}{2}\right) \tag{5.20}$$

From this equation, the authors obtained expressions that generate two Gaussian deviates, z_1 and z_2, from two uniform deviates, r_1 and r_2:

$$\begin{aligned} z_1 &= \sqrt{-2 \ln r_1} \cos 2\pi r_2 \\ z_2 &= \sqrt{-2 \ln r_1} \sin 2\pi r_2 \end{aligned} \tag{5.21}$$

Example 5.4. A uniform 10-cm long rod has one end held at 0°C and the other at 100°C so that the temperature along the rod is expected to vary linearly from 0° to 100°C. Let us attempt to simulate data that would be obtained by measuring the temperature at regular intervals along the rod. We shall assume that the parent population is described by the equation

$$T = a_0 + b_0 x \tag{5.22}$$

with $a_0 = 0$°C and $b_0 = 10$°C/cm, and that 10 measurements are made at 1-cm intervals from $x = 0.5$ to $x = 9.5$ cm, with negligible uncertainties in x_i and uniform measuring uncertainties in T_i of $\sigma_T = 1.0$°C.

Example 5.4 illustrates a common Monte Carlo technique: simulating the effects of measuring uncertainties by smearing data points. If a particular variable has a mean value T_i, with uncertainties σ_i and Gaussian uncertainties are assumed, then we obtain the smeared value of T_i from the relation

$$T_i' = T_i + \sigma_i r_i \tag{5.23}$$

where r_i is a random variable drawn from the standard Gaussian distribution with mean 0 and standard deviation 1. The calculation is equivalent to drawing the random variable T_i' directly from a Gaussian distribution with mean T_i and standard deviation σ_i.

Program 5.1. HOTROD (Appendix E) A simple Monte Carlo calculation to simulate the measurements described in Example 5.4. The program uses routines in the program unit MONTELIB.

Program 5.3. MONTELIB (Appendix E) Some useful Monte Carlo routines.

The data generated by the program HOTROD are shown in Table 5.2, with values of T_i for the parent population, predicted by Equation (5.22), and of T_i for the sample population, calculated from Equation (5.23) for various values of x_i. Note that, as we should expect, the modified values of T are scattered about the values calculated from Equation (5.22).

Choice of a Method

Which of these methods for generating samples from the Gaussian probability distribution is the best? The answer depends on need and circumstance. For general use it is convenient to keep a version of the Box-Müller method in your program library.

TABLE 5.2
Simulated temperature versus position data for a 10-cm rod held at $T = 0°C$ at $x = 0.0$ cm and at $T = 100°C$ at $x = 10.0$ cm

i	x_i (cm)	T_i (°C)	T_i' (°C)
1	0.5	5.00	4.71
2	1.5	15.00	15.43
3	2.5	25.00	23.24
4	3.5	35.00	35.77
5	4.5	45.00	45.39
6	5.5	55.00	52.26
7	6.5	65.00	65.71
8	7.5	75.00	76.96
9	8.5	85.00	85.97
10	9.5	95.00	93.77

Note: A uniform temperature gradient was assumed. The uncertainty in the measurement of T was assumed to be $\sigma_T = 1.0$ °C.

This routine produces a continuous range of samples limited only by the computer word size. For high-precision work, however, we should be aware that subtle correlations between adjacent uniform deviates have been shown to distort the tails of the Gaussian distribution of these numbers. If highest speed is essential, then the transformation method with a precalculated table of the integral and some pointers for quick access to the table should be the choice. This method requires making decisions on the range and resolution of the generated variable and some extra programming to create and access the integral table, but the lookup method can be very fast. Finally, if you are stranded on a desert island with only your laptop computer and have an urgent need for random selections from a Gaussian distribution, the method of summing N random numbers is sufficiently simple that you should be able to write and debug the routine in a few minutes, provided you can remember that the magic number is $N = 12$ for a variance of 1.

Poisson Distribution

Poisson statistics are important in most Monte Carlo calculations, but they are usually implied rather than calculated explicitly. Nevertheless, we sometimes wish to generate data that are distributed according to the Poisson function, and application of the transformation method to the problem is particularly simple and instructive. To find an integer x drawn from the Poisson distribution with mean μ, a *Poisson deviate*, we generate a random variable r from the uniform distribution, replace the integral of Equation (5.10) by the sum

$$r = \sum_{x=0}^{x} P_P(x; \mu) = \sum_{x=0}^{x} \frac{\mu^x}{x!} e^{-\mu} \tag{5.24}$$

and solve Equation (5.24) for x.

Although the Poisson function does not have the convenient scaling properties of the Gaussian function, and thus different calculations are required for each value of the mean μ, very few calculations are actually needed because we are interested in this distribution only at small values of μ, say $\mu \leq 16$, and only at integral values of the argument x. At larger values of μ, the Poisson distribution becomes indistinguishable from the Gaussian and it is generally more convenient to employ the Gaussian function in calculations.

Example 5.5. An instructor is preparing an exercise on Poisson statistics for his class. He plans to provide each student with a simulated data set corresponding to 200 Geiger counter measurements of cosmic ray flux recorded in 10-s intervals with an assumed mean counting rate of 8.4 counts per interval. The data will correspond to the number of counts recorded in each 10-s interval.

Students will be asked to make histograms of their individual data samples, find the means and standard deviations of the data, and compare their distributions with the predictions of Gaussian and Poisson probability functions.

For each student, a set of values of x is generated from Equation (5.24) with $\mu = 8.4$ and 200 different random numbers. The transformation method is used with a precalculated table of sums so that the value of x associated with each value

of r can be selected by a simple search. To assure that each student's data set is independent, either all sets are generated in a single computer run or else the random number seeds are saved at the end of each run and used to start the next run.

Program 5.2. POISDECAY (Appendix E) Generates 200 random variables drawn from the Poisson probability distribution with mean $\mu = 8.4$ to illustrate Example 5.5. The program uses routines in the program unit MONTELIB.

The program calls the function POISSONDEVIATE with second argument INIT = TRUE to set up a table of sums of $P_P(i; \mu)$ from $i = 0$ to n indexed by n; that is, to form the array

$$S_n = \sum_{i=0}^{n} P_P(i; \mu) \quad \text{for } n = 1, 2, \ldots, n_{\max} \tag{5.25}$$

so that

$$S_n = S_{n-1} + P_P(n; \mu) \quad \text{with } S_0 = P_P(0; \mu) = e^{-\mu} \tag{5.26}$$

where $n_{\max} = N + 8\sqrt{\mu}$ is selected as a reasonable upper range for the Poisson curve.

For each event, the program calls POISSONDEVIATE with second argument INIT = FALSE to select a value from the table. The routine POISSONDEVIATE generates a random number r from the uniform distribution and searches the table beginning at S_0, to find the value of n for which $S_h \geq r$. The value of n at which this occurs is the desired random sample from the Poisson distribution. As the samples are generated they are entered in a histogram by calls to the routine HISTOGRAM.

A histogram of 200 variables drawn from the Poisson distribution Program 5.2 is shown in Figure 5.3 with the parent distribution represented as a solid curve (although it is, of course, not defined between integer values of the abscissa). The values of the Poisson function, calculated by the routine POISSONRECUR, and the sums, calculated by the routine POISSONDEVIATE, for $\mu = 8.4$ and for n ranging from 0 to 31, are displayed in Table 5.3.

We note that with the precalculated table it is only necessary to increment a counter a few times and compare two real numbers to obtain each random variable, whereas, without the table, it would have been necessary to calculate the Poisson function several times for each generated sample, in addition to comparing the two real numbers.

Exponential Distribution

If the Monte Carlo problem includes the generation of unstable states, random numbers drawn from an exponential distribution will be needed. Here the transformation method is clearly the method of choice because the integral equation (5.10) and resultant equation can be solved analytically.

Example 5.6. Consider an experiment to study the decay rate of a radioactive source with estimated mean life of τ seconds. The experiment involves collecting counts over successive time intervals Δt with a Geiger counter and scaler combination and plotting the number of counts in each interval against the mean interval time.

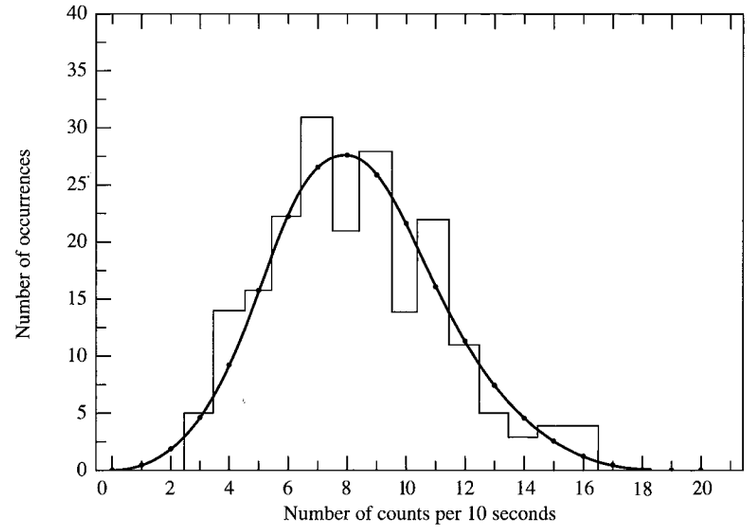


FIGURE 5.3
Histogram of 200 random variables generated by Program 5.3 from the Poisson distribution with mean $\mu = 8.4$.

TABLE 5.3
Poisson probability $P_P(i; \mu)$ and summed probability $S_i = \sum_{i=0}^{n} P_P(i; \mu)$ for $\mu = 8.4$

n	$P_P(n; \mu)$	S_n	n	$P_P(n; \mu)$	S_n
0	0.0002248673	0.0002248673	16	0.0066035175	0.9940781736
1	0.0018888855	0.0021137528	17	0.0032629145	0.9973410882
2	0.0079333192	0.0100470720	18	0.0015226935	0.9988637816
3	0.0222132938	0.0322603658	19	0.0006731908	0.9995369724
4	0.0466479169	0.0789082827	20	0.0002827401	0.9998197126
5	0.0783685004	0.1572767830	21	0.0001130961	0.9999328086
6	0.1097159005	0.2669926835	22	0.0000431821	0.9999759908
7	0.1316590806	0.3986517641	23	0.0000157709	0.9999917616
8	0.1382420346	0.5368937988	24	0.0000055198	0.9999972814
9	0.1290258990	0.6659196977	25	0.0000018547	0.9999991361
10	0.1083817551	0.7743014529	26	0.0000005992	0.9999997353
11	0.0827642494	0.8570657023	27	0.0000001864	0.9999999217
12	0.0579349746	0.9150006768	28	0.0000000559	0.9999999776
13	0.0374349066	0.9524355835	29	0.0000000162	0.9999999938
14	0.0224609440	0.9748965275	30	0.0000000045	0.9999999983
15	0.0125781286	0.9874746561	31	0.0000000012	1.0000000000

Note: The summation was terminated arbitrarily at $n \simeq \mu + 8\sqrt{\mu} \simeq 31$, and $P_P(31; \mu)$ was set to 1.

We wish to simulate this experiment with a Monte Carlo calculation. The normalized probability density function for obtaining a count at time t from an exponential distribution with mean life τ is given by

$$P_e(t;\tau) = \begin{cases} 0 & \text{for } t < 0 \\ \dfrac{e^{-t/\tau}}{\tau} & \text{for } t \geq 0 \end{cases} \tag{5.27}$$

We can obtain an expression for random samples t_i from this distribution by applying Equation (5.10) to obtain

$$t_i = -\tau \ln r_i \tag{5.28}$$

Thus, to obtain each value of t_i, we find a random number from the uniform distribution and calculate t_i from Equation (5.28).

Let us consider a second method of generating a histogram of data for this example, a method that is much more efficient, but that severely limits any later treatment of the data.

We can calculate the fraction of events that the parent distribution predicts would fall into each of the Δt wide histogram bins from the equation

$$\Delta N'(t) = \int_{t-d}^{t+d} \frac{e^{-x/\tau}}{\tau}\, dx = e^{-t/\tau}\Big|_{t-d}^{t+d} \simeq \frac{\Delta t}{\tau}\, e^{-t/\tau} \tag{5.29}$$

where we have written $d = \Delta t/2$. The effect of the statistical errors is to smear each of these calculated values in a way consistent with the Poisson distribution with mean $\mu = \Delta N_i'$. For small values of $\Delta N_i'$ we find the smeared value ΔN_i directly from Equation (5.24):

$$r = \sum_{x=0}^{\Delta N} P_p(x; \Delta N') \tag{5.30}$$

For larger values of $\Delta N_i'$ calculation with the Poisson equation would be too tedious, but we can use Gaussian smearing as in Example 5.4 with $\sigma_i = \sqrt{\mu}$. Note that the Poisson equation *must* be used for bins with low statistics to assure a positive number of counts in each bin. (A reminder: The overall distribution of events in this example is exponential; the expected distribution of events *in each individual bin* follows the Poisson distribution, as discussed in Section 4.3.)

Although these two methods of generating a data set or histogram produce equivalent statistical results for Example 5.6, they differ in important details. The full Monte Carlo method required generating individual "events" that can be recorded and studied. For example, we could check the statistical behavior of the data by subdividing the sample into several smaller groups. We could also investigate the effect of decreasing as well as increasing the binning intervals Δt. Finally, if we should wish to expand the study, perhaps to consider experimental geometry and detector efficiency, the full Monte Carlo method will allow that. The smearing method, on the other hand, produces only the ten numbers, representing the counts in the ten bins. Aside from merging the bins, we have no control over the data for future calculations. It is strictly a fast, "one-shot" procedure with a specific limited aim.

Example 5.7. Consider an experiment to determine the mean life of an elementary particle, the short-lived K_S^0 meson (which we shall refer to as the kaon), from

measurements of the decay in flight of many such particles. In principle, we can determine the mean life τ by measuring the distribution of decay times, fitting the probability density function of Equation (5.27) to the data and solving for τ. In practice, we must make corrections for biases resulting from detection inefficiencies, including those associated with the finite sizes of our detectors. We can use a Monte Carlo calculation to estimate these biases and enable us to apply the appropriate correction.

The experimental arrangement is sketched in Figure 5.4. A high-energy charged particle p_i interacts in the target at the production vertex V_1 to produce several charged and neutral secondary particles, including a neutral kaon. The kaon travels a distance L before decaying into two pions, π_1 and π_2, at the decay vertex V_2. We determine the coordinates of the production vertex by measuring in the *production vertex detector* the trajectories of charged particles that are produced with the kaon, and tracing back these trajectories to their intersection point in the target. Similarly, we determine the coordinates of the decay vertex by measuring in the *decay vertex detector* the trajectories of the two charged pions from the kaon decay, and tracing these trajectories back to their intersection point, V_2. (The trajectories of neutral particles are much more difficult to measure than those of the charged particles.) We calculate the momentum of the neutral kaon from measurements of the momentum vectors of its two decay products, π_1 and π_2.

The geometry of the detector plays a critical role in the analysis of the data. We can make useful measurements only on events in which the trajectories of the charged particles can be measured in the vertex detectors. To assure precise measurements of the secondary tracks from the decay of the kaon, we define a *fiducial* region in which the decay must occur. The dashed rectangle on Figure 5.4 indicates the *fiducial* region with its limits d_1 and d_2 along the *x-axis*. With these limits, very short-lived and long-lived particles will be eliminated from the data sample, introducing a bias into the determination of the mean life.

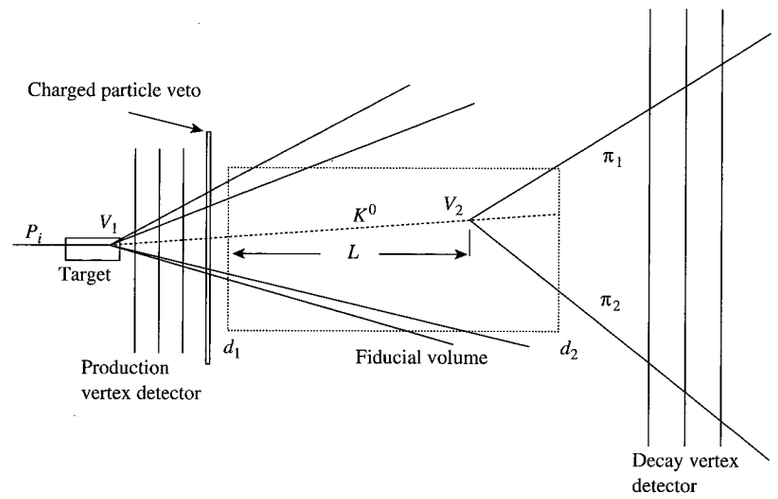


FIGURE 5.4
Experimental arrangement to measure the lifetime of an elementary particle.

In a Monte Carlo study of these biases, we could take the following steps to simulate measurements of decaying kaons:

1. Generate the production vertex coordinates and kaon momentum vector P from the known cross section for kaon production in the interaction of the incident and target particles.
2. Consider the efficiency of the production detector. If the detector successfully records charged particles produced in the initial interaction, proceed to step 3; if not, mark the event a failure and go to step 8.
3. Apply Equation (5.28) to find the time of flight (or lifetime) T of each individual kaon in its own rest frame. Use the current best-known value for the mean life τ.
4. Apply the Lorentz transformation to T to find the lifetime T' in the laboratory system.
5. Calculate the range r of the kaon in the laboratory and from this, the coordinate of the decay point.
6. Check that the kaon decays within the fiducial volume. If so, proceed to step 7; otherwise, mark the event a failure and go to step 8.
7. In the rest frame of the kaon, generate the pair of pion vectors. Transform to the laboratory system and check whether or not both particles can be detected in the decay vertex detector. If they can be detected, mark the event a success; if not, mark the event a failure.
8. Record details of the event and return to step 1 to generate a new event, or terminate if the desired number of events has been generated.

Program 5.4. KDECAY (website) Illustration of Example 5.7.

For this sample program, we simplify the problem by treating it in two dimensions and simplify or skip some of the steps as noted below.

1. Assume that each kaon is produced in the plane illustrated in Figure 5.4 and travels along the x-axis. Generate a vertex x-coordinate x_0 and the magnitude of the kaon's momentum P from suitable Gaussian distributions.
2. Skip
3. Find the lifetime T of the kaon in its own rest frame from the published value of the kaon mean life τ and Equation (5.28).
4. Apply the Lorentz transformation to T to find the lifetime T' in the laboratory system:

$$T' = \gamma T_{cm}, \text{ where } \gamma = 1/\sqrt{1-\beta^2} \text{ and } \beta = v/c$$

where v is the velocity of the kaon in the laboratory and c is the velocity of light.

5. Calculate the range r and decay point x_d:

$$r = \beta c T' \text{ and } x_d = x_0 + r$$

6. Check that the decay is within the fiducial area, that is, that

$$d_1 \leq x_d < d_2$$

If it is not, mark the event as a failure; otherwise, mark the event as a success.

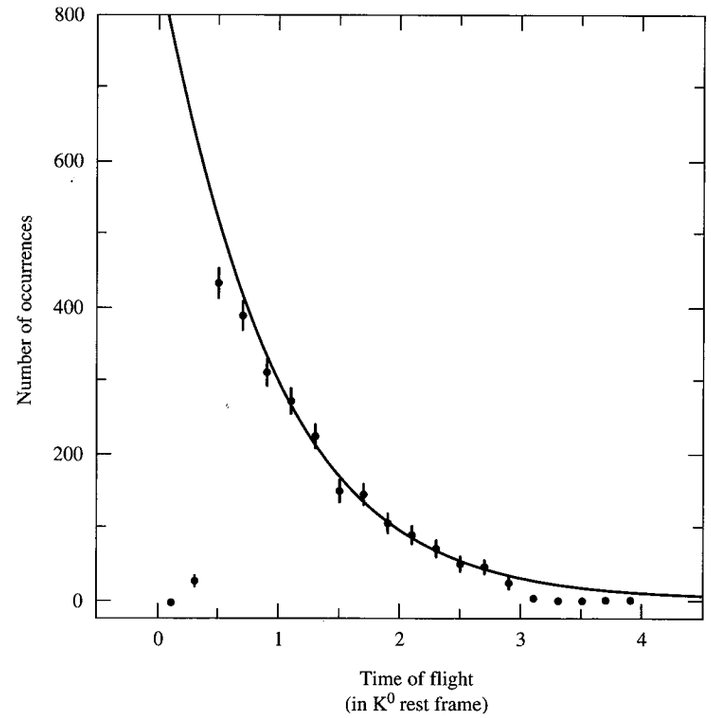


FIGURE 5.5
Distribution of times of flight (in units of 10^{-10} s) of 2355 successful K^0 decays from a total sample of 4000 generated events. The curve shows the predicted exponential distribution of the total 4000-event sample.

7. Skip this step.
8. Increment the event counters and record only successful events. If the desired number of events has been generated, terminate the calculation; otherwise, go to step 1 to begin generating the next event.

The properties of the two Gaussians and the other constants of the calculation are listed in Table 5.4. Note that we must use as input to our program a reasonable value of the kaon mean life, the quantity that we are attempting to measure. If the quantity had been only poorly measured previously, or perhaps not at all, it might be necessary to run the Monte Carlo program with several different trial values of τ, bracketing the expected value.

For this example, we generated 4000 events of which 2355 passed the fiducial cut. Figure 5.5 shows the distribution of the times of flight T (or lifetimes) in the rest frame of the kaon for successful events. The curve shows the expected distribution of the times of flight if no events had been rejected. We obtain the efficiency of the system as a function of the time of flight T by calculating the ratio of the number $N'(T)$ of successful events to the total number $N(T)$ generated

TABLE 5.4
Constants used in the Monte Carlo generation of Program 5.3

TauKaon (K_0 mean life)	0.894×10^{-10} s
MassKaon (K_0 mass)	497.7 Mev/c^2
*d*1 (Lower limit of fiducial range)	10 m
*d*2 (Upper limit of fiducial range)	40 m
xMean (mean coordinate of the production vertex, V_1)	5.00 cm
xSig (Standard deviation of production vertex)	0.50 cm
pMean (mean K_0 momentum)	2000 MeV/c
pSig (Standard deviation of K_0 momentum)	100 MeV/c
c (velocity of light)	3.00×10^{10} cm/s

$$\epsilon(T) = N'(T)/N(T) \tag{5.31}$$

We note that there are large losses of events at short times, below about $T = 0.5 \times 10^{-10}$ s, caused by the gap between the production vertex V_1 and the beginning of the fiducial region d_1, and smaller, but significant losses at long times of events that decayed beyond the end of the fiducial region, d_2.

To correct data obtained in an actual experiment and distributed as $N_{exp}(T)$, we should first run the Monte Carlo to generate sufficient numbers of events so that the uncertainties in the $N'(T)$ are negligible compared to the uncertainties in the experimental data sample. We should then select a continuous region of our data sample where the efficiency is reasonably good (and definitely not zero!) and correct the measurements by scaling $N_{exp}(T)$ by $1/\epsilon(T)$. Note that the statistical uncertainties in the measured data must also be scaled, so there is little point in including data from very low-efficiency regions of the sample. We can then obtain our estimate of the mean life of the kaon from a least-squares fit of Equation (5.27) to the corrected data. (A reminder: Although the overall distribution of events in this example is exponential, the expected distribution of events *in each individual bin* follows the Poisson distribution, as discussed in Section 4.4.)

A more detailed discussion of analysis techniques for this experiment is in Chapter 10.

5.5 EFFICIENT MONTE CARLO GENERATION

Because the relative error in a result calculated by the Monte Carlo method is inversely proportional to the square root of the number of *successful* events generated, it is important, especially for a long calculation, to have the highest possible program efficiency. Rejected events do not improve the statistical accuracy and every effort should be made to reduce the time spent on calculations that lead to "misses" rather than "hits." There are several ways to improve generation efficiency:

1. Don't be a purist. The Monte Carlo method is basically a way of doing complicated multidimensional integrals. If you can save time by doing part of the problem by analytic methods, do so.

2. Program carefully. Do not repeat calculations if the results can be saved for later use.
3. If possible, test the low-yield sections of the simulation early and cut out as soon as a "miss" occurs. Except for particular loss studies, it is usually not profitable to follow the calculation of an event that is known to end in failure.
4. Try to reduce the variance of the results by limiting ranges wherever possible. One application of this technique can be illustrated in Example 5.1, where the area of a circle of radius r_c is calculated by inscribing it within a square. Making the side of the square larger than the diameter of the circle would be wasteful and would increase the variance of the area determination.
5. When repeating a calculation to find the effects of varying a parameter, consider setting up the program in such a way that the identical sequence of random numbers is repeated throughout the calculation, except for calculations specifically associated with the change. This technique will not improve the variance of the *overall* calculation, but will reduce the variance of the *difference* of results from two calculations.
6. Inspect each probability function carefully before beginning a calculation and estimate the resolution and detail that will be required in the calculation. If a distribution has fine structure, try to determine whether or not such structure is of interest and must be preserved. If necessary, consider breaking the calculations into separate regions and varying the sampling sensitivity as appropriate for each region.
7. Be critical. Examine your generated variables to see that they fall within the expected ranges and follow expected distributions. In a large program, errors that affect the results in subtle ways may be buried within the program and be very difficult to detect. The only way to prevent problems is to make detailed checks at every stage of the program.

SUMMARY

Pseudorandom numbers: Numbers created by a computer algorithm such that successive numbers appear to be uncorrelated with previous numbers. They are referred to as *random numbers* or *random deviates.*
Uniform deviates: Pseudorandom numbers that are uniformly distributed between 0 and 1:

$$p(r) = \begin{cases} 1 & \text{for } 0 \le r < 1 \\ 0 & \text{otherwise} \end{cases}$$

Normalized distribution: A distribution that is scaled so that its integral over a specified range is equal to unity.
Transformation integral: Transforms the variable r drawn randomly from the uniform distribution into a variable x drawn randomly from the distribution $P(x)$:

$$\int_{r=0}^{r} 1\, dr = \int_{x=-\infty}^{x} P(x)\, dx$$

Rejection method: A method of generating random numbers drawn from particular distributions by rejecting those that fall outside the geometrical limits of the specified distribution.
Gaussian deviate: Random number drawn from a Gaussian distribution.
Quick Gaussian deviate: The sum of N random numbers is approximately Gaussian distributed with $\mu = N/2$ and $\sigma = \sqrt{N/12}$. Choose $N = 12$ and calculate $r_G = \Sigma r_i - N/2$ to obtain r_G drawn from
the standard Gaussian distribution with $\mu = 0$ and $\sigma = 1$.
Box-Müller method for Gaussian deviates: Select r_1 and r_2 from the uniform distribution and calculate

$$z_1 = \sqrt{-2 \ln r_1} \cos 2\pi r_2 \quad \text{and} \quad z_2 = \sqrt{-2 \ln r_1} \sin 2\pi r_2$$

to obtain z_1 and z_2 drawn from the standard Gaussian distribution.
Data smearing: Method for adding random variations to calculations to simulate the effects of finite measuring errors, $T_i' = T_i + \sigma_i r_i$.
Random numbers from the exponential distribution: To obtain a random number t_i drawn from the exponential distribution, calculate $t_i = -\tau \ln r_i$ from a random deviate r_i.

EXERCISES

5.1. Write a computer program that incorporates the Wichmann and Hill pseudorandom number generator and use it to generate 100 random numbers beginning with seeds $s_1 = 13$, $s_2 = 117$, and $s_3 = 2019$. Make a histogram of the numbers and draw a line representing the expected number of events in each bin. Calculate χ^2 for the agreement between the expected and generated number of events and find the associated probability.

5.2. (*a*) Generate 1000 random numbers uniformly distributed between $-\pi$ and $+\pi$.
(*b*) Generate 1000 random numbers between $x = 0$ and 1, distributed according to the distribution function $P(x) = (5x + 3)$. Use the transformation method with an analytic integration.
(*c*) Find the mean and standard deviation of each distribution and compare them to the predicted values.
(*d*) Make a 20-bin histogram of each distribution and plot on each the predicted distribution.
(*e*) Calculate χ^2 to compare each generated distribution to its parent distribution.

5.3. Write a general routine to generate random integers drawn from the binomial distribution by the transformation method. Use the routine to generate 1000 events corresponding to the distribution of heads or tails when a coin is tossed 50 times. Plot your results and compare them to the direct prediction of Equation (2.4).

5.4. Write a Monte Carlo routine to simulate 200 rolls of a pair of dice and find the frequency of occurrences of each possible sum. Plot a histogram of the occurrences with statistical error bars and plot the prediction of the binomial distribution. Calculate χ^2 for the agreement between the prediction and the data, and find the χ^2 probability. Compare your results to the exact probability calculation of Exercise 2.4.

5.5. Make a histogram of 200 random numbers that follow the Gaussian distribution by finding the distribution of the sums of groups of 12 random variates drawn from the uniform distribution. Calculate the mean and standard deviation of the generated numbers and the uncertainty in the mean.

5.6. Generate 1000 random numbers between $x = -3$ and $+3$, distributed according to the Lorentzian distribution with mean $\mu = 0$ and half-width $\Gamma = 1.0$. Use the transformation method with a numerical integration and interpolation. (See Appendix A.1 and A.3.) Make a 20-bin histogram of the generated numbers and plot Lorentzian the curve on the distribution. Calculate χ^2 to compare the generated distribution to the parent distribution.

5.7. Use the transformation method to produce a sequence of 200 random numbers x drawn from the distribution

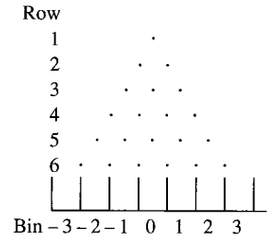
$$P(x) = \sin x \quad \text{for } 0 \le x < \pi$$
$$= 0 \quad \text{elsewhere}$$

Make a histogram of the events and compare it to the expected distribution. Note that the calculation can be done analytically and requires an inverse trigonometric function.

5.8. Use the rejection method to generate 500 random deviates between $x = 0$ and $x = 1$, drawn from the distribution $y(x) = a_1 + a_2 x^2$, with $a_1 = 3.4$ and $a_2 = 12.1$. Find the mean and standard deviation of the generated numbers and compare them to the expected values.

5.9. Write a Monte Carlo program to generate 200 cubes with sides $a = 2.0 \pm 0.1$ cm, $b = 3.0 \pm 0.1$ cm, and $c = 4.0 \pm 0.2$ cm. Plot the distribution of the volumes of the cubes and find the mean volume, the standard deviation of the distribution, and the uncertainty in the mean. Compare the standard deviation of the distribution to the value predicted by the error propagation equation.

5.10. A *Pascal triangle* provides an interesting illustration of the relation between the binomial and Gaussian probability distributions. Assume an arrangement of pins in the form of a triangle as illustrated.



A ball, dropped into the device strikes the top pin and has a 50% probability of striking either of the two pins below it in the next row. The ball bounces down until it reaches the bottom where it is collected in one of the vertical bins.
(*a*) Find a general expression for the probability that a ball will land in a given bin after dropping through N rows of pins.
(*b*) Assume that 512 balls are dropped onto the top pin. Find the number of balls in each bottom bin for a device with three rows of pins above the bins. Repeat for devices with four, five, and six rows of pins.
(*c*) Find the standard deviation of the distribution of balls for each example; that is, assume that the bin number is the independent variable so that $\bar{x} = 0$.
(*d*) Plot histograms of the distribution of the balls with Gaussian curves with the means and standard deviations determined in (*c*).

5.11. Write a Monte Carlo program to simulate the Pascal triangle device described in the previous exercise. Compare the results obtained by the two methods.

CHAPTER 6

LEAST-SQUARES FIT TO A STRAIGHT LINE

6.1 DEPENDENT AND INDEPENDENT VARIABLES

We often wish to determine one characteristic y of an experiment as a function of some other quantity x. That is, instead of making a number of measurements of a single quantity x, we make a series of N measurements of the pair (x_i, y_i), one for each of several values of the index i, which runs from 1 to N. Our object is to find a function $y = y(x)$ that describes the relation between these two measured variables. In this chapter we consider the problem of pairs of variables (x_i, y_i) that are linearly related to one another, and refer to data from two undergraduate laboratory experiments as examples. In the following chapters, we shall discuss methods of finding relationships that are not linear.

Example 6.1. A student is studying electrical currents and potential differences. He has been provided with a 1-m nickel-silver wire mounted on a board, a lead-acid battery, and an analog voltmeter. He connects cells of the battery across the wire and measures the potential difference or voltage between the negative end and various positions along the wire. From examination of the meter, he estimates the uncertainty in each potential measurement to be 0.05 V. The uncertainty in the position of the probe is less than 1 mm and is considered to be negligible.

The data are listed in Table 6.1 and are plotted in Figure 6.1 to show the potential difference as a function of wire length x. The estimated common uncertainty in each measured potential difference is indicated on the graph by the vertical error bars. From these measurements, we wish to find the linear function $y(x)$ (shown as a solid line) that describes the way in which the voltage V varies as a function of position x along the wire.

Example 6.2. In another experiment, a student is provided with a radioactive source enclosed in a small 8-mm-diameter plastic disk and a Geiger counter with a 1-cm-diameter end window. Her object is to investigate the $1/r^2$ law by recording Geiger counter measurements over a fixed period of time at various distances from the source between 20 and 100 cm. Because the counting rate is not expected to vary from measurement to measurement, except for statistical fluctuations, the student can record data long enough to obtain good statistics over the entire range of the experiment. She uses an automatic recording system and records counts for thirty 15-s intervals at each position. For analysis in this experiment, she sums the counts from each set of 30 measurements to obtain the number of counts in 7.5 m intervals. The separate 15-s interval measurements at each position can be used in other statistical studies.

The data are listed in Table 6.2 and plotted against $x = 1/r^2$ in Figure 6.2. The vertical error bars on the data points represent the statistical uncertainties in the measured numbers of counts and are equal to the square roots of the numbers of counts. The uncertainties in the measurements of the distances from the source to the counter were assumed to be negligible.

Linear Approximation

In both of these examples, the functional relationship between the dependent and independent variables can be approximated by a straight line of the form

$$y(x) = a + bx \tag{6.1}$$

We shall consider in this chapter a method for determining the most probable values for the coefficients a and b.

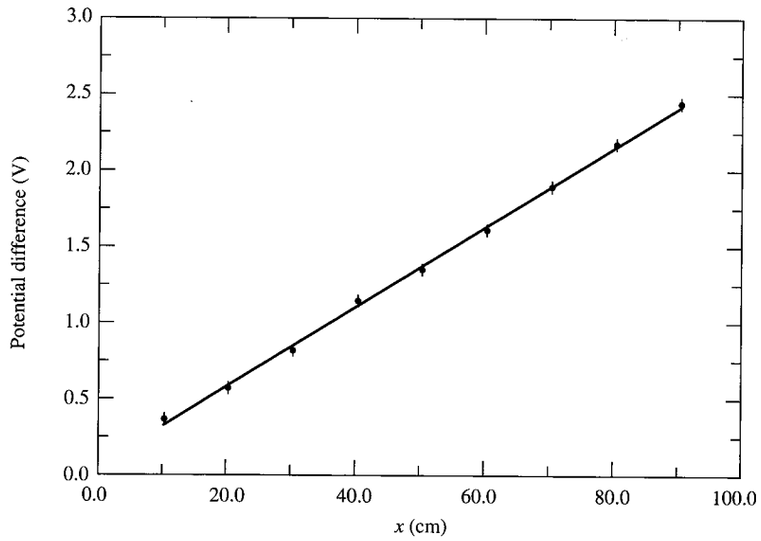


FIGURE 6.1
Potential difference as a function of position along a conducting wire (Example 6.1). The uniform uncertainties in the potential measurements are indicated by the vertical error bars. The straight line is the result of a least-squares fit to the data.

TABLE 6.1
Potential difference V as a function of position along a current-carrying nickel-silver wire

Point number	Postition x_i (cm)	Potential difference V_i (V)	x_i^2	$x_i V_i$	Fitted potential difference $a + bx$
1	10.0	0.37	100	3.70	0.33
2	20.0	0.58	400	11.60	0.60
3	30.0	0.83	900	24.90	0.86
4	40.0	1.15	1,600	46.00	1.12
5	50.0	1.36	2,500	68.00	1.38
6	60.0	1.62	3,600	97.20	1.64
7	70.0	1.90	4,900	133.00	1.91
8	80.0	2.18	6,400	174.40	2.17
9	90.0	2.45	8,100	220.50	2.43
Sums	450.0	12.44	28,500	779.30	

$\Delta = N\Sigma x_i^2 - (\Sigma x_i)^2 = (9 \times 28{,}500) - (450)^2 = 54{,}000$
$a = (\Sigma x_i^2 \Sigma V_i - \Sigma x_i \Sigma x_i V_i)/\Delta = (28{,}500 \times 12.44 - 450.0 \times 779.30)/54{,}000 = 0.0714$
$b = (N\Sigma x_i V_i - \Sigma x_i \Sigma V_i)/\Delta = (9 \times 779.30 - 450.0 \times 12.44)/54{,}000 = 0.0262$
$\sigma_a^2 \simeq \sigma_V^2 \Sigma x_i^2/\Delta = 0.05^2 \times 28{,}500 / 54{,}000 = 0.001319 \qquad \sigma_a \simeq 0.036 \; \sigma'_a = 0.019$
$\sigma_b^2 \simeq N\sigma_V^2/\Delta = 9 \times 0.05^2 / 54{,}000 = 0.417 \times 10^{-6} \qquad \sigma_b \simeq 0.00065 \; \sigma'_b = 0.00034$

Note: A uniform uncertainty in V of 0.05 V is assumed. A linear fit to the data, calculated by the method of determinants, gives $a = 0.07 \pm 0.04$ V and $b = 0.0262 \pm 0.0006$ V/cm, with $\chi^2 = 1.95$ for 7 degrees of freedom. The χ^2 probability for the fit is approximately 96%.

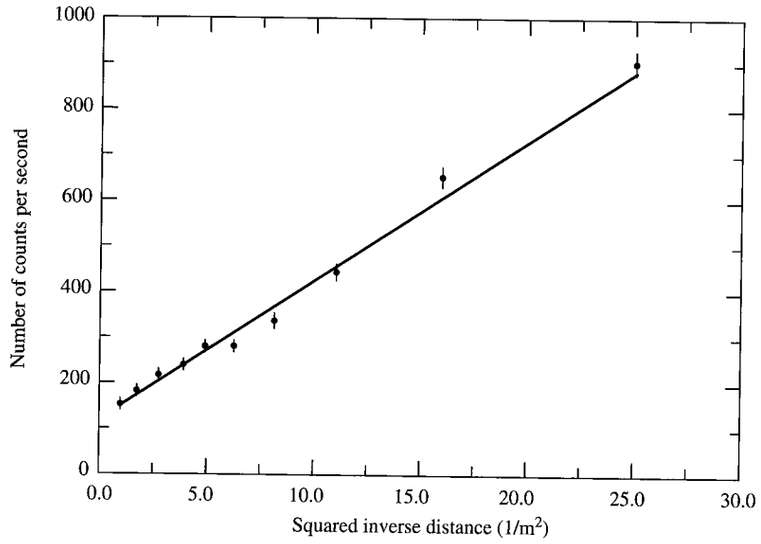


FIGURE 6.2
Number of counts in constant time intervals from a radioactive source as a function of the inverse distance from source to Geiger counter (Example 6.2). The vertical error bars indicate the statistical uncertainties in the counts. The straight line is the result of a least-squares fit to the data.

TABLE 6.2
Number of counts detected in 7½-min intervals as a function of distance from the source

i	Distance d_i (m)	$x_i = 1/d_i^2$ (m^{-2})	Counts C_i	σ_{C_i}	Weight $(1/C_i^2)$ w_i	$w_i x_i$	$w_i C_i$	$w_i x_i^2$	$w_i x_i C_i$	Fitted counts $a + bx_i$
1	0.20	25.00	901	30.0	0.00111	0.0278	1	0.694	25.0	887
2	0.25	16.00	652	25.5	0.00153	0.0254	1	0.393	16.0	610
3	0.30	11.11	443	21.0	0.00226	0.0251	1	0.279	11.1	461
4	0.35	8.16	339	18.4	0.00295	0.0241	1	0.197	8.2	370
5	0.40	6.25	283	16.8	0.00353	0.0221	1	0.138	6.3	311
6	0.45	4.94	281	16.8	0.00356	0.0176	1	0.087	4.9	271
7	0.50	4.00	240	15.5	0.00417	0.0167	1	0.067	4.0	242
8	0.60	2.78	220	14.8	0.00455	0.0126	1	0.035	2.8	205
9	0.75	1.78	180	13.4	0.00556	0.0099	1	0.018	1.8	174
10	1.00	1.00	154	12.4	0.00649	0.0065	1	0.007	1.0	150
Sums					0.03570	0.1868	10	1.912	81.0	

$\sigma_i = \sqrt{y_i} \;\; w_i = 1/\sigma_i^2 = 1/y_i$
$\Delta = \Sigma w_i \Sigma w_i x_i^2 - (\Sigma w_i x_i)^2 = 0.03570 \times 1.912 - (0.1868)^2 = 0.0334$
$a = [\Sigma w_i C_i \Sigma w_i x_i^2 - \Sigma w_i x_i \Sigma w_i x_i C_i]/\Delta = [10 \times 1.912 - 0.1868 \times 81.0]/\Delta = 119.5$
$b = [\Sigma w_i \Sigma w_i x_i C_i - \Sigma w_i x_i \Sigma w_i C_i]/\Delta = [0.03570 \times 81.0 - 0.1868 \times 10]/\Delta = 30.7$
$\sigma_a^2 \simeq \Sigma w_i x_i^2/\Delta = 1.912 / 0.0334 = 57.3 \qquad \sigma_a \simeq 7.6$
$\sigma_b^2 \simeq \Sigma w_i/\Delta = 0.03570 / 0.0334 = 1.07 \qquad \sigma_b \simeq 1.1$

Note: A linear fit to the data of the function $C = a + bx$ by the method of determinants gives $a = 119 \pm 8$ and $b = 31 \pm 1$, with $\chi^2 = 11.1$ for 8 degrees of freedom. The χ^2 probability for the fit is about 20%.

We cannot fit a straight line to the data exactly in either example because it is impossible to draw a straight line through all the points. For a set of N arbitrary points, it is always possible to fit a polynomial of degree $N - 1$ exactly, but for our experiments, the coefficients of the higher-order terms would have questionable significance. We assume that the fluctuations of the individual points above and below the solid curves are caused by experimental uncertainties in the individual measurements. In Chapter 11 we shall develop a method for testing whether higher-order terms are significant.

Measuring Uncertainties

If we were to make a series of measurements of the dependent quantity y_i for one particular value x_i of the independent quantity, we would find that the measured values were distributed about a mean in the manner discussed in Chapter 5 with a probability of ~68% that any single measurement of y_i be within 1 standard deviation of the mean. By making a number of measurements for each value of the independent quantity x_i, we could determine mean values $\bar{y}_i$ with any desired precision. Usually, however, we can make only one measurement y_i for each value of $x = x_i$, so that we must determine the value of y corresponding to that value of x with an uncertainty that is characterized by the standard deviation σ_i of the distribution of data for that point.

We shall assume for simplicity in all the following discussions that we can ascribe all the uncertainty in each measurement to the dependent variable. This is equivalent to assuming that the precision of the determination of x is considerably higher than that of y. This difference is illustrated in Figures 6.1 and 6.2 by the fact that the uncertainties are indicated by error bars for the dependent variables but not for the independent variables.

Our condition, that we neglect uncertainties in x and consider just the uncertainties in y, will be valid only if the uncertainties in y that would be produced by variations in x corresponding to the uncertainties in the measurement of x are much smaller than the uncertainties in the measurement of y. This is equivalent, in first order, to the requirement at each measured point that

$$\sigma_x \frac{dy}{dx} \ll \sigma_y$$

where dy/dx is the slope of the function $y = y(x)$.

We are not always justified in ascribing all uncertainties to the dependent parameter. Sometimes the uncertainties in the determination of both quantities x and y are nearly equal. But our fitting procedure will still be fairly accurate if we estimate the indirect contribution σ_{yI} from the uncertainty σ_x in x to the total uncertainty in y by the first-order relation

$$\sigma_{yI} = \sigma_x \frac{dy}{dx} \tag{6.2}$$

and combine this with the direct contribution σ_{yD}, which is the measuring uncertainty in y, to get

$$\sigma_y^2 = \sigma_{yI}^2 + \sigma_{yD}^2 \tag{6.3}$$

For both Examples 6.1 and 6.2 the condition would be reasonable because we predict a linear dependence of y with x. With the linear assumption, we treat the uncertainties in our data as if they were in the dependent variable only, while realizing that the corresponding fluctuations may have been originally derived from uncertainties in the determinations of both dependent and independent variables.

In those cases where the uncertainties in the determination of the independent quantity are considerably greater than those in the dependent quantity, it might be wise to interchange the definition of the two quantities.

6.2 METHOD OF LEAST SQUARES

Our data consist of pairs of measurements (x_i, y_i) of an independent variable x and a dependent variable y. We wish to find values of the parameters a and b that minimize the discrepancy between the measured values y_i and calculated values $y(x)$. We cannot determine the parameters exactly with only a finite number of observations, but can hope to extract the most probable estimates for the coefficients in the same way that we extracted the most probable estimate of the mean in Chapter 4.

Before proceeding, we must define our criteria for minimizing the discrepancy between the measured and predicted values y_i. For any arbitrary values of a and b, we can calculate the deviations Δy_i between each of the observed values y_i and the corresponding calculated or fitted values

$$\Delta y_i = y_i - y(x_i) = y_i - a - bx_i \tag{6.4}$$

With well chosen parameters, these deviations should be relatively small. However, the sum of these deviations is not a good measure of how well our calculated straight line approximates the data because large positive deviations can be balanced by negative ones to yield a small sum even when the fit of the function $y(x)$ to the data is bad. We might consider instead summing the absolute values of the deviations, but this leads to difficulties in obtaining an analytical solution. Instead we sum the squares of the deviations.

There in no correct unique method for optimizing the parameters valid for all problems. There exists, however, a method that can be fairly well justified, that is simple and straightforward, and that is well established experimentally. This is the *method of least squares,* similar to the method discussed in Chapter 4, but extended to include more than one variable. It may be considered as a special case of the more general *method of maximum likelihood.*

Method of Maximum Likelihood

Our data consist of a sample of observations drawn from a parent distribution that determines the probability of making any particular observation. For the particular problem of an expected linear relationship between dependent and independent variables, we define parent parameters a_0 and b_0 such that the actual relationship between y and x is given by

$$y_0(x) = a_0 + b_0 x \tag{6.5}$$

We shall assume that each individual measured value of y_i is itself drawn from a Gaussian distribution with mean $y_0(x_i)$ and standard deviation σ_i. We should be aware that the Gaussian assumption may not always be exactly true. In Example 6.2 the $y_i = C_i$ were obtained in a counting experiment and therefore follow a Poisson distribution. However, for a sufficiently large number of counts y_i the distribution may be considered to be Gaussian. We shall discuss fitting with Poisson statistics in Section 6.6.

With the Gaussian assumption, the probability P_i for making the observed measurement y_i with standard deviation σ_i for the observations about the actual value $y_0(x_i)$ is

$$P_i = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[\frac{y_i - y_0(x_i)}{\sigma_i}\right]^2\right\} \tag{6.6}$$

The probability for making the observed set of measurements of the N values of y_i is the product of the probabilities for each observation:

$$P(a_0, b_0) = \Pi P_i = \prod\left(\frac{1}{\sigma_i \sqrt{2\pi}}\right) \exp\left\{-\frac{1}{2}\sum\left[\frac{y_i - y_0(x_i)}{\sigma_i}\right]^2\right\} \tag{6.7}$$

where the product Π is taken with i ranging from 1 to N and the product of the exponentials has been expressed as the exponential of the sum of the arguments. In these products and sums, the quantities $1/\sigma_i^2$ act as weighting factors.

Similarly, for any *estimated* values of the parameters a and b, we can calculate the probability of obtaining the observed set of measurements

$$P(a, b) = \prod\left(\frac{1}{\sigma_i\sqrt{2\pi}}\right)\exp\left\{-\frac{1}{2}\sum\left[\frac{y_i - y(x_i)}{\sigma_i}\right]^2\right\} \tag{6.8}$$

with $y(x)$ defined by Equation (6.1) and evaluated at each of the values x_i.

We assume that the observed set of measurements is more likely to have come from the parent distribution of Equation (6.5) than from any other similar distribution with different coefficients and, therefore, the probability of Equation (6.7) is the maximum probability attainable with Equation (6.8). Thus, the maximum-likelihood estimates for a and b are those values that maximize the probability of Equation (6.8).

Because the first factor in the product of Equation (6.8) is a constant, independent of the values of a and b, maximizing the probability $P(a, b)$ is equivalent to minimizing the sum in the exponential. We define this sum to be our goodness-of-fit parameter χ^2:

$$\chi^2 = \sum\left[\frac{y_i - y(x_i)}{\sigma_i}\right]^2 = \sum\left[\frac{1}{\sigma_i}(y_i - a - bx_i)\right]^2 \tag{6.9}$$

We use the same symbol χ^2, defined earlier in Equation (4.32), because this is essentially the same definition in a different context.

Our method for finding the optimum fit to the data will be to find values of a and b that minimize this weighted sum of the squares of the deviations χ^2 and hence, to find the fit that produces the smallest sum of the squares or the *least-squares fit.* The magnitude of χ^2 is determined by four factors:

1. Fluctuations in the measured values of the variables y_i, which are random samples from a parent population with expectation values $y_0(x_i)$.
2. The values assigned to the uncertainties σ_i in the measured variables y_i. Incorrect assignment of the uncertainties σ_i will lead to incorrect values of χ^2.
3. The selection of the analytical function $y(x)$ as an approximation to the "true" function $y_0(x)$. It might be necessary to fit several different functions in order to find the appropriate function for a particular set of data.
4. The values of the parameters of the function $y(x)$. Our objective is to find the "best values" of these parameters.

6.3 MINIMIZING χ^2

To find the values of the parameters a and b that yield the minimum value for χ^2, we set to zero the partial derivatives of χ^2 with respect to each of the parameters

$$\begin{aligned}\frac{\partial}{\partial a}\chi^2 &= \frac{\partial}{\partial a}\sum\left[\frac{1}{\sigma_i^2}(y_i - a - bx)^2\right]\\ &= -2\sum\left[\frac{1}{\sigma_i^2}(y_i - a - bx_i)\right] = 0\\ \frac{\partial}{\partial b}\chi^2 &= \frac{\partial}{\partial b}\sum\left[\frac{1}{\sigma_i^2}(y_i - a - bx_i)^2\right]\\ &= -2\sum\left[\frac{1}{\sigma_i^2}(y_i - a - bx_i)\right] = 0\end{aligned} \tag{6.10}$$

These equations can be rearranged as a pair of linear simultaneous equations in the unknown parameters a and b:

$$\begin{aligned}\sum\frac{y_i}{\sigma_i^2} &= a\sum\frac{1}{\sigma_i^2} + b\sum\frac{x_i}{\sigma_i^2}\\ \sum\frac{x_i y_i}{\sigma_i^2} &= a\sum\frac{x_i}{\sigma_i^2} + b\sum\frac{x_i^2}{\sigma_i^2}\end{aligned} \tag{6.11}$$

The solutions can be found in any one of a number of different ways, but, for generality we shall use the method of determinants. (See Appendix B.) The solutions are

$$\begin{aligned}a &= \frac{1}{\Delta}\begin{vmatrix}\sum\frac{y_i}{\sigma_i^2} & \sum\frac{x_i}{\sigma_i^2}\\ \sum\frac{x_i y_i}{\sigma_i^2} & \sum\frac{x_i^2}{\sigma_i^2}\end{vmatrix} = \frac{1}{\Delta}\left(\sum\frac{x_i^2}{\sigma_i^2}\sum\frac{y_i}{\sigma_i^2} - \sum\frac{x_i}{\sigma_i^2}\sum\frac{x_i y_i}{\sigma_i^2}\right)\\ b &= \frac{1}{\Delta}\begin{vmatrix}\sum\frac{1}{\sigma_i^2} & \sum\frac{y_i}{\sigma_i^2}\\ \sum\frac{x_i}{\sigma_i^2} & \sum\frac{x_i y_i}{\sigma_i^2}\end{vmatrix} = \frac{1}{\Delta}\left(\sum\frac{1}{\sigma_i^2}\sum\frac{x_i y_i}{\sigma_i^2} - \sum\frac{x_i}{\sigma_i^2}\sum\frac{y_i}{\sigma_i^2}\right)\\ \Delta &= \begin{vmatrix}\sum\frac{1}{\sigma_i^2} & \sum\frac{x_i}{\sigma_i^2}\\ \sum\frac{x_i}{\sigma_i^2} & \sum\frac{x_i^2}{\sigma_i^2}\end{vmatrix} = \sum\frac{1}{\sigma_i^2}\sum\frac{x_i^2}{\sigma_i^2} - \left(\sum\frac{x_i}{\sigma_i^2}\right)^2\end{aligned} \tag{6.12}$$

For the special case in which all the uncertainties are equal ($\sigma = \sigma_i$), they cancel and the solutions may be written

$$\begin{aligned}a &= \frac{1}{\Delta'}\begin{vmatrix}\Sigma y_i & \Sigma x_i\\ \Sigma x_i y_i & \Sigma x_i^2\end{vmatrix} = \frac{1}{\Delta'}(\Sigma x_i^2\Sigma y_i - \Sigma x_i\Sigma x_i y_i)\\ b &= \frac{1}{\Delta'}\begin{vmatrix}N & \Sigma y_i\\ \Sigma x_i & \Sigma x_i y_i\end{vmatrix} = \frac{1}{\Delta'}(N\Sigma x_i y_i - \Sigma x_i\Sigma y_i)\\ \Delta' &= \begin{vmatrix}N & \Sigma x_i\\ \Sigma x_i & \Sigma x_i^2\end{vmatrix} = N\Sigma x_i^2 - (\Sigma x_i)^2\end{aligned} \tag{6.13}$$

Examples

For the data of Example 6.1 (Table 6.1), we assume that the uncertainties in the measured voltages V are all equal and that the uncertainties in x_i are negligible. We can therefore use Equation (6.13). We accumulate four sums Σx_i, $\Sigma y_i = \Sigma V_i$, Σx_i^2, and $\Sigma x_i y_i = \Sigma x_i V_i$ and combine them according to Equation (6.13) to find numerical values for a and b. The steps of the calculation are illustrated in Table 6.1, and the resulting fit is shown as a solid line on Figure 6.1.

Determination of the parameters a and b from Equation (6.12) is somewhat more tedious, because the uncertainties σ_i must be included. Table 6.2 shows steps in the calculation of the data of Example 6.2 with the uncertainties σ_i in the numbers of counts C_i determined by Poisson statistics so that $\sigma_i^2 = C_i$. The values of a and b found in this calculation were used to calculate the straight line through the data points in Figure 6.2.

It is important to note that the value of C_i to be used in determining the uncertainty σ_i must be the actual number of events observed. If, for example, the student had decided to improve her statistics by collecting data at the larger distances over longer time periods Δt_i and to normalize all her data to a common time interval Δt_c,

$$C_i' = C_i \times \Delta t_c / \Delta t_i$$

then the statistical uncertainty in C' would be given by

$$\sigma_i' = \sqrt{C_i} \times \Delta t_c / \Delta t_i$$

Program 6.1. FITLINE (Appendix E) Solution of Equations (6.11) by the determinant method of Equation (6.12).

The program uses routines in the programs units FITVARS, FITUTIL, and GENUTIL, which are also used by other fitting programs. The sample programs use single precision variables for simplicity, although double, or higher, precision is highly recommended.

Program 6.1 uses Equation (6.12) to solve both Examples 6.1 and 6.2, although separate routines written for each problem would be slightly more efficient. Because the measurements of Example 6.1 have common errors, we could, for example, increase the fitting speed by using Equations (6.13) rather than Equations (6.12). Similarly, for Example 6.2, we could simplify the fitting routine by replacing the statistical errors SIGY[I] by the explicit expression for $\sqrt{y_i}$. However, in most calculations that involve statistical errors, there are also other errors to be considered, such as those arising from background subtractions, so the loss of generality would more than compensate for any increased efficiency in the calculations.

Program 6.2. FITVARS (website) Include file of constants, variables, and arrays for least-squares fits.

Program 6.3. FITUTIL (website) Utility routines for fitting programs Input/output routine, χ^2 calculation, χ^2-density, and χ^2-integral probability.

Program 6.4. GENUTIL (website) General Utility Routines
Includes approximate gamma function, Simpson's rule integration.

6.4 ERROR ESTIMATION

Common Uncertainties

If the standard deviations σ_i for the data points y_i are unknown but we can assume that they are all equal, $\sigma_i^2 = \sigma^2$, then we can estimate them from the data and the results of our fit. The requirement of equal errors may be satisfied if the uncertainties are instrumental and all the data are recorded with the same instrument and on the same scale, as was assumed in Example 6.1.

In Chapter 2 we obtained, for our best estimate of the variance of the data sample,

$$\sigma^2 \simeq s^2 \equiv \frac{1}{N-m} \sum (y_i - \bar{y})^2 \tag{6.14}$$

where $N - m$ is the number of degrees of freedom and is equal to the number of measurements minus the number of parameters determined from the fit. In Equation (6.14) we identify y_i with the measured value of the dependent variable, and for $\bar{y}$, the expected mean value of y_i, we use the value calculated from Equation (6.1) for each data point with the fitted parameters a and b. Thus, our estimate $\sigma_i = \sigma$ for the standard deviation of an individual measurement is

$$\sigma^2 \simeq s^2 = \frac{1}{N-2} \sum (y_i - a - bx_i)^2 \tag{6.15}$$

By comparing Equation (6.15) with Equation (6.9), we see that it is just this common uncertainty that we have minimized in the least-squares fitting procedure. Thus, we can obtain the common error in our measurements of y from the fit, although at the expense of any information about the quality of the fit.

Variable Uncertainties

In general the uncertainties σ_i in the dependent variables y_i will not all be the same. If, for example, the quantity y represents the number of counts in a detector per unit time interval (as in Example 6.2), then the errors are statistical and the uncertainty in each measurement y_i is directly related to the magnitude of y (as discussed in Section 4.2), and the standard deviations σ_i associated with these measurements is

$$\sigma_i^2 = C_i \tag{6.16}$$

In principle, the value of y_i, which should be used in calculating the standard deviations σ_i by Equation (6.16), is the value $y_0(x_i)$ of the parent population. In practice we use the measured values that are only samples from that population. In the limit of an infinite number of determinations, the average of all the measurements would very closely approximate the parent value, but generally we cannot make more than one measurement of each value of x, much less an infinite number. We

could approximate the parent value $y_0(x_i)$ by using the calculated value $y(x)$ from our fit, but that would complicate the fitting procedure. We shall discuss this possibility further in the following section.

Contributions from instrumental and other uncertainties may modify the simple square root form of the statistical errors. For example, uncertainties in measuring the time interval during which the events of Example 6.2 were recorded might contribute, although statistical fluctuations generally dominate in counting experiments. Background subtractions are another source of uncertainty. In many counting experiments, there is a background under the data that may be removed by subtraction, or may be included in the fit. In Example 6.2, cosmic rays and other backgrounds contribute to a counting rate even when the source is moved far away from the detector, as indicated by the nonzero intercept of the fitted line of Figure 6.2 on the C axis. If the student had chosen to record the radiation background counts C_b in a separate measurement and to subtract C_b from each of her measurements C_i to obtain

$$C_i' = C_i - C_b$$

then the uncertainty in C' would have been given by combining in quadrature the uncertainties in the two measurements:

$$\sigma_i'^2 = \sigma_i^2 + \sigma_b^2$$

χ^2 Probability

For those data for which we know the uncertainties σ_i in the measured values y_i we can calculate the value of χ^2 from Equation (6.9) and test the goodness of our fit. For our two-parameter fit to a straight line, the number of degrees of freedom will be $N - 2$. Then, for the data of Example 6.2, we should hope to obtain $\chi^2 \simeq 10 - 2 = 8$. The actual value, $\chi^2 = 11.1$, is listed in Table 6.2, along with the probability ($p = 20\%$). (See Table C.4.) We interpret this probability in the following way. Suppose that we have obtained a χ^2 probability of $p\%$ for a certain set of data. Then, we should expect that, if we were to repeat the experiment many times, approximately $p\%$ of the experiments would yield χ^2 values as high as the one that we obtained or higher. This subject will be discussed further in Chapter 11.

In Example 6.1, we obtained a value of $\chi^2 = 1.95$ for 7 degrees of freedom, corresponding to a probability of about 96%. Although this probability may seem to be gratifyingly high, the very low value of χ^2 gives a strong indication that the common uncertainty in the data may have been overestimated and it might be wise to use the value of χ^2 to obtain a better estimate of the common uncertainty. From Equations (6.15) and (6.9), we obtain an expression for the revised common uncertainty σ_c' in terms of χ^2 and the original estimate, σ_c:

$$\sigma_c'^2 \simeq \sigma_i^2 \times \chi^2/(N-2) \tag{6.17}$$

or, more generally

$$\sigma_c'^2 \simeq \sigma_i^2 \times \chi_\nu^2 \tag{6.18}$$

where $\chi_\nu^2 = \chi^2/\nu$ and ν is the number of degrees of freedom in the fit. Thus, for Example 6.1, we find $\sigma_c'^2 = 0.05^2 \times 1.95/(9-2) = 0.0007$, or $\sigma_c' = \sim 0.03$ V.

Uncertainties in the Parameters

In order to find the uncertainty in the estimation of the parameters a and b in our fitting procedure, we use the error propagation method discussed in Chapter 3. Each of our data points y_i has been used in the determination of the parameters and each has contributed some fraction of its own uncertainty to the uncertainty in our final determination. Ignoring systematic errors, which would introduce correlations between uncertainties, the variance σ_z^2 of the parameter z is given by Equation (3.14) as the sum of the squares of the products of the standard deviations σ_i of the data points with the effects that the data points have on the determination of z:

$$\sigma_z^2 = \sum \left[\sigma_i^2 \left(\frac{\partial z}{\partial y_i} \right)^2 \right] \tag{6.19}$$

Thus, to determine the uncertainties in the parameters a and b, we take the partial derivatives of Equation (6.12):

$$\begin{aligned} \frac{\partial a}{\partial y_j} &= \frac{1}{\Delta} \left(\frac{1}{\sigma_j^2} \sum \frac{x_i^2}{\sigma_i^2} - \frac{x_j}{\sigma_j^2} \sum \frac{x_i}{\sigma_i^2} \right) \\ \frac{\partial b}{\partial y_j} &= \frac{1}{\Delta} \left(\frac{x_j}{\sigma_j^2} \sum \frac{1}{\sigma_i^2} - \frac{1}{\sigma_j^2} \sum \frac{x_i}{\sigma_i^2} \right) \end{aligned} \tag{6.20}$$

We note that the derivatives are functions only of the variances and of the independent variables x_i. Combining these equations with the general expression of Equation (6.19) and squaring, we obtain for σ^2,

$$\begin{aligned} \sigma_a^2 &\simeq \sum_{j=1}^{N} \frac{\sigma_j^2}{\Delta^2} \left[\frac{1}{\sigma_j^4} \left(\sum \frac{x_i^2}{\sigma_i^2} \right)^2 - \frac{2x_j}{\sigma_j^4} \sum \frac{x_i^2}{\sigma_i^2} \sum \frac{x_i}{\sigma_i^2} + \frac{x_j^2}{\sigma_j^4} \left(\sum \frac{x_i}{\sigma_i^2} \right)^2 \right] \\ &= \frac{1}{\Delta^2} \left[\sum \frac{1}{\sigma_j^2} \left(\sum \frac{x_i^2}{\sigma_i^2} \right)^2 - 2 \sum \frac{x_j}{\sigma_j^2} \sum \frac{x_i^2}{\sigma_i^2} \sum \frac{x_i}{\sigma_i^2} + \sum \frac{x_j^2}{\sigma_j^2} \left(\sum \frac{x_i}{\sigma_i^2} \right)^2 \right] \\ &= \frac{1}{\Delta^2} \left(\sum \frac{x_i^2}{\sigma_i^2} \right) \left[\sum \frac{1}{\sigma_j^2} \sum \frac{x_i^2}{\sigma_i^2} - \left(\sum \frac{x_i}{\sigma_i^2} \right)^2 \right] \\ &= \frac{1}{\Delta} \sum \frac{x_i^2}{\sigma_i^2} \end{aligned} \tag{6.21}$$

and for σ_b^2,

$$\begin{aligned} \sigma_b^2 &\simeq \sum_{j=1}^{N} \frac{\sigma_j^2}{\Delta^2} \left[\frac{x_j^2}{\sigma_j^4} \left(\sum \frac{1}{\sigma_i^2} \right)^2 - \frac{2x_j}{\sigma_j^4} \sum \frac{1}{\sigma_i^2} \sum \frac{x_i}{\sigma_i^2} + \frac{1}{\sigma_j^4} \left(\sum \frac{x_i}{\sigma_i^2} \right)^2 \right] \\ &= \frac{1}{\Delta^2} \left[\sum \frac{x_j^2}{\sigma_j^2} \left(\sum \frac{1}{\sigma_i^2} \right)^2 - 2 \sum \frac{x_j}{\sigma_j^2} \sum \frac{1}{\sigma_i^2} \sum \frac{x_i}{\sigma_i^2} + \sum \frac{1}{\sigma_j^2} \left(\sum \frac{x_i}{\sigma_i^2} \right)^2 \right] \\ &= \frac{1}{\Delta^2} \left(\sum \frac{x_j^2}{\sigma_i^2} \right) \left[\sum \frac{1}{\sigma_j^2} \sum \frac{1}{\sigma_i^2} - \left(\sum \frac{x_i}{\sigma_i^2} \right)^2 \right] \\ &= \frac{1}{\Delta} \sum \frac{1}{\sigma_i^2} \end{aligned} \tag{6.22}$$

For the special case of common uncertainties in y_i, $\sigma_i = \sigma$, these equations reduce to

$$\sigma_a^2 = \frac{\sigma^2}{\Delta'}\Sigma x_i^2 \qquad \text{and} \qquad \sigma_b^2 = N\frac{\sigma^2}{\Delta'} \tag{6.23}$$

with σ given by Equation (6.15) and Δ' given by Equation (6.13).

The uncertainties in the parameters σ_a and σ_b, calculated from the original error estimates, are listed in Tables 6.1 and 6.2. For Example 6.1, revised uncertainties σ_a' and σ_b', based on the revised common data uncertainty calculated from Equation (6.18), are also listed.

6.5 SOME LIMITATIONS OF THE LEAST-SQUARES METHOD

When a curve is fitted by the least-squares method to a collection of statistical counting data, the data must first be *histogrammed*; that is, a histogram must be formed of the corrected data, either during or after data collection. In Example 6.2, the data were collected over intervals of time Δt, with the size of the interval chosen to assure that a reasonable number of counts would be collected in each time interval. For data that vary linearly with the independent variable, this treatment poses no special problems, but one could imagine a more complex problem in which fine details of the variation of the dependent variable y with the independent variable x are important. Such details might well be lost if the binning were too coarse. On the other hand, if the binning interval were too fine, there might not be enough counts in each bin to justify the Gaussian probability hypothesis. How does one choose the appropriate bin size for the data?

A handy rule of thumb when considering the Poisson distribution is to assume that *large enough* = 10. A comparison of the Gaussian and Poisson distributions for mean $\mu \simeq 10$ and standard deviation $\sigma = \sqrt{\mu}$ (see Figures 2.4 and 2.5) shows very little difference between the two distributions. We might expect this because the mean is more than 3 standard deviations away from the origin. Thus, we may be reasonably confident about the results of a fit if no histogram contains less than ten counts and if we are not placing excessive reliance on the actual value of χ^2 obtained from the fit. If a bin does have fewer than the allowed minimum number of counts, it may be possible to merge that bin with an adjacent one. Note that there is no requirement that intervals on the abscissa be equal, although we must be careful in our choice of the appropriate value of x_i for the merged bin. We should also be aware that such mergers necessarily reduce the resolution of our data and may, when fitting functions more complicated than a straight line, obscure some interesting features.

In general, the choice of bin width will be a compromise between the need for sufficient statistics to maintain a small relative error in the values of y_i and thus in the fitted parameters, and the need to preserve interesting structure in the data. When full details of any structure in the data must be preserved, it might be advisable to apply the maximum-likelihood method directly to the data, event by event, rather than to use the least-squares method with its necessary binning of the data. We return to this subject in Chapter 10.

There is also a question about our use of the experimental errors in the fitting process, rather than the errors predicted by our estimate of the parent distribution. For Example 6.2, this corresponds to our choosing $\sigma_i^2 = y_i$ rather than $\sigma_i^2 = y(x_i) = a + bx_i$. We shall consider the possibility of using errors from our estimate of the parent distribution, as well as the direct application of the Poisson probability function, in the following section.

Another important point to consider when fitting curves to data is the possibility of rounding errors, which can reduce the accuracy of the results. With manual calculations, it is important to avoid rounding the numbers until the very end of the calculation. With computers, problems may arise because of finite computer word length. This problem can be especially severe with matrix and determinant calculations, which often involve taking small differences between large numbers. Depending on the computer and the software, it may be necessary to use double-precision variables in the fitting routine.

We discuss in Chapter 7 the interaction of parameters in a multiparameter fit. For now, it is worth noting that, for a nominally "flat" distribution of data, the intercept obtained from a fit to a straight line may not be identical to the mean value of the data points on the ordinate. See Exercise 6.7 for an example of this effect.

6.6 ALTERNATE FITTING METHODS

In this section we attempt to solve the problem of fitting a straight line to a collection of data points by using errors determined from the estimated parent distribution rather than from the measurements, and by directly applying Poisson statistics, rather than Gaussian statistics. Because it is not possible to derive a set of independent linear equations for the parameters with these conditions, explicit expressions for the parameters a and b cannot be obtained. However, with fast computers, solving coupled, nonlinear equations is not difficult, although the clarity and elegance of the straightforward least-squares method can be lost.

Poisson Uncertainties

Let us consider a collection of purely statistical data that obey Poisson statistics (as in Example 6.2) so that the uncertainties can be expressed by Equation (6.16). We begin by substituting the approximation $\sigma_i^2 = y(x_i) = a + bx_i$ into the definition of χ^2 in Equation (6.9), which is based on Gaussian probability, and minimizing the value of χ^2 as in Equations (6.10). The result is a pair of simultaneous equations that can be solved for a and b:

$$\begin{aligned} N &= \sum \frac{y_i^2}{(a + bx_i)^2} \\ \Sigma x_i &= \sum \frac{x_i y_i^2}{(a + bx_i)^2} \end{aligned} \tag{6.24}$$

Poisson Probability

Next, let us replace the Gaussian probability $P(a, b)$ of Equation (6.8) by the corresponding probability for observing y_i counts from a Poisson distribution with mean $\mu_i = y(x_i)$,

$$P(a, b) = \prod \left(\frac{[y(x_i)]^{y_i}}{y_i!} e^{-y(x_i)} \right) \tag{6.25}$$

and apply the method of maximum likelihood to this probability. It is easier and equivalent to maximize the natural logarithm of the probability with respect to each of the parameters a and b:

$$\ln P(a, b) = \Sigma[y_i \ln y(x_i)] - \Sigma y(x_i) + \text{constant} \tag{6.26}$$

where the constant term is independent of the parameters a and b. The result of taking partial derivatives of Equation (6.26) is a pair of simultaneous equations similar to those of Equation (6.24),

$$\begin{aligned} N &= \sum \frac{y_i}{a + bx_i} \\ \Sigma x_i &= \sum \frac{x_i y_i}{a + bx_i} \end{aligned} \tag{6.27}$$

but with less emphasis on fitting the larger values of y_i.

Neither the coupled simultaneous Equations (6.24) nor the Equations (6.27) can be solved directly for a and b, but each pair can be solved by an iterative method in which values of a and b are chosen and then adjusted until the two simultaneous equations are satisfied. (See Appendix A.5.)

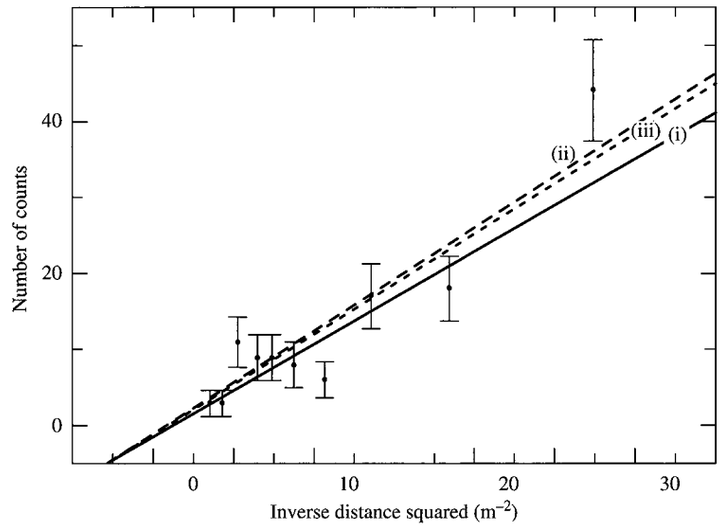


FIGURE 6.3
Least-squares fit of a straight line to the data by three different methods. (i) Standard least-squares method with Gaussian statistics and experimental uncertainties; (ii) Gaussian statistics and analytic uncertainties; (iii) Poisson statistics and analytic uncertainties. The analytic errors are expressed as $\sigma_i^2 = a + bx_i$.

TABLE 6.3
Comparison of fits to a selection of statistical data from Example 6.2 for three different fitting methods

i	Inverse distance squared x_i (m^{-2})	Number of counts C_i	(1) Standard	(2) Gaussian $\sigma^2 = y(x_i)$	(3) Poisson $\sigma^2 = y(x_i)$
1	25.00	44	32.0	36.3	35.1
2	16.00	18	21.0	24.1	23.2
3	11.11	17	15.1	17.5	16.8
4	8.16	6	11.5	13.5	12.9
5	6.25	8	9.1	10.9	10.4
6	4.94	9	7.5	9.2	8.6
7	4.00	9	6.4	7.9	7.4
8	2.78	11	4.9	6.3	5.8
9	1.78	3	3.7	4.9	4.5
10	1.00	3	2.7	3.9	3.4
Sums		128	114.0	134.4	128.0
		a	1.52	2.50	2.11
		b	1.22	1.35	1.32
		χ^2	13.7	17.6	15.5

Note: (1) Standard least-squares method with Gaussian statistics and experimental uncertainties; (2) Gaussian statistics and analytic uncertainties; (3) Poisson statistics and analytic uncertainties. The analytic uncertainties are expressed as $\sigma^2 = a + bx_i$.

Example 6.3. Because we expect the methods discussed here to be equivalent to the standard method for large data samples, we selected a low statistics sample to emphasize the differences. We chose from the measurements of Example 6.2 only those events collected at each detector position during the first 15-s interval, a total of 128 events at ten different positions. The results of (i) calculations by the standard method, (ii) calculations with Gaussian statistics and with errors given by $\sigma_i = y(x_i) = a + bx_i$, and (iii) calculations with Poisson statistics with errors as in method (ii) are listed in Table 6.3 and illustrated in Figure 6.3. We note that method (i) appears to underestimate the number of events in the sample, whereas method (ii) overestimates the number. Method (iii) with Poisson statistics and errors calculated as in method (ii) finds the exact number.

We can avoid questions of finite binning and the choice of statistics by making direct use of the maximum-likelihood method, treating the fitting function as a probability distribution. This method also allows detailed handling of problems in which the probability associated with individual measurements varies in a complex way from observation to observation. We shall pursue this subject further in Chapter 10.

In general, however, the simplicity of the least-squares method and the difficulty of solving the equations that result from other methods, particularly with more complicated fitting functions, leads us to choose the standard method of least squares for most problems. We make the following two assumptions to simplify the calculation:

1. The shapes of the individual Poisson distributions governing the fluctuations in the observed y_i are nearly Gaussian.
2. The uncertainties σ_i in the observations y_i may be obtained from the uncertainties in the data and may be approximated by $\sigma_i^2 \simeq y_i$ for statistical uncertainties.

SUMMARY

Linear function: $y(x) = a + bx$.
Chi-square:

$$\chi^2 = \sum \left[\frac{1}{\sigma_i}(y_i - a - bx_i)\right]^2$$

Least-squares fitting procedure: Minimize χ^2 with respect to each of the parameters simultaneously.
Solutions for least-squares fit of a straight line:

$$a = \frac{1}{\Delta}\begin{vmatrix} \sum \frac{y_i}{\sigma_i^2} & \sum \frac{x_i}{\sigma_i^2} \\ \sum \frac{x_i y_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} \end{vmatrix} = \frac{1}{\Delta}\left(\sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2}\right)$$

$$b = \frac{1}{\Delta}\begin{vmatrix} \sum \frac{1}{\sigma_i^2} & \sum \frac{y_i}{\sigma_i^2} \\ \sum \frac{x_i}{\sigma_i^2} & \sum \frac{x_i y_i}{\sigma_i^2} \end{vmatrix} = \frac{1}{\Delta}\left(\sum \frac{1}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2}\right)$$

$$\Delta = \begin{vmatrix} \sum \frac{1}{\sigma_i^2} & \sum \frac{x_i}{\sigma_i^2} \\ \sum \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} \end{vmatrix} = \sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - \left(\sum \frac{x_i}{\sigma_i^2}\right)^2$$

Estimated uniform variance s^2:

$$\sigma^2 \simeq s^2 = \frac{1}{N-2} \sum (y_i - \bar{y})^2$$

Statistical fluctuations:

$$\sigma_i^2 \simeq y_i \qquad \text{(raw data counts)}$$

Uncertainties in coefficients:

$$\sigma_a^2 = \frac{1}{\Delta} \sum \frac{x_i^2}{\sigma_i^2} \qquad \sigma_b^2 = \frac{1}{\Delta} \sum \frac{1}{\sigma_i^2}$$

EXERCISES

6.1. Fit the data of Example 6.2 as if all the data had equal uncertainties $\sigma_i = \bar{\sigma} = 18.5$, where $\bar{\sigma}$ is the average of the given values of σ. Note that the fitted parameters are independent of the value of $\bar{\sigma}$, but the values of χ^2, σ_a, and σ_b are not.

6.2. Derive Equation (6.23) from Equations (6.21) and (6.22).

6.3. Show that Equation (6.12) reduces to Equation (6.13) if $\sigma_i = \sigma$.

6.4. Derive a formula for making a linear fit to data with an intercept at the origin so that $y = bx$. Apply your method to fit a straight line through the origin to the following coordinate pairs. Assume uniform uncertainties $\sigma_i = 1.5$ in y_i. Find χ^2 for the fit and the uncertainty in b.

x_i	2	4	6	8	10	12	14	16	18	20	22	24
y_i	5.3	14.4	20.7	30.1	35.0	41.3	52.7	55.7	63.0	72.1	80.5	87.9

6.5. A student hangs masses on a spring and measures the spring's extension as a function of the applied force in order to find the spring constant k. Her measurements are:

Mass (kg)	200	300	400	500	600	700	800	900
Extension (cm)	5.1	5.5	5.9	6.8	7.4	7.5	8.6	9.4

There is an uncertainty of 0.2 in each measurement of the extension. The uncertainty in the masses is negligible. For a perfect spring, the extension ΔL of the spring will be related to the applied force by the relation $k\Delta L = F$, where $F = mg$, and $\Delta L = L - L_0$, and L_0 is the unstretched length of the spring. Use these data and the method of least squares to find the spring constant k, the unstretched length of the spring L_0, and their uncertainties. Find χ^2 for the fit and the associated probability.

6.6. Outline a procedure for solving the simultaneous Equations (6.27). Refer to Appendix A.

6.7. A student measures the temperature (T) of water in an insulated flask at times (t) separated by 1 minute and obtains the following values:

t(s)	0	1	2	3	4	5	6	7	8
T(°C)	98.51	98.50	98.50	98.49	98.52	98.49	98.52	98.45	98.47

(*a*) Calculate the mean temperature and its standard error.
(*b*) To test whether or not the water is cooling, plot a graph of the temperatures versus the time and make a least-squares fit of a straight line to the data. Is there a statistically significant slope to the graph?
(*c*) Note that the intercept is not identical to the mean value of the temperature you calculated in part (*a*). Now, shift the time coordinates by 4 s so that the mean time is 0. Refit the data with the new values of T. Is the intercept now identical to the mean value of T?
(*d*) Clearly, the results of this experiment cannot depend upon the time at which the measurements were made. Show that, if the mean value of x is equal to zero, then the intercept b calculated from Equation (6.13) is identically equal to the mean

CHAPTER 7

LEAST-SQUARES FIT TO A POLYNOMIAL

7.1 DETERMINANT SOLUTION

So far we have discussed fitting a straight line to a group of data points. However, suppose our data (x_i, y_i) were not consistent with a straight line fit. We might construct a more complex function with extra parameters and try varying the parameters of this function to fit the data more closely. A very useful function for such a fit is a power-series polynomial

$$y(x) = a_1 + a_2x + a_3x^2 + a_4x^3 + \cdots + a_mx^{m-1} \tag{7.1}$$

where the dependent variable y is expressed as a sum of power series of the independent variable x with coefficients a_1, a_2, a_3, a_4, and so forth.

For problems in which the fitting function is linear in the parameters, the method of least squares is readily extended to any number of terms m, limited only by our ability to solve m linear equations in m unknowns and by the precision with which calculations can be made. We can rewrite Equation (7.1) as

$$y(x) = \sum_{k=1}^{m} a_k x^{k-1} \tag{7.2}$$

where the index k runs from 1 to m. In fact, we can generalize the method even further by writing Equation (7.2) as

$$y(x) = \sum_{k=1}^{m} a_k f_k(x) \tag{7.3}$$

where the functions $f_k(x)$ could be the powers of x as in Equation (7.2), $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = x^2$, and so forth, or they could be other functions of x as long as they *do not involve the parameters* a_1, a_2, a_3, and so forth.

With this definition, the probability function of Equation (6.8) can be written as

$$P(a_1, a_2, \ldots, a_m) = \prod \left(\frac{1}{\sigma_i\sqrt{2\pi}}\right) \exp\left\{-\frac{1}{2}\sum \frac{1}{\sigma_i^2}\left[y_i - \sum_{k=1}^{m} a_k f_k(x_i)\right]^2\right\} \tag{7.4}$$

and Equation (6.9) for χ^2 becomes

$$\chi^2 = \sum \left[\frac{1}{\sigma_i}\left[y_i - \sum_{k=1}^{m} a_k f_k(x_i)\right]\right]^2 \tag{7.5}$$

The method of least squares requires that we minimize χ^2, our measure of the goodness of fit to the data, with respect to the parameters a_1, a_2, a_3, and so forth. The minimum is determined by taking partial derivatives with respect to each parameter in the expression for χ^2 of Equation (7.5), and setting them to zero:

$$\frac{\partial}{\partial a_l}\chi^2 = \frac{\partial}{\partial a_l}\sum \left[\frac{1}{\sigma_i}\left[y_i - \sum_{k=1}^{m} a_k f_k(x_i)\right]\right]^2$$

$$= -2\sum \left\{\frac{f_l(x_i)}{\sigma_i^2}\left[y_i - \sum_{k=1}^{m} a_k f_k(x_i)\right]\right\} = 0 \tag{7.6}$$

Thus, we obtain a set of m coupled linear equations for the m parameters a_l, with the index l running from 1 to m:

$$\sum y_i \frac{f_l(x_i)}{\sigma_i^2} = \sum_{k=1}^{m} \left\{a_k \sum \left[\frac{1}{\sigma_i^2} f_l(x_i) f_k(x_i)\right]\right\}$$

or

$$\begin{aligned}
\sum y_i \frac{f_1(x_i)}{\sigma_i^2} &= \sum \frac{f_1(x_i)}{\sigma_i^2}[a_1 f_1(x_i) + a_2 f_2(x_i) + a_3 f_3(x_i) \cdots] \\
\sum y_i \frac{f_2(x_i)}{\sigma_i^2} &= \sum \frac{f_2(x_i)}{\sigma_i^2}[a_1 f_1(x_i) + a_2 f_2(x_i) + a_3 f_3(x_i) \cdots] \\
\sum y_i \frac{f_3(x_i)}{\sigma_i^2} &= \sum \frac{f_3(x_i)}{\sigma_i^2}[a_1 f_1(x_i) + a_2 f_2(x_i) + a_3 f_3(x_i) \cdots]
\end{aligned} \tag{7.7}$$

and so forth.

The solutions can be found by the method of determinants, as in Chapter 6. We shall display the full solution for the particular case of $m = 3$:

$$a_1 = \frac{1}{\Delta}\begin{vmatrix} \sum y_i \frac{f_1(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_3(x_i)}{\sigma_i^2} \\ \sum y_i \frac{f_2(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_3(x_i)}{\sigma_i^2} \\ \sum y_i \frac{f_3(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_3(x_i)}{\sigma_i^2} \end{vmatrix}$$

$$a_2 = \frac{1}{\Delta}\begin{vmatrix} \sum \frac{f_1(x_i)f_1(x_i)}{\sigma_i^2} & \sum y_i \frac{f_1(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_3(x_i)}{\sigma_i^2} \\ \sum \frac{f_2(x_i)f_1(x_i)}{\sigma_i^2} & \sum y_i \frac{f_2(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_3(x_i)}{\sigma_i^2} \\ \sum \frac{f_3(x_i)f_1(x_i)}{\sigma_i^2} & \sum y_i \frac{f_3(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_3(x_i)}{\sigma_i^2} \end{vmatrix} \tag{7.8}$$

$$a_3 = \frac{1}{\Delta}\begin{vmatrix} \sum \frac{f_1(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_2(x_i)}{\sigma_i^2} & \sum y_i \frac{f_1(x_i)}{\sigma_i^2} \\ \sum \frac{f_2(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_2(x_i)}{\sigma_i^2} & \sum y_i \frac{f_2(x_i)}{\sigma_i^2} \\ \sum \frac{f_3(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_2(x_i)}{\sigma_i^2} & \sum y_i \frac{f_3(x_i)}{\sigma_i^2} \end{vmatrix}$$

with $$\Delta = \begin{vmatrix} \sum \frac{f_1(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_3(x_i)}{\sigma_i^2} \\ \sum \frac{f_2(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_3(x_i)}{\sigma_i^2} \\ \sum \frac{f_3(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_3(x_i)}{\sigma_i^2} \end{vmatrix}$$

We note that, as in the straight-line fits in Chapter 6, the denominator Δ is a function only of the independent variable x and the uncertainties σ_i in the dependent variable, and is not a function of the dependent variable y_i itself. For the special case of a quadratic power series in x, $y(x_i) = a_1 + a_2 x_i + a_3 x_i^2$, we have $f_1(x_i) = 1$, $f_2(x_i) = x_i$, and $f_3(x_i) = x^2$, so that Equations (7.8) become

$$a_1 = \frac{1}{\Delta}\begin{vmatrix} \sum y_i \frac{1}{\sigma_i^2} & \sum \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} \\ \sum y_i \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} & \sum \frac{x_i^3}{\sigma_i^2} \\ \sum y_i \frac{x_i^2}{\sigma_i^2} & \sum \frac{x_i^3}{\sigma_i^2} & \sum \frac{x_i^4}{\sigma_i^2} \end{vmatrix}$$

$$a_2 = \frac{1}{\Delta}\begin{vmatrix} \sum \frac{1}{\sigma_i^2} & \sum y_i \frac{1}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} \\ \sum \frac{x_i}{\sigma_i^2} & \sum y_i \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^3}{\sigma_i^2} \\ \sum \frac{x_i^2}{\sigma_i^2} & \sum y_i \frac{x_i^2}{\sigma_i^2} & \sum \frac{x_i^4}{\sigma_i^2} \end{vmatrix} \tag{7.9}$$

$$a_3 = \frac{1}{\Delta}\begin{vmatrix} \sum \frac{1}{\sigma_i^2} & \sum \frac{x_i}{\sigma_i^2} & \sum y_i \frac{1}{\sigma_i^2} \\ \sum \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} & \sum y_i \frac{x_i}{\sigma_i^2} \\ \sum \frac{x_i^2}{\sigma_i^2} & \sum \frac{x_i^3}{\sigma_i^2} & \sum y_i \frac{x_i^2}{\sigma_i^2} \end{vmatrix}$$

with $$\Delta = \begin{vmatrix} \sum \frac{1}{\sigma_i^2} & \sum \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} \\ \sum \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} & \sum \frac{x_i^3}{\sigma_i^2} \\ \sum \frac{x_i^2}{\sigma_i^2} & \sum \frac{x_i^3}{\sigma_i^2} & \sum \frac{x_i^4}{\sigma_i^2} \end{vmatrix}$$

Example 7.1. A student plans to use a thermocouple to monitor temperatures and must first calibrate it against a thermometer. The thermocouple consists of a junction of a copper wire and a constantan wire. In order to measure the junction voltage with high precision, she connects the sample junction in series with a reference junction that is held at 0°C in an ice water bath. The data, therefore, will be valid only for calibrating the relative variation of the junction voltage with temperature. The absolute voltage must be determined in a separate experiment by measuring it at one specific temperature.

The student measures the difference in output voltage between the two junctions for a temperature variation in the sample junction from 0 to 100°C in steps of 5°C. The measurements are made on the 3-mV scale of the voltmeter, and fluctuations of the

TABLE 7.1
Experimental data for the determination of the relative output voltage V of a thermocouple junction as a function of temperature T of the junction

Trial i	Temperature T (°C)	Measured voltage V (mV)	Calculated voltage $V(T)$ (mV)
1	0.	−0.849	−0.918
2	5.	−0.738	−0.728
3	10.	−0.537	−0.536
4	15.	−0.354	−0.341
5	20.	−0.196	−0.143
6	25.	−0.019	0.058
7	30.	0.262	0.261
8	35.	0.413	0.467
9	40.	0.734	0.676
10	45.	0.882	0.888
11	50.	1.258	1.102
12	55.	1.305	1.319
13	60.	1.541	1.539
14	65.	1.768	1.761
15	70.	1.935	1.987
16	75.	2.147	2.215
17	80.	2.456	2.446
18	85.	2.676	2.679
19	90.	2.994	2.915
20	95.	3.200	3.155
21	100.	3.318	3.396

$a_1 = -0.918 \pm 0.030$
$a_2 = 0.0377 \pm 0.0013$
$a_3 = 0.000055 \pm 0.000013$

Note: The common uncertainty in the voltage measurement is assumed to be 0.05 V. The value of χ^2 for the fit was $\chi^2 = 26.6$ for 18 degrees of freedom, with a probability of 8.8%. Parameters obtained from the fit are listed at the bottom of the table.

needle indicate that the uncertainties in the measurements are approximately 0.05 mV for all readings.

Data from the experiment are listed in Table 7.1 and are plotted in Figure 7.1. To a first approximation, the variation of V with T is linear, but close inspection of the graph reveals a slight curvature. Theoretically, we expect a good fit to these data with a quadratic curve of the form $V = a_1 + a_2T + a_3T^2$.

The parameters for the fit to the data of Example 7.1 have been obtained by evaluating the sums and determinants of Equations (7.9). For a second-degree polynomial with 21 data points, Equation (7.5) becomes

$$\chi^2 \equiv \sum_{i=1}^{21} \frac{1}{\sigma_i^2}[y_i - a_1 - a_2x_i - a_3x_i^2]^2 \tag{7.10}$$

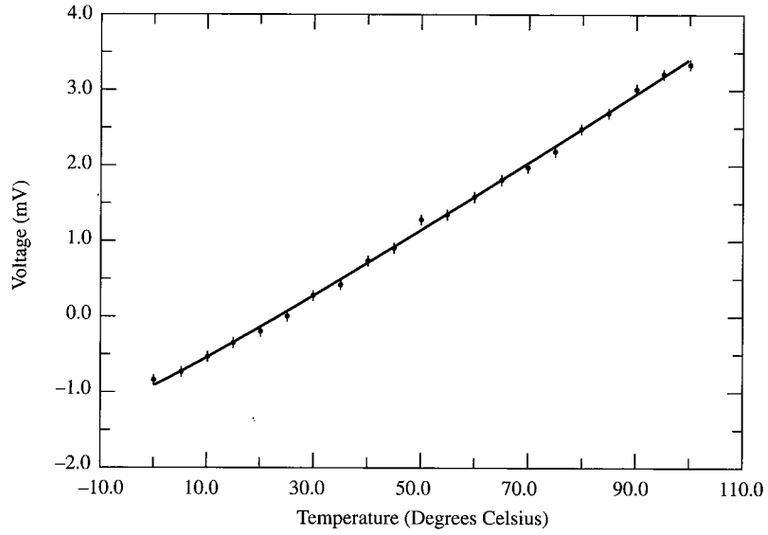


FIGURE 7.1
Thermocouple voltage versus temperature (Example 7.1). The curved line was calculated by fitting to the data second-degree polynomial $V = a_1 + a_2T + a_3T^2$ by the least-squares method. Uniform uncertainties were assumed.

The values of χ^2 and the parameters a_1, a_2, and a_3 determined from the fit are listed in Table 7.1, as are the calculated values of $V(T_i) = y(x_i)$. The calculated values of V are also represented by the solid line on the graph of Figure 7.1. We obtain $\chi^2 = 26.6$ for this fit, or $\chi_\nu^2 = \chi^2/\nu = 1.5$, where the number of degrees of freedom ν is related to the number of events N and the number of free parameters m by $\nu = N - m$. The probability for obtaining χ^2 this high or higher can be determined from the χ^2-probability distribution (see Table C.4) and is about 8.8%, indicating a reasonable fit to the data.

As an alternative to calculating χ^2 from the fit, we could extend Equation (6.15) to three parameters and calculate the average uncertainty in the temperature readings to obtain

$$\sigma^2 \simeq s^2 = \frac{1}{N-m}\sum_{i=1}^{21}[[y_i - (a_1 + a_2x_i + a_3x_i^2)]]^2 = \frac{\chi^2}{N-m} \tag{7.11}$$

which is just the value of the uncertainty that would make $\chi_\nu^2 = 1$. For Example 7.1, we obtain for an estimate of the variance,

$$\sigma'^2 = \sigma^2 \times \chi^2/(N-n) = 0.05 \times 26.6/18 = 0.06°\text{C}$$

suggesting, perhaps, that the student slightly underestimated the uncertainty in her measurements of V.

7.2 MATRIX SOLUTION

The techniques of least-squares fitting fall under the general name of regression analysis. Because we have been considering only problems in which the fitting function

$$y(x_i) = \sum_{k=1}^{m} a_k f_k(x_i) \tag{7.12}$$

is linear in the *parameters* a_k, we are considering only linear regression or multiple linear regression, usually shortened to multiple regression. In Chapter 8 we deal with techniques for handling problems with fitting functions that are not linear in the parameters.

Matrix Equations

We have not yet determined the uncertainties in the three parameters we obtained when we fitted the second-order equation to the data of Example 7.1. We could find the uncertainties by extending the method used for the linear fits of Examples 6.1 and 6.2. However, the algebra becomes even more tedious as the number of terms in the fitted equation increases, and in fact, our method only yielded estimates of the variances σ_k^2 and not of the covariances σ_{kl}^2, which are often important for fitted parameters. Rather than pursue the determinant method, we shall discuss immediately the more elegant and general matrix method for solving the multiple regression problem. Some of the properties of matrices are discussed in Appendix B.

Equations (7.7) can be expressed in matrix form as the equivalence between a row matrix $\boldsymbol{\beta}$ and the product of a row matrix $\mathbf{a}$ with a symmetric matrix $\boldsymbol{\alpha}$, all of order m:

$$\boldsymbol{\beta} = \mathbf{a}\boldsymbol{\alpha} \tag{7.13}$$

The elements of the row matrix $\boldsymbol{\beta}$ are defined by

$$\beta_k \equiv \sum \left[\frac{1}{\sigma_i^2} y_i f_k(x_i) \right] \tag{7.14}$$

those of the symmetric matrix $\boldsymbol{\alpha}$ by

$$\alpha_{lk} \equiv \sum \left[\frac{1}{\sigma_i^2} f_l(x_i) f_k(x_i) \right] \tag{7.15}$$

and the elements of the row matrix $\mathbf{a}$ are the parameters of the fit. For $m = 3$, the matrices may be written as

$$\boldsymbol{\beta} = [\beta_1 \ \ \beta_2 \ \ \beta_3] \qquad \mathbf{a} = [a_1 \ \ a_2 \ \ a_3] \tag{7.16}$$

and

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \tag{7.17}$$

To solve for the parameter matrix $\mathbf{a}$ we multiply both sides of Equation (7.13) on the right by the inverse $\boldsymbol{\epsilon}$ of the matrix $\boldsymbol{\alpha}$, defined such that $\boldsymbol{\alpha}\boldsymbol{\epsilon} = \boldsymbol{\alpha}\boldsymbol{\alpha}^{-1} = \mathbf{1}$, the unity matrix. We obtain

$$\boldsymbol{\beta}\boldsymbol{\epsilon} = \mathbf{a}\boldsymbol{\alpha}\boldsymbol{\epsilon} = \mathbf{a} \tag{7.18}$$

which gives

$$\mathbf{a} = \boldsymbol{\beta}\boldsymbol{\epsilon} = \boldsymbol{\beta}\boldsymbol{\alpha}^{-1} \tag{7.19}$$

Equation (7.19) can also be expressed as

$$a_l = \sum_{k=1}^{m} (\beta_k \epsilon_{kl}) = \sum_{k=1}^{m} \left\{ \epsilon_{kl} \sum \left[\frac{1}{\sigma_i^2} y_i f_k(x_i) \right] \right\} \tag{7.20}$$

where the β_k is given by Equation (7.14).

The solution of Equation (7.19) requires that the matrix $\boldsymbol{\alpha}$ be inverted. This generally is not a simple procedure, except for matrices of very low order, but computer routines are readily available. The inversion of a matrix is discussed in Appendix B.

The symmetric matrix $\boldsymbol{\alpha}$ is called the *curvature matrix* because of its relationship to the curvature of the χ^2 function in parameter space. The relationship becomes apparent when we take the second derivatives of χ^2 with respect to the parameters. From Equation (7.6), we have for the partial derivative of χ^2 with respect to any arbitrary parameter a_l,

$$\frac{\partial \chi^2}{\partial a_l} = -2 \sum \left\{ \frac{f_l(x_i)}{\sigma_i^2} \left[y_i - \sum_{k=1}^{m} a_k f_k(x_i) \right] \right\} \tag{7.21}$$

and the second cross-partial derivative with respect to two such parameters is

$$\frac{\partial^2 \chi^2}{\partial a_l \, \partial a_k} = 2 \sum \left[\frac{1}{\sigma_i^2} f_l(x_i) f_k(x_i) \right] = 2\alpha_{lk} \tag{7.22}$$

Estimation of Errors

The variance $\sigma_{a_l}^2$ of any parameter a_l is the sum of the variances of each of the data points σ_i multiplied by the square of the effect that each data point has on the determination of the parameter a_l [see Equation (6.19)]. Similarly, the covariance of two parameters a_j and a_l is given by

$$\sigma_{a_j a_l}^2 = \sum \left[\sigma_i^2 \frac{\partial a_j}{\partial y_i} \frac{\partial a_l}{\partial y_i} \right] \tag{7.23}$$

(which also gives the variance for $j = l$), where we have assumed that there are no correlations between uncertainties in the measured variables y_i. Taking the derivatives in Equation (7.23) of a_l with respect to y_i we obtain

$$\frac{\partial a_l}{\partial y_i} = \sum_{k=1}^{m} \left[\epsilon_{lk} \frac{1}{\sigma_i^2} f_k(x_i) \right] \tag{7.24}$$

and, substituting into Equation (7.23), we obtain for the weighted sum of the squares of the derivatives,

$$\begin{aligned}\sigma^2_{a_j a_l} &= \sum\left\{\sigma_i^2 \sum_{k=1}^{m}\left[\epsilon_{jk}\frac{1}{\sigma_i}f_k(x_i)\right]\sum_{p=1}^{m}\left[\epsilon_{lp}\frac{1}{\sigma_i}f_p(x_i)\right]\right\}\\ &= \sum_{k=1}^{m}\left\{\epsilon_{jk}\sum_{p=1}^{m}\left[\epsilon_{lp}\sum\left(\frac{1}{\sigma_i^2}f_p(x_i)f_k(x_i)\right)\right]\right\}\\ &= \sum_{k=1}^{m}\left\{\epsilon_{jk}\sum_{p=1}^{m}[\epsilon_{lp}\cdot\alpha_{pk}]\right\}\\ &= \sum_{k=1}^{m}[\epsilon_{kj}\cdot\mathbf{1}_{lk}] = \epsilon_{jl}\end{aligned} \tag{7.25}$$

where we have switched the order of the sums over the dummy indices i, k, and l and have used the fact that because the curvature matrix $\boldsymbol{\alpha}$ is symmetric, its inverse $\boldsymbol{\epsilon}$ must also be symmetric, so that $\epsilon_{kj} = \epsilon_{jk}$. The elements of the unity matrix, which result from the summed products of the elements of $\boldsymbol{\alpha}$ with its inverse $\boldsymbol{\epsilon}$, are represented by $\mathbf{1}_{lk}$.

The inverse matrix $\boldsymbol{\epsilon} \equiv \boldsymbol{\alpha}^{-1}$ is called the error matrix or the covariance matrix because its elements are the variances and covariances of the fitted parameters $\sigma^2_{a_j a_l} = \epsilon_{jl}$.

Example 7.2. The matrix method is illustrated by a straight-line fit $V = a_1 + a_2 T$ to a selection of data from Example 7.1. To show clearly each step of the calculation, we have selected just six points spaced at 25° intervals between 0 and 100° and have assumed a common uncertainty in the dependent variable $\sigma_v = 0.05$ mV. The data are listed in the columns 2 and 3 of Table 7.2*a*.

We begin by calculating each of the fitting functions $f_1 = 1$ and $f_2 = x$ at each value of the independent variable T. These are listed in columns 4 and 5 of Table 7.2*a*. For each measured value of x, the values of β_k, the elements of the column matrix $\boldsymbol{\beta}$, and of α_{lk}, the elements of the symmetric matrix $\boldsymbol{\alpha}$, are calculated according to Equations (7.14) and (7.15). The individual terms in the calculation of β_1 and β_2 are listed in columns 6 and 7 of Table 7.2*a* and the individual terms in the calculation of α_{lk} are listed in columns 8 through 10. (We assume symmetry in $\boldsymbol{\alpha}$.) The resulting matrices are displayed in Table 7.2*b*.

The symmetric matrix $\boldsymbol{\alpha}$ is inverted to obtain the variance matrix $\boldsymbol{\epsilon}$ with elements ϵ_{kl}, shown in Table 7.2*b*, and the product matrix of the fitted parameters $\boldsymbol{\alpha} = \boldsymbol{\beta}\boldsymbol{\epsilon}$ is calculated and displayed in Table 7.2*b*. The calculated values of the fitted variable V for each value of the independent variable T are listed in the last column of Table 7.2*a*.

Program 7.1. MULTREGR (Appendix E) Least-squares fitting with matrices.

Multiple regression problems are usually solved by computer. The program MULTREGR calls a set of routines for fitting any function that is linear in the parameters $a_1, a_2, \ldots, a_m$ to a set of N data points. Branches in the program on the character variable PAE permit selection of the fitting function for each example in this chapter, with PAE = 'P' for the power series in x, PAE = 'A' for all terms of a fourth-order Legendre polynomial, or PAE = 'E' for only the even terms in the Legendre polynomial. The program uses several program units in addition to those referred to in Chapter 6.

TABLE 7.2
Matrix solution for linear fit to data of Example 7.2

(*a*) Data and components of matrix elements

i	T	V	$f_1(x_i)$	$f_2(x_i)$	β_1	β_2	α_{11}	α_{12}	α_{22}	V_{fit}
1	0	−0.849	1	0	−339.6	0	400	0	0	−0.947
2	20	−0.196	1	20	−78.4	−1,458	400	8,000	160,000	−0.101
3	40	0.734	1	40	293.6	11,744	400	16,000	640,000	0.745
4	60	1.541	1	60	616.4	36,984	400	24,000	1,440,000	1.590
5	80	2.456	1	80	982.4	78,592	400	32,000	2,560,000	2.436
6	100	3.318	1	100	1327.6	132,720	400	40,000	4,000,000	3.281
					2802.0	258,472	2,400	120,000	8,800,000	

(*b*) Matrices

$$\boldsymbol{\alpha} = \begin{bmatrix} 2{,}400 & 120{,}000 \\ 120{,}000 & 8{,}800{,}00 \end{bmatrix} \qquad \boldsymbol{\epsilon} = \begin{bmatrix} 1.310\times10^{-03} & -1.786\times10^{-05} \\ -1.786\times10^{-05} & 3.571\times10^{-07} \end{bmatrix}$$

$$\boldsymbol{\beta} = [2{,}802 \quad 258{,}472] \qquad \mathbf{a} = [-0.947 \quad 0.0423]$$

Note: The uniform uncertainty in V was assumed to be 0.05 mV as in Example 7.1. The columns labeled β_1 and α_{11}, etc. correspond to the individual contributions by each measured coordinate pair to the summed values of β and α. The value of χ^2 for the fit was 9.1 for 4 degrees of freedom corresponding to a probability of 5.5%.

Program 7.2. FITFUNC7 (Appendix E) Fitting functions and χ^2 calculation.

In general, every fitting problem requires such a routine. The function POWERFUNC calculates the individual terms in a power function of any order in x for Example 7.2, or Legendre polynomials for Example 7.3.

Program 7.3. MAKEAB7 (Appendix E) Form the arrays for the matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

Program B.1. MATRIX (Appendix E) Matrix products and inversion.

When we use the matrix method to fit a polynomial function to a data sample, the resulting parameters must be identical to those calculated by the determinant method, but we also obtain the full error matrix. The error matrix obtained by fitting a second-degree polynomial to the complete data sample of Example 7.1 is listed in Table 7.3.

The error matrix can be used to estimate the uncertainty in a calculated result, including the effects of the correlations of the errors. As an example, let us suppose that we wish to find the predicted value of the voltage V and its uncertainty for a temperature of exactly 80°C. We should calculate

$$V = a_1 + a_2 T + a_3 T^2 \tag{7.26}$$

using the parameters determined by the fit to the data. The uncertainty in the calculated value of V, which results from the uncertainty in the parameters, is given by Equation (3.13),

TABLE 7.3
Error matrix from a fit by the matrix method to the data of Table 7.1

$$\begin{bmatrix} 8.907 \times 10^{-04} & -3.473 \times 10^{-05} & 2.823 \times 10^{-07} \\ -3.473 \times 10^{-05} & 1.913 \times 10^{-06} & -1.783 \times 10^{-08} \\ 2.823 \times 10^{-07} & -1.783 \times 10^{-08} & 1.783 \times 10^{-10} \end{bmatrix}$$

Note: The table gives the variances and covariances of the fitted parameters. The values of the parameters and of χ^2 are listed in Table 7.1.

$$s^2 = \left(\frac{\partial V}{\partial a_1}\right)^2 \sigma_1^2 + \left(\frac{\partial V}{\partial a_2}\right)^2 \sigma_2^2 + \left(\frac{\partial V}{\partial a_3}\right)^2 \sigma_3^2 + 2\left(\frac{\partial V}{\partial a_1}\frac{\partial V}{\partial a_2}\right)\sigma_{12}^2 + 2\left(\frac{\partial V}{\partial a_1}\frac{\partial V}{\partial a_3}\right)\sigma_{13}^2 + 2\left(\frac{\partial V}{\partial a_2}\frac{\partial V}{\partial a_3}\right)\sigma_{23}^2 \tag{7.27}$$
$$= 1 \cdot \epsilon_{11} + T^2 \cdot \epsilon_{22} + T^4 \cdot \epsilon_{33} + 2(T \cdot \epsilon_{12} + T^2 \cdot \epsilon_{13} + T^3 \cdot \epsilon_{23})$$

where ϵ_{12} and so on are the covariant terms in the symmetric error matrix. If we used only the diagonal terms in the error matrix, our result would be $V = (2.45 \pm 0.14)$ V. However, the off-diagonal terms are mainly negative, and including them reduces the uncertainty by almost a factor of 10 to 0.015, so that we should quote $V = (2.45 \pm 0.02)$ V.

Linear Least-Squares Fitting with a Spreadsheet

Table 7.4 illustrates the use of a spreadsheet (without taking advantage of the spreadsheet's built-in least-squares fitting routine) to fit a straight line to the data of Example 7.2 by the matrix method. We entered the data in columns labeled T, V, and σ_v and calculated component terms to be summed for $\boldsymbol{\beta_1}$, $\boldsymbol{\beta_2}$, and $\boldsymbol{\alpha_{11}}$, $\boldsymbol{\alpha_{12}}$, and $\boldsymbol{\alpha_{22}}$ in the labeled columns using the indicated equations. We summed each $\boldsymbol{\alpha}$ column to form the elements of the square matrix $\boldsymbol{\alpha}$, and the $\boldsymbol{\beta}$ columns to form the linear matrix $\boldsymbol{\beta}$. The spreadsheet's matrix-handling routines were applied to invert the $\boldsymbol{\alpha}$-matrix to form the $\boldsymbol{\epsilon}$-matrix, and to multiply $\boldsymbol{\epsilon}$ by $\boldsymbol{\beta}$ to find the parameter matrix $\mathbf{a}$. Uncertainties in the parameters were calculated from the square roots of the diagonal terms in the $\boldsymbol{\epsilon}$-matrix. Although we used absolute cell addresses to illustrate the procedure, we could have simplified the calculation by naming the arrays of cells and using the array-handling capabilities of the spreadsheet.

It may seem inefficient to write a program to solve such a simple problem, which most spreadsheets can handle with ease. However, there are advantages. First, it would be relatively easy to expand the program to fit more parameters, or to fit a series of functions more complicated than simple powers of the independent variable. A second advantage is that the solution provides the full error matrix. While most fitting programs should provide the uncertainties in the fitted parameters, the covariances may not be available. In some problems, they are essential.

We used *Quatro Pro* for this example, but the procedure with *Excel* is similar.

TABLE 7.4
Matrix solution by spreadsheet calculation for linear fit to data of Example 7.2

(*a*) Data and components of the matrix elements and sums

T(°C)	V (mV)	σ_v (mV)	β_1	β_2	α_{11}	$\alpha_{12} = \alpha_{21}$	α_{22}	Y_{calc}	χ^2
Column Equations			**V*1/σ²**	**V*T/σ²**	**1*1/σ²**	**1*T/σ²**	**T*T/σ²**	**$a_1 + a_2$*T**	**$[(Y - Y_{calc})/\sigma]^2$**
0	−0.849	0.050	−339.6	0	400	0	0	−0.947	3.81
20	−0.196	0.050	−78.4	−1,568	400	8,000	160,000	−0.101	3.61
40	0.734	0.050	293.6	11,744	400	16,000	640,000	0.745	0.04
60	1.541	0.050	616.4	36,984	400	24,000	1,440,000	1.590	0.97
80	2.456	0.050	982.4	78,592	400	32,000	2,560,000	2.436	0.16
100	3.318	0.050	1327.2	132,720	400	40,000	4,000,000	3.281	0.54
SUMS			2801.6	258,472	2400	120,000	8,800,000		9.13

(*b*) Matrices and fitted coefficients with uncertainties (*Quatro Pro* matrix algebra used to calculate ϵ and a)

$\boldsymbol{\alpha}$

$$\begin{bmatrix} 2400 & 120000 \\ 120000 & 8800000 \end{bmatrix}$$

$\boldsymbol{\epsilon}$

$$\begin{bmatrix} 1.310\text{E-}03 & -1.786\text{E-}05 \\ -1.786\text{E-}05 & 3.571\text{E-}07 \end{bmatrix}$$

$\boldsymbol{\beta}$

$$\begin{vmatrix} 2801.6 & 258472 \end{vmatrix}$$

$\mathbf{a}$

$$[-0.947 \quad 0.0423]$$

$\chi^2 = 9.13$

σ_a

$$\boxed{0.036 \quad 0.006}$$

7.3 INDEPENDENT PARAMETERS

Suppose we take the data of Example 6.1 or Example 6.2 and fit to them the quadratic polynomial function $y = a_1 + a_2x + a_3x^2$ as we did for Example 7.1. We should expect to find a rather small and possibly meaningless result for the coefficient a_3 of the quadratic term, but, because a_3 was not set equal to zero by definition, as in the analysis of Chapter 6, we might also find that the values of a_1 and a_2 have changed, sometimes considerably, from the values obtained in the linear fit. In general, the polynomial fitting procedure that we have considered will yield values for the coefficients that depend on the degree of the polynomial fitted to the data.

This interdependence arises from the fact that we have specified our coordinate system without regard to the region of parameter space from which our data points are extracted. The value of a_1 represents the intercept on the ordinal axis, the coefficient a_2 represents the slope at this same point, and other coefficients represent higher orders of curvature at this same intercept point. If the data are not clustered about this intercept point, its location might be highly dependent on the polynomial used to fit the data.

We might be able to extract more meaningful information about the data if we were to determine instead coefficients a_1', a_2', a_3', and so forth, which represent the

average value, the average slope, the average curvature, and so forth, of the data. Such coefficients would be independent of our choice of coordinate system and would represent physical characteristics of the data that are independent of the degree of the fitted polynomial.

Orthogonal Polynomials

We want to fit the data to a function that is similar to that of Equation (7.1) but that yields the desired independence of the coefficients. The appropriate function to use is the sum of orthogonal polynomials,[1] which has the form

$$y(x) = a_1 + a_2(x - \beta) + a_3(x - \gamma_1)(x - \gamma_2) + a_4(x - \delta_1)(x - \delta_2)(x - \delta_3) + \cdots \tag{7.28}$$

Following the development of Section 7.1, we must minimize χ^2 to determine the coefficients a_1, a_2, a_3, a_4, and so on, with the further criterion that the addition of higher-order terms to the polynomial will not affect the evaluation of lower-order terms. This criterion will be used to determine the extra parameters β, γ_1, γ_2, and so on.

The goodness-of-fit parameter χ^2 is defined as

$$\chi^2 \equiv \sum \left[\frac{\Delta y_i}{\sigma_i}\right]^2 = \sum \left[\frac{1}{\sigma_i^2}[y_i - y(x_i)]^2\right] \tag{7.29}$$

Setting the derivatives of χ^2 with respect to each of the m coefficients a_1, a_2, and so forth to 0 yields m simultaneous equations

$$\sum y_i = Na_1 + a_2\Sigma(x_i - \beta) + a_3\Sigma(x_i - \gamma_1)(x_i - \gamma_2) + a_4\Sigma(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3) + \cdots \tag{7.30}$$

$$\sum x_i y_i = a_1\Sigma x_i + a_2\Sigma x_i(x_i - \beta) + a_3\Sigma x_i(x_i - \gamma_1)(x_i - \gamma_2) + a_3\Sigma x_i(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3) + \cdots \tag{7.31}$$

$$\sum x_i^2 y_i = a_1\Sigma x_i^2 + a_2\Sigma x_i^2(x_i - \beta) + a_3\Sigma x_i^2(x_i - \gamma_1)(x_i - \gamma_2) + a_4\Sigma x_i^2(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3) + \cdots \tag{7.32}$$

$$\sum x_i^3 y_i = a_1\Sigma x_i^3 + a_2\Sigma x_i^3(x_i - \beta) + a_3\Sigma x_i^3(x_i - \gamma_1)(x_i - \gamma_2) + a_4\Sigma x_i^3(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3) + \cdots \tag{7.33}$$

where we have omitted a factor of σ_i^2 in the denominator for clarity.

[1] Any polynomial such as that of Equation (7.1) can be rewritten as a sum of orthogonal polynomials

$$y = a + \sum_{j=1}^{n} [b_j f_j(x_i)]$$

with the orthogonal property that $\Sigma[f_j(x_i)f_k(x_i)] = 0$ for $j \neq k$.

Additional Parameters

Let us examine Equation (7.30). If we restrict ourselves to a zeroth-degree polynomial, that is, to only one coefficient a_1, all the other coefficients are equal to 0 by definition. The coefficient a_1, therefore, is specified completely by the first term on the right-hand side of Equation (7.30):

$$a_1 = \frac{1}{N}\sum y_i = \bar{y} \tag{7.34}$$

If we restrict ourselves to a first-degree polynomial, the coefficient a_2 of the second term of Equation (7.30) is not 0. However, if a_1 is to be independent of the value of a_2, the second term itself must be 0. Hence, the requirement that

$$\sum (x_i - \beta) = 0$$

leads to the value for β,

$$\beta = \frac{1}{N}\sum x_i = \bar{x} \tag{7.35}$$

and a_2 can be determined directly from Equation (7.31) by substituting the values of a_1 and β with higher-order coefficients (a_3, a_4, etc.) set to 0.

Similarly, if we consider a quadratic function, the third term of Equation (7.30) must be 0 even when the coefficient a_3 is not 0. This constraint leads to a quadratic equation in γ_1 and γ_2 that is not sufficient to specify either parameter. We have the additional constraint, however, that the coefficient a_2 must be specified completely by Equations (7.30) and (7.31). Thus, the third term in both Equations (7.30) and (7.31) must be 0 regardless of the value of the coefficient a_3, and we have two simultaneous quadratic equations for the parameters γ_1 and γ_2,

$$\sum (x_i - \gamma_1)(x_i - \gamma_2) = 0 \quad \text{and} \quad \sum x_i(x_i - \gamma_1)(x_i - \gamma_2) = 0 \tag{7.36}$$

Similarly, the coefficient a_3 must be determined completely by Equation (7.32) (and the predetermined values of a_1 and a_2), and this constraint yields three simultaneous equations for the parameters δ_1, δ_2, and δ_3:

$$\sum (x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3) = 0$$
$$\sum x_i(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3) = 0 \tag{7.37}$$
$$\sum x_i^2(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3) = 0$$

The extrapolation to higher order is straightforward. (Note that these *additional parameters* are functions only of the independent variable x_i.)

Estimates of the Coefficients

Once the parameters β, γ, δ, and so on have been determined by the constraint equations, estimates of the coefficients a_1, a_2, and so on can be found from the resulting

$n + 1$ simultaneous equations. The value for the first coefficient a_1 is specified completely by minimizing χ^2 with respect to a_1 in Equation (7.30) and is given in Equation (7.34). The value of the second coefficient a_2 is determined by minimizing χ^2 with respect to both a_1 and a_2 in Equations (7.30) and (7.31) and substituting the value of a_1 from Equation (7.34) into Equation (7.31). Similarly, the value of a_3 can be determined from Equation (7.32) after substituting the values of a_1 and a_2 determined from Equations (7.30) and (7.31). Each succeeding equation yields a value for the next higher-order coefficient.

Note that the value determined for any coefficient is thus independent of the value specified for any higher-order coefficient, but is not independent of the value of lower-order coefficients. The parameters, representing our best estimates of the coefficients, are given by

$$
\begin{aligned}
a_1 &= \frac{1}{N}\sum y_i = \bar{y} \\
a_2 &= \frac{\sum y_i(x_i - \beta)}{\sum (x_i - \beta)^2} \\
a_3 &= \frac{\sum y_i(x_i - \gamma_1)(x_i - \gamma_2)}{\sum [(x_i - \gamma_1)(x_i - \gamma_2)]^2} \\
a_4 &= \frac{\sum y_i(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3)}{\sum [(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3)]^2}
\end{aligned}
\tag{7.38}
$$

and so forth.

Simplification

For the general case of arbitrarily chosen data points (x_i, y_i), this procedure is cumbersome even with computer techniques because it requires the solution of coupled, nonlinear equations. There is, however, a special type of data for which the calculations can be considerably simplified, namely, data that meet the following two criteria: (1) the independent variables x_i are equally spaced, and (2) the uncertainties are constant, $\sigma_i = \sigma$, and can therefore be ignored.

Consider the experiments of Examples 6.1 (measurement of temperature versus position) and 7.1 (voltage versus temperature). Those data satisfy the required conditions and, therefore, we could use a simplified method of independent parameters to obtain a fit. The resulting values of the coefficients for these particular experiments might not have any great physical significance (that is, $a_1 = \bar{T}$ the average temperature of the data points in Example 6.1 is not a particularly useful number), but by using this technique of fitting orthogonal polynomials we could try fitting higher-degree polynomials without changing the values of the coefficients already calculated for a straight-line or quadratic fit. The experiment of Example 6.2 (the decay of a radioactive state) fulfills only the first of the two criteria, because the x data points are equally spaced but the uncertainties are statistical, so that we cannot ignore the factor of σ_i^2 that belongs in the denominators of the fitting Equations (7.30) through (7.33).

For an experiment similar to that of Example 7.1, where we have made N measurements of equally spaced values of the independent variable x ranging from x_1 to x_N in steps of Δ,

$$\Delta = x_{i+1} + x_i$$

and the uncertainties are due to instrumental errors with a common standard deviation $\sigma_i = \sigma$, Equations (7.35) through (7.37) reduce to

$$
\begin{aligned}
\beta &= \frac{1}{N}\sum x_i = \bar{x} = \frac{1}{2}(x_i - x_n) \\
\gamma &= \beta \pm \sqrt{\frac{1}{N}\sum (x_i - \beta)^2} = \beta \pm \Delta\sqrt{\frac{1}{12}(N^2 - 1)} \\
\delta &= \beta,\ \beta \pm \sqrt{\frac{\sum [x_i(x_i - \beta)^3]}{\sum [x_i(x_i - \beta)]}} = \beta,\ \beta \pm \Delta\sqrt{\frac{1}{20}(3N^2 - 7)}
\end{aligned}
\tag{7.39}
$$

A more comprehensive list of parameters for orthogonal polynomials can be found in Anderson and Houseman (1942).

Table 7.5 shows coefficients a_1, a_2, a_3, and a_4 as well as the values of χ^2 and the χ^2-probability obtained when we fit the data of Example 7.1, by the standard least-squares method and by the independent parameter method of Equation (7.39). We have made separate fits with first-, second-, and third-degree polynomials ($m = 2$, 3, and 4). As expected, adding extra terms does not change the values of the lower-order coefficients obtained by the independent parameter method and therefore we display them only once in Table 7.5. There is a marked improvement in χ^2 in going from the two-parameter (linear) fits to three-parameter (quadratic) fits. Unless a theoretical reason dictates that our data should follow a cubic distribution, there is no justification in making a four-parameter (cubic) fit to these data, because the value of χ^2 for $m = 3$ is satisfactory (26.6 for 18 degrees of freedom,

TABLE 7.5
Values of χ^2 and parameters obtained by fitting the data of Example 7.1 by the standard least-squares method and by the method of independent parameters, as a function of the number of parameters m of the fit

	Standard least squares			Independent parameters
m	2	3	4	
χ^2	43.5 (0.12%)	26.6(8.8%)	24.9(9.4%)	
a_1	-1.01 ± 0.02	(-0.92 ± 0.03)	(-0.89 ± 0.03)	1.15
a_2	$(4.31 \pm 0.04)10^{-2}$	$(3.8 \pm 0.1)10^{-2}$	$(3.4 \pm 0.3)10^{-2}$	4.31×10^{-2}
a_3		$(5.5 \pm 1.3)10^{-5}$	$(1.5 \pm 0.8)10^{-4}$	5.49×10^{-5}
a_4			$(-6.5 \pm 5.1)10^{-7}$	6.51×10^{-7}

Note: The values of χ^2 are the same for both methods. The numbers in parentheses correspond to the χ^2 probability for the fit with 21-m degrees of freedom.

corresponding to $P = 8.8\%$), and adding more terms does not improve the fit. If a cubic function had been predicted by theoretical considerations, we should be obligated to say that our data are not sensitive to the presence of a cubic term.

Legendre Polynomials

Although the method of fitting to orthogonal polynomials outlined in the previous section can be tedious, there are predefined sets of orthogonal polynomials that are often useful in fitting data. One important set is the *Legendre polynomials*

$$y(x) = a_0P_0(x) + a_1P_1(x) + \cdots = \sum_{L=0}^{m-1}[a_LP_L(x)] \tag{7.40}$$

where $x = \cos\theta$ and the terms $P_L(x)$ in the function are given by

$$\begin{aligned} P_0(x) &= 1 & P_2(x) &= \tfrac{1}{2}(3x^2 - 1) \\ P_1(x) &= x & P_3(x) &= \tfrac{1}{2}(5x^3 - 3x) \end{aligned} \tag{7.41}$$

and higher-order terms can be determined from the recursion relation

$$P_L(x) = \frac{1}{L}[(2L-1)xP_{L-1}(x) - (L-1)P_{L-2}(x)] \tag{7.42}$$

Legendre polynomials are orthogonal when averaged over all values of $x = \cos\theta$:

$$\int_{-1}^{1}[P_L(x)P_M(x)\,dx = \begin{cases} 2/(2L+1) & \text{for } L = M \\ 0 & \text{for } L \neq M \end{cases} \tag{7.43}$$

Example 7.3. Let us consider an experiment in which ^{13}C is bombarded by 4.5-MeV protons. In the subsequent reaction, some of the protons are captured by the ^{13}C nucleus, which then decays by gamma emission, producing gamma rays with energies up to 11 MeV. A measurement of the angular distribution of the emitted gamma rays gives information about the angular momentum states of the energy levels in the residual nucleus ^{14}N.

Table 7.6 lists simulated data for this experiment. Gamma ray counts were recorded at 17 angles from 0 to 160°. Columns 1 through 4 list the angles at which the measurements were made, the cosine of the angle ($x = \cos\theta$), the measured number of counts (C_i), and the uncertainties σ_{C_i} in the counts. The uncertainties are assumed to be purely statistical. These data are plotted in Figure 7.2 as a function of the angle θ. There appears to be symmetry around $\theta = 90°$, and consideration of the reaction process predicts that the data should be described by a fourth-order Legendre polynomial with only even terms:

$$C = a_0P_0(x) + a_2P_2(x) + a_4P_4(x) \quad \text{with } x = \cos\theta \tag{7.44}$$

Let us apply the matrix method of least squares of Section 7.2 to this problem to fit the series of Legendre polynomials of Equation (7.41) to these data. We shall first fit a fourth-order Legendre polynomial that includes both odd and even terms. The fitting function is of the form

$$y(x) = a_0f_0(x) + a_1f_1(x) + \cdots + a_{m-1}f_{m-1}(x) \tag{7.45}$$

which is linear in the fitting parameters, the coefficients a_i.

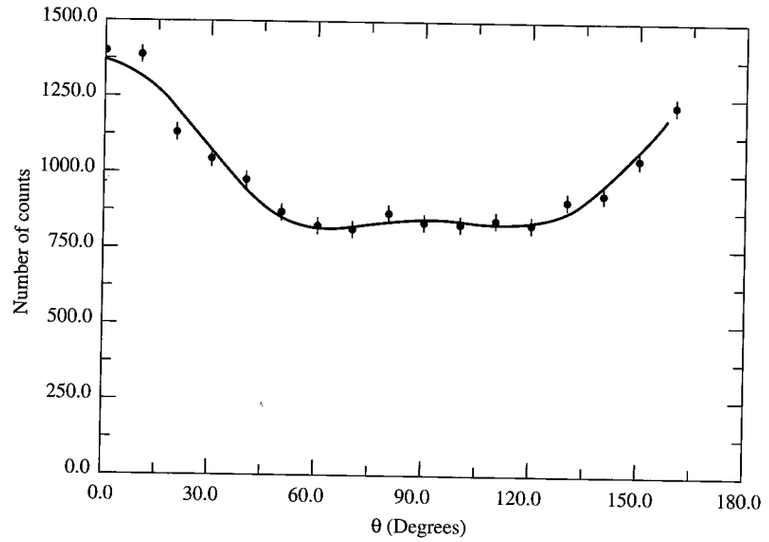


FIGURE 7.2
Angular distribution of gamma rays emitted from the simulated reaction $^{13}C(p, \gamma)^{14}N$ produced by incident protons at $E_p = 4.5$ MeV (Example 7.3). The calculated curve represents a fit to the data of a series of even Legendre polynomials up to $L = 4$. Statistical uncertainties were assumed.

TABLE 7.6
Angular distribution of gamma rays emitted in the reaction $^{13}C(p, \gamma)^{14}N$ produced by incident protons at $E_P = 4.5$ MeV

θ (degrees)	$x = \cos\theta$	C_i counts	σ_{C_i}	Y_i all terms	Y_i even term
0	1.000	1400	37.4	1365.8	1361.3
10	0.985	1386	37.2	1325.2	1321.1
20	0.940	1130	33.6	1217.0	1213.9
30	0.866	1045	32.3	1075.8	1074.5
40	0.766	971	31.2	943.5	944.4
50	0.643	862	29.4	852.5	855.6
60	0.500	819	28.6	813.9	818.6
70	0.342	808	28.4	816.9	821.9
80	0.174	862	29.4	836.5	840.2
90	0.000	829	28.2	848.6	849.6
100	−0.174	824	28.7	842.8	840.2
110	−0.342	839	29.0	827.5	821.9
120	−0.500	819	28.6	825.4	818.6
130	−0.643	901	30.0	861.0	855.6
140	−0.766	925	30.4	945.7	944.4
150	−0.866	1044	32.3	1069.8	1074.5
160	−0.940	1224	35.0	1202.9	1213.9

Note: The calculated numbers of counts were obtained from least-squares Legendre polynomials fits to the data of the form $Y_i(x) = \sum_{L=1}^{5} a_LP_{L-1}(x_i)$, for separate fits with all terms and with even terms only.

Computer fits Routines used for fitting a series of Legendre polynomials to these data are included in Program 7.1. The procedure LEGPOLY in the program unit FITFUNC7 calculates the terms of the Legendre polynomials through tenth order. The procedure is selected through a branch on the variable PAE in the function Funct with PAE = 'A' for all terms to order $n = m - 1$, or PAE = 'E' to fit with just the even terms. Note that the index k of the term in the fitting function, in general, does not correspond to the order L of the Legendre polynomial.

The efficiency of the calculation (and therefore the speed of the linear regression calculation) could be improved in a number of ways. The simplest change would be to calculate the functions once at each value of the independent variable and store the calculated values in an array.

Parameters obtained by fitting a series in Legendre polynomials for terms up to $L = 4$ are listed in Table 7.7. Separate fits were made with all terms and with only the even terms in the series. As expected, the coefficients of terms involving odd orders are comparable to their uncertainties and negligible compared to those involving even poynomials. The full error matrix for the fit with even terms is listed in Table 7.8.

In view of the strong theoretical argument that only even Legendre polynomials are required for this reaction, it would be appropriate to fit a series that includes only the even terms. The parameters obtained in this fit are also displayed in Table 7.7, and the numbers of counts calculated from these parameters are listed. The function calculated with even terms is illustrated as a curve on the data of Figure 7.2.

Because we are fitting with orthogonal functions, we might have expected to obtain identical values for the coefficient a_0 from both fits. (We expect the higher-order even coefficients to change because the presence or absence of lower-order coefficients must affect the higher coefficients.) The fact that there is some dependence of a_0 on higher-order terms is a result of the fact that a given experiment does not sample uniformly the entire range of the Legendre polynomial, so the orthogonality relation Equation (7.43) is not satisfied by a finite data set. This is in contrast to the situation in the previous section, where we set up orthogonal functions based on the data themselves. Nevertheless, it is generally good practice to use orthogonal

TABLE 7.7
Coefficients and χ^2 from least-squares fit to Legendre polynomial series

	χ^2	a_0	a_1	a_2	a_3	a_4
All terms	17.2(14%)	937.4 ± 7.6	0.7 ± 12.8	259 ± 14	10 ± 17	158 ± 18
Even terms	17.6(22%)	938.1 ± 7.5		261 ± 14		161 ± 16

TABLE 7.8
Error matrix for a least-squares fit to even Legendre polynomials

$$\begin{bmatrix} 56.24 & -5.256 & -6.272 \\ 5.256 & 186.5 & -26.90 \end{bmatrix}$$

fitting functions whenever possible to minimize both the correlations between coefficients and the dependence of higher coefficients on the presence of lower ones.

The values of χ^2 and the χ^2-probability for the two fits are also given in Table 7.7. We note that χ^2 for the three-parameter fit is necessarily higher than that for the five-parameter fit, but χ^2 per degree of freedom is smaller and the χ^2-probability is higher.

7.4 NONLINEAR FUNCTIONS

In all the procedures developed so far we have assumed that the fitting function was linear in the coefficients. By that we mean that the function can be expressed as a sum of separate terms each multiplied by a single coefficient. How can we fit data with a function that is not linear in the coefficients? For example, suppose we have measured the distribution of decay times of an unstable state and that the distribution can be represented by the normalized function $P(t) = (1/\tau)e^{-t/\tau}$, where τ is the mean lifetime of the state. Can we find the parameter τ by the least-squares method? The method of least squares does not yield a straightforward analytical solution for such functions. In Chapter 8 we investigate methods of searching parameter space for values of the coefficients that will minimize the goodness-of-fit criterion χ^2. Here we consider approximate solutions to such problems using linear-regression techniques.

Linearization

It is possible to transform some functions into linear functions. For example, if we were to fit an exponential decay problem of the form

$$y = ae^{-bx} \tag{7.46}$$

where a and b are the unknown parameters, it would seem reasonable to take logarithms of both sides and to fit the resulting straight line equation

$$\ln y = \ln a - bx \tag{7.47}$$

The method of least squares minimizes the value of χ^2 with respect to each of the coefficients $\ln a$ and $\ln b$ where χ^2 is given by

$$\chi^2 = \sum \left\{ \frac{1}{\sigma_i'^2} [\ln y_i + \ln a - bx_i]^2 \right\} \tag{7.48}$$

where we must use weighted uncertainties σ_i' instead of σ_i to account for the transformation of the dependent variable:

$$\sigma_i' = \frac{d(\ln y_i)}{dy} \sigma_i = \frac{1}{y_i} \sigma_i \tag{7.49}$$

The importance of weighting the uncertainties is illustrated in Figure 7.3, which shows the function of Equation (7.46) graphed both on a linear and on a logarithmic

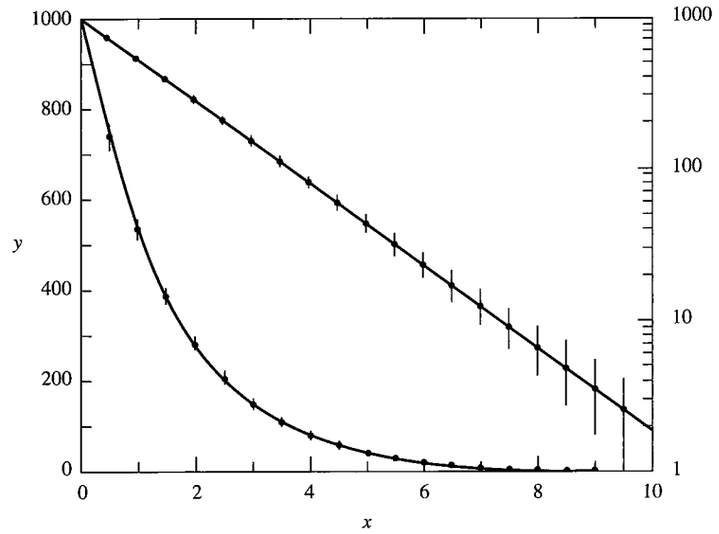


FIGURE 7.3
Graph of the function $y = ae^{-bx}$ calculated on a linear and a logarithmic scale. The error bars are given by $\sigma_i = \sqrt{y_i}$. The curved line corresponds to the linear scale on the left, and the straight line to the logarithmic scale on the right.

However, on the logarithmic scale, they appear to decrease with increasing y_i and are very large for very small $\ln y_i$. If we were to ignore this effect in fitting Equation (7.47), we would overemphasize the uncertainties for small values of y_i.

In general, if we fit the function $f(y)$ rather than y, the uncertainties σ_i in the measured quantities must be modified by

$$\sigma_i' = \frac{df(y)}{dy_i}\sigma_i \tag{7.50}$$

Errors in the Parameters

If we modify the fitting function so that instead of fitting the data points y_i with the coefficient $a, b, \ldots$, we fit modified data points $y_i' = f(y_i)$ with coefficients $a', b', \ldots$, then our estimates of the errors in the coefficients will pertain to the uncertainties in the modified coefficients $a', b', \ldots$, rather than to the desired coefficients $a, b, \ldots$. If the relationship between the two sets of coefficients is defined to be

$$a' = f_a(a) \qquad b' = f_b(b) \tag{7.51}$$

then the correspondence between the uncertainties $\sigma_a', \sigma_b', \ldots$ in the modified coefficients and the uncertainties $\sigma_a, \sigma_b, \ldots$ in the desired coefficients is obtained in a manner similar to that for σ_i' and σ_i in Equation (7.50):

$$\sigma_a' = \frac{df_a(a)}{da}\sigma_a \qquad \sigma_b' = \frac{df_b(b)}{db}\sigma_b \tag{7.52}$$

Thus, if the modified coefficient is $a' = \ln a$, the estimated error in a is determined from the estimated error in a', according to Equation (7.52) with $f_a = \ln a$:

$$\sigma_a' = \frac{d(\ln a)}{da}\sigma_a = \frac{\sigma_a}{a} \tag{7.53}$$

Values of χ^2 for testing the goodness of fit should be determined from the original uncertainties of the data σ_i and from the unmodified equation, although Equation (7.48) should give approximately equivalent results when weighted with the modified uncertainties σ_i'.

In Example 6.2, we considered an experiment to check the decrease in the number of counts C as a function of distance r from a radiative source. We expected a relation of the form

$$C(r) = b/r^2 \tag{7.54}$$

and therefore changed the independent variable to $x = 1/r^2$ and fitted a straight line to the C versus x data. Because uncertainties were assigned only to the dependent variable C, the fit was not distorted by that transformation.

Suppose, instead, that our objective had been to determine the exponent a in the expression for C:

$$C(r) = br^{-a} \tag{7.55}$$

Taking logarithms of both sides, we obtain the linear equation

$$\ln(C) = \ln(b) - a \ln r$$

or

$$C_i = b' - ar' \tag{7.56}$$

with $C' = \ln C$, $r' = \ln r$, and $b' = \ln b$. The uncertainties σ' in C' would be given by Equation (7.49) as

$$\sigma' = \sigma/C$$

and we could find the exponent a by fitting a straight line to Equation (7.56) using these weighted uncertainties.

Although the method of taking logarithms of an exponential or a power function to produce a function that is linear in the parameters may be convenient for quick estimates, with fast computers it is generally better to solve such problems by one of the approximation methods developed for fitting nonlinear functions. These methods will be explored in Chapter 8.

SUMMARY

Linear function: Function that is linear in its parameter a_k:

$$y(x) = \sum_{k=1}^{m} a_k f_k(x)$$

Least-squares fit to a function that is linear in its parameters:

$$\Delta = \begin{vmatrix} \sum \frac{f_1(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_3(x_i)}{\sigma_i^2} & \cdots \\ \sum \frac{f_2(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_3(x_i)}{\sigma_i^2} & \cdots \\ \sum \frac{f_3(x_i)f_1(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_3(x_i)}{\sigma_i^2} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

$$a_1 = \frac{1}{\Delta} \begin{vmatrix} \sum y_i \frac{f_1(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_1(x_i)f_3(x_i)}{\sigma_i^2} & \cdots \\ \sum y_i \frac{f_2(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_2(x_i)f_3(x_i)}{\sigma_i^2} & \cdots \\ \sum y_i \frac{f_3(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_2(x_i)}{\sigma_i^2} & \sum \frac{f_3(x_i)f_3(x_i)}{\sigma_i^2} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

For the jth coefficient, a_j is found by replacing the jth column in the expression for Δ with the first column in the expression for a_1.

Chi square:

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{1}{\sigma} [y_i - y(x_i)] \right]^2 = \sum_{i=1}^{N} \left[\frac{1}{\sigma} \left[y_i - \sum_{k=1}^{m} a_k f_k \right] \right]^2$$

Sample variance σ^2:

$$\sigma^2 \simeq s^2 = \frac{1}{N-m} \sum_{i=1}^{N} [y_i - y(x_i)]^2$$

Matrix solution: $\mathbf{a} = \boldsymbol{\beta}\boldsymbol{\epsilon} = \boldsymbol{\beta}\boldsymbol{\alpha}^{-1}$ where $\mathbf{a}$ is a linear matrix of the coefficients and

$$\beta_k \equiv \sum \left[\frac{1}{\sigma_i^2} y_i f_k(x_i) \right]$$

$$\alpha_{lk} \equiv \sum \left[\frac{1}{\sigma_i^2} f_l(x_i) f_k(x_i) \right]$$

Error or variance matrix: The diagonal elements of the square matrix $\boldsymbol{\epsilon} = \boldsymbol{\alpha}^{-1}$ are the variances of the parameters a_k and the off-diagonal elements are the covariances:

$$\sigma_{a_l}^2 = \epsilon_{ll} \qquad \sigma_{a_{lk}}^2 = \epsilon_{lk}$$

Orthogonal polynomials:

$$y(x) = a_1 + a_2(x - \beta) + a_3(x - \gamma_1)(x - \gamma_2) + a_4(x - \delta_1)(x - \delta_2)(x - \delta_3) + \cdots$$

$$a_1 = \bar{y} \qquad a_2 = \frac{\Sigma y_i (x_i - \beta)}{\Sigma (x_i - \beta)^2}$$

$$a_3 = \frac{\Sigma y_i (x_i - \gamma_1)(x_i - \gamma_2)}{\Sigma [(x_i - \gamma_1)(x_i - \gamma_2)]^2} \qquad a_4 = \frac{\Sigma y_i (x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3)}{\Sigma [(x_i - \delta_1)(x_i - \delta_2)(x_i - \delta_3)]^2}$$

For equally spaced values of x, $x_{i+1} - x_i = \Delta$,

$$\beta = \tfrac{1}{2}(x_i + x_N) \qquad \gamma = \beta \pm \Delta \sqrt{\tfrac{1}{12}(N^2 - 1)}$$

$$\delta = \beta, \beta \pm \Delta \sqrt{\tfrac{1}{20}(3N^2 - 7)}$$

Legendre polynomials:

$$y(x) = \sum_{L=1}^{m-1} [a_L P_L(x)]$$

$$P_0(x) = 1 \qquad P_1(x) = x$$

$$P_L(x) = \frac{1}{L}[(2L-1)xP_{L-1}(x) - (L-1)P_{L-2}(x)] \quad \text{(recursion relation)}$$

Nonlinear functions:

If $y_i' = f(y_i)$, then

$$\sigma_i' = \frac{df(y)}{dy_i} \sigma_i$$

and if $a' = f_a(a)$ and $b' = f_b(b)$, then

$$\sigma_a' = \frac{df_a(a)}{da} \sigma_a \qquad \sigma_b' = \frac{df_b(b)}{db} \sigma_b$$

EXERCISES

7.1. Show by direct calculation using the data of Example 7.2 listed in Table 7.2 that $\boldsymbol{\alpha}\boldsymbol{\epsilon} = \mathbf{1}$ where $\mathbf{1}$ is the unity matrix.

7.2. The tabulated data represent the lower bin limit x and the bin contents y of a histogram of data that fall into two peaks.

i	1	2	3	4	5	6	7	8	9	10
x_i	50	60	70	80	90	100	110	120	130	140
y_i	5	7	11	13	21	43	30	16	15	10
i	11	12	13	14	15	16	17	18	19	20
x_i	150	160	170	180	190	200	210	220	230	240
y_i	13	42	90	75	29	13	8	4	6	3

Use the method of least squares to find the amplitudes a_1 and a_2 and their uncertainties by fitting to the data the function

$$y(x) = a_1 L(x; \mu_1, \Gamma_1) + a_2 L(x; \mu_2, \Gamma_2)$$

with $\mu_1 = 102.1$, $\Gamma_1 = 30$, $\mu_2 = 177.9$, and $\Gamma_2 = 20$. The function $L(x; \mu, \Gamma)$ is the Lorentzian function of Equation (2.32). Assume statistical uncertainties ($\sigma_i = \sqrt{y_i}$). Find χ^2 for the fit and the full error matrix.

7.3. From the parameters listed in Table 7.7 for the fit of even terms to the data of Example 7.3, determine the predicted value of the cross section for $\theta = 90°$ and its uncertainty. Calculate the uncertainty from the diagonal errors, listed in Table 7.7 and from the full error matrix listed in Table 7.8 and compare the two results.

7.4. Fit fourth-degree power series polynomials instead of Legendre polynomials to the data of Example 7.3. Let $x = \cos\theta$ and fit a polynomial with all terms to x^4 and another polynomial with only the even terms. Compare your results to those obtained from the fit to Legendre polynomials displayed in Table 7.7.

7.5. Derive the expression for γ_1 and γ_2 of Equation (7.36).

7.6. Derive an expression for $P_4(\cos\theta)$. [See Equation (7.42).]

7.7. Show by direct integration that $P_0(x)$, $P_1(x)$, and $P_2(x)$ are orthogonal and obey Equation (7.43).

7.8. In an experiment to measure the angular distribution of elastically scattered particles, a beam of particles strikes a liquid hydrogen target and counts are recorded at selected angles to the direction of the incident beam. Measurements are made both with the target filled with liquid hydrogen *(full target)* and with an empty target *(empty target)*. The empty-target measurements were made with one-half the number of incident particles used for the full-target signal. By subtracting the suitably scaled empty-target signal from the full-target signal, the angular distribution of scattering on pure hydrogen can be determined.

Assume that the following data were obtained in such an experiment. Uncertainties in the numbers of counts are statistical.

$\cos\theta$ (lower limit)	−1.0	−0.8	−0.6	−0.4	−0.2	0.0	0.2	0.4	0.6	0.8
$\cos\theta$ (upper limit)	−0.8	−0.6	−0.4	−0.2	0.0	0.2	0.4	0.6	0.8	1.0
Counts, full target	184	128	99	49	53	55	70	81	136	216
Counts, empty target	5	4	4	1	3	1	4	9	8	7

(*a*) Scale the empty-target data to the same number of incident antiprotons used in recording the full-target data and make a subtraction to obtain the number of interactions on the hydrogen. Pay particular attention to the uncertainties in the difference.

(*b*) Use the least-squares method to fit the function

$$y(x) = a_1 P_0(x) + a_2 P_1(x) + a_3 P_2(x)$$

to the subtracted data, to obtain the coefficients a_1, a_2, and a_3, where the functions $P_L(x)$ are the Legendre polynomials defined in Equation (7.41).

7.9. Follow the procedure outlined in Section 7.4 to find the exponent a in Equation (7.55), using the data of Example 6.2 (Table 6.2).

7.10. A 1-m-long plastic plate with rulings at 10-cm intervals is dropped through a photogate to measure the acceleration of gravity g in an undergraduate laboratory experiment. The time is recorded as each ruling passes through the gate. The passage of the first ruling starts the timer. Data from such an experiment are tabulated. The recorded time is related to the distance that the ruler has fallen by $y = y_0 - v_0 t - 1/2gt^2$. Note that neither the initial height y_0 nor the initial speed v_0 are known.

Ruling #	0	1	2	3	4	5	6	7	8	9	10
Time(s)	0.000	0.079	0.132	0.174	0.212	0.244	0.271	0.301	0.325	0.349	0.373

Use the least-squares method with a second-degree polynomial to find g and its uncertainty. Measure y from the photogate so that you can set $y = 0$ when ruling #0 passes the gate, $y = 1$ when ruling #1 passes, and so forth. Choose t as the independent and y as the dependent variable. Assume a uniform uncertainty in t of 0.001 s and a negligible uncertainty in y. Because the uncertainty is in the independent variable, it must be transformed to the dependent variable by the method discussed in Section 6.1. This will require initial estimates of g and v_0. After the fit has been made you may wish to repeat the fit using estimates of g and v_0 from the previous fit to improve the results.

CHAPTER 8

LEAST-SQUARES FIT TO AN ARBITRARY FUNCTION

8.1 NONLINEAR FITTING

The methods of least squares and multiple regression developed in the previous chapters are restricted to fitting functions that are linear in the parameters as in Equation (7.3):

$$y(x) = \sum_{j=1}^{m} [a_j f_j(x)] \tag{8.1}$$

This limitation is imposed by the fact that, in general, minimizing χ^2 can yield a set of coupled equations that are linear in the m unknown parameters only if the fitting functions $y(x)$ are themselves linear in the parameters. We shall distinguish between the two types of problems by referring to *linear fitting* for problems that involve equations that are linear in the parameters, such as those discussed in Chapters 6 and 7, and *nonlinear fitting* for those problems that are nonlinear in the parameters.

Example 8.1. In a popular undergraduate physics laboratory experiment, a real silver quarter is irradiated with thermal neutrons to create two short-lived isotopes of silver, $_{47}Ag^{108}$ and $_{47}Ag^{110}$, that subsequently decay by beta emission. Students count the emitted beta particles in 15-s intervals for about 4 min to obtain a decay curve. Data collected from such an experiment are listed in Table 8.1 and plotted on a semi-

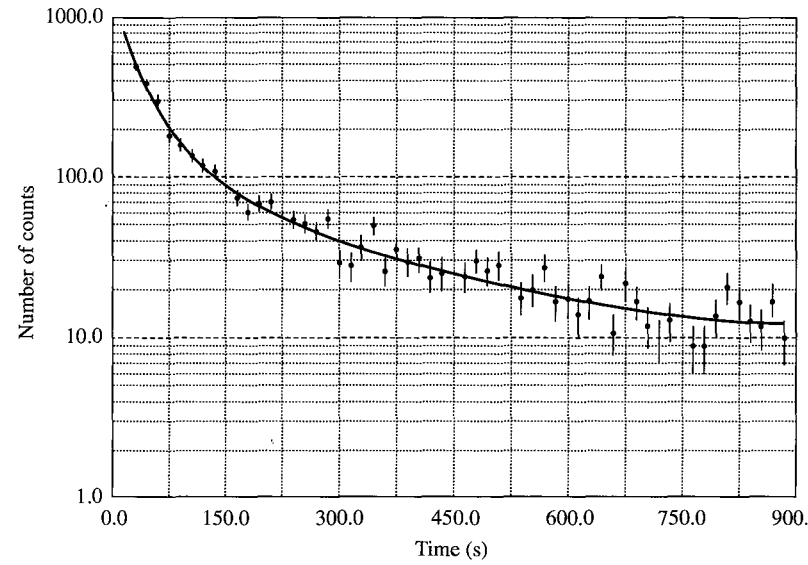


FIGURE 8.1
Number of counts detected from the decay of two excited states of silver as a function of time (Example 8.1). Time is reported at the end of each interval. Statistical uncertainties are assumed. The curve was obtained by a nonlinear least-squares fit of Equation (8.2) to the data.

line because the probability function that describes the process is the sum of two exponential functions plus a constant background. We can represent the decay by the fitting function

$$y(x_i) = a_1 + a_2 e^{-t/a_4} + a_3 e^{-t/a_5} \tag{8.2}$$

where the parameter a_1 corresponds to the background radiation and a_2 and a_3 correspond to the amplitudes of the two excited states with mean lives a_4 and a_5, respectively. Clearly, Equation (8.2) is not linear in the parameters a_4 and a_5, although it is linear in the parameters a_1, a_2, and a_3.

We can use a graphical analysis method to find the two mean lifetimes by plotting the data on semilogarithmic paper after first subtracting from each data point the constant background contribution, which has been measured separately. (Note that the background counts have not been subtracted in Figure 8.1.) We then consider two regions of the plot: region a, at small values of T (e.g., $T < 120$ s) in which the short-lived state dominates the plot, and region b, at large values of T (e.g., $T > 200$ s) in which only the long-lived state contributes to the data. We can estimate the mean lifetime of the long-lived state by finding the slope of our best estimate of the straight line that passes through the data points in region b. From this result we can estimate the contribution of the long-lived component to region a and subtract that contribution from each of the data points, and thus make a new plot of the number of counts in region a, which we attribute to the short-lived state alone. The slope of the line through the corrected points gives us the mean lifetime of the short-lived state. Linear regres-

TABLE 8.1
Geiger counter data from an irradiated silver piece, recorded in 15-s intervals

Point number	Time	Measured counts	Calculated counts	Point number	Time	Measured counts	Calculated counts
1	15	775	748.3	31	465	24	24.0
2	30	479	519.8	32	480	30	23.0
3	45	380	370.4	33	495	26	22.1
4	60	302	272.0	34	510	28	21.3
5	75	185	206.7	35	525	21	20.5
6	90	157	162.7	36	540	18	19.8
7	105	137	132.5	37	555	20	19.2
8	120	119	111.5	38	570	27	18.5
9	135	110	96.3	39	585	17	18.0
10	150	89	85.0	40	600	17	17.4
11	165	74	76.5	41	615	14	16.9
12	180	61	69.7	42	630	17	16.5
13	195	66	64.2	43	645	24	16.0
14	210	68	59.5	44	660	11	15.6
15	225	48	55.5	45	675	22	15.2
16	240	54	51.9	46	690	17	14.9
17	255	51	48.8	47	705	12	14.6
18	270	46	45.9	48	720	10	14.3
19	285	55	43.3	49	735	13	14.0
20	300	29	40.9	50	750	16	13.8
21	315	28	38.7	51	765	9	13.5
22	330	37	36.7	52	780	9	13.3
23	345	49	34.8	53	795	14	13.1
24	360	26	33.1	54	810	21	12.9
25	375	35	31.5	55	825	17	12.7
26	390	29	30.0	56	840	13	12.6
27	405	31	28.6	57	855	12	12.4
28	420	24	27.3	58	870	18	12.3
29	435	25	26.1	59	885	10	12.1
30	450	35	25.0				

Note: The time is reported at the end of each interval. The calculated number of counts was found by method 4.

Because analytic methods of least-squares fitting cannot be used for nonlinear fitting problems, we must consider approximation methods and make searches of parameter space. In the following sections we discuss four nonlinear fitting methods: a simple grid-search method in which we simply calculate χ^2 at trial values of the parameters, and search for those values of the parameters that yield a minimum value of χ^2, a gradient-search method that uses the slope of the function to improve the efficiency of the search, and two semianalytic methods that make use of the matrix method developed in Chapter 7, with a linear approximation to the nonlinear functions. As examples, we shall determine the parameters $(a_1 \ldots a_5)$ by fitting Equation (8.2) to the data of Example 8.1 using each of the four methods. The curve

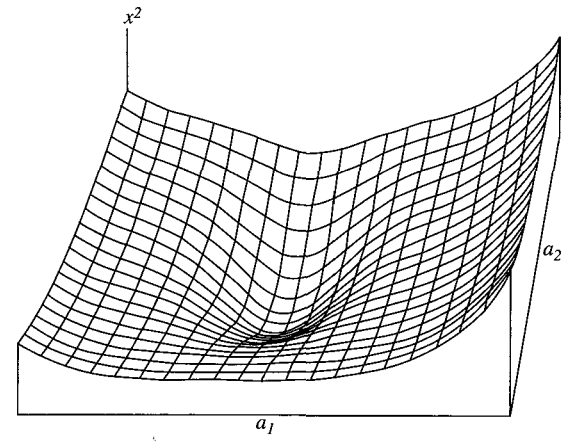


FIGURE 8.2
Chi-square hypersurface as a function of two parameters.

Method of Least Squares

We can generalize the probability function, or *likelihood function,* of Equation (6.7) to any number of parameters,

$$P(a_1, a_2, \ldots, a_m) = \prod\left[\frac{1}{\sigma_i\sqrt{2\pi}}\right]\exp\left\{-\frac{1}{2}\sum\left[\frac{y_i - y(x_i)}{\sigma_i}\right]^2\right\} \tag{8.3}$$

and, as in the previous chapters, maximize the likelihood with respect to the parameters by minimizing the exponent, or the goodness-of-fit parameter χ^2:

$$\chi^2 \equiv \sum\left\{\frac{1}{\sigma_i^2}[y_i - y(x_i)]^2\right\} \tag{8.4}$$

where x_i and y_i are the measured variables, σ_i is the uncertainty in y_i, and $y(x_i)$ are values of the function calculated at x_i. According to the method of least squares, the optimum values of the parameters a_j are obtained by minimizing χ^2 simultaneously with respect to each parameter,

$$\begin{aligned}\frac{\partial\chi^2}{\partial a_j} &= \frac{\partial}{\partial a_j}\sum\left\{\frac{1}{\sigma_i^2}[y_i - y(x_i)]^2\right\} = 0 \\ &= -2\sum\left\{\frac{1}{\sigma_i^2}[y_i - y(x_i)]\frac{\partial y(x_i)}{\partial a_j}\right\}\end{aligned} \tag{8.5}$$

Taking partial derivatives of χ^2 with respect to each of the m parameters a_j will yield m coupled equations in the m unknown parameters a_j as in Section 7.1. If these equations are not linear in all the parameters, we must, in general, treat χ^2 as a continuous function of the m parameters, describing a hypersurface in an m-dimensional space,

value of χ^2. Figure 8.2 illustrates such a hyperspace for a function of two parameters. Alternatively, we may apply to the m equations obtained from Equations (8.5) approximation methods developed for finding roots of coupled, nonlinear equations. A combination of both methods is often used.

Variation of χ^2 Near a Minimum

For a sufficiently large event sample, the likelihood function becomes a Gaussian function of each parameter centered on those values a_j' that minimize χ^2:

$$P(a_j) = Ae^{-(a_j - a_j')^2/2\sigma_j^2} \tag{8.6}$$

where A is a function of the other parameters, but not of a_j. Comparing Equation (8.3) for the likelihood function with Equation (8.4) for χ^2, we observe that we can express χ^2 as

$$\chi^2 = -2\ln[P(a_1, a_2, \ldots, a_m)] + 2\sum \ln(\sigma_i\sqrt{2\pi}) \tag{8.7}$$

Then, from Equation (8.6), we can write

$$\chi^2 = \frac{(a_j - a_j')^2}{\sigma_j^2} + C \tag{8.8}$$

to show the variation of χ^2 with any single parameter a_j in the vicinity of a minimum with respect to that parameter. The constant C is a function of the uncertainties σ_i and the parameters a_k for $k \neq j$. Thus χ^2 varies as the square of distance from a minimum, and an increase of 1 standard deviation (σ) in the parameter from the value a_j' at the minimum increases χ^2 by 1. For a more general proof, see Arndt and MacGregor (1966), appendix II.

We can see that this result is consistent with that obtained from a second-order Taylor expansion of χ^2 about the values a_j', where the values of χ^2 and its derivatives at $a = a'$ are written as χ_0^2, $\partial\chi_0^2/\partial a_j$, and so forth:

$$\chi^2 \simeq \chi_0^2 + \sum_{j=1}^{m}\left\{\frac{\partial\chi_0^2}{\partial a_j}(a_j - a_j')\right\} + \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m}\left\{\frac{\partial^2\chi_0^2}{\partial a_k \partial a_j}(a_k - a_k')(a_j - a_j')\right\} \tag{8.9}$$

Because the condition for minimizing χ^2 is that the first partial derivative with respect to each parameter vanish (i.e., $\partial\chi^2/\partial a_j = 0$), we can expect that near a local minimum in any parameter a_j, χ^2 will be a quadratic function of that parameter.

We can obtain another useful relation from Equation (8.8) by taking the second derivative of χ^2 with respect to the parameter a_j to obtain

$$\frac{\partial^2\chi^2}{\partial a_j^2} = \frac{2}{\sigma_j^2} \tag{8.10}$$

We obtain the following expression for the uncertainty in the parameter in terms of the curvature of the χ^2 function in the region of the minimum:

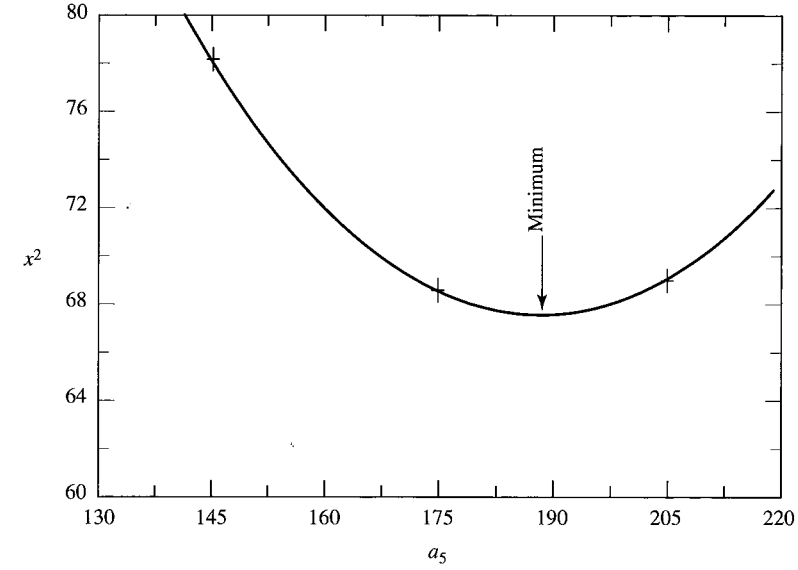


FIGURE 8.3
Plot of χ^2 versus a single parameter a in the region of a local minimum. The location of the minimum is calculated by fitting a parabola through the three indicated data points.

We note that for uncorrelated parameters, Equation (8.11) is equivalent to Equation (7.22) with Equation (7.25) for obtaining the uncertainties from the curvature matrix.

We can also use the quadratic relation to find the approximate location of a χ^2 minimum by considering the equation of a parabola that passes through three points that straddle the minimum, and solving for the value of the parameter at the minimum, as illustrated in Figure 8.3. If we have calculated three values of χ^2, $\chi_1^2 = \chi^2(a_{j1})$, $\chi_2^2 = \chi^2(a_{j2})$, and $\chi_3^2 = \chi^2(a_{j3})$, where $a_{j2} = a_{j1} + \Delta a_j$ and $a_{j3} = a_{j2} + \Delta a_j$, then the value a_j' of the parameter at the minimum of the parabola is given by

$$a_j' = a_{j3} - \Delta a_j\left[\frac{\chi_3^2 - \chi_2^2}{\chi_1^2 - 2\chi_2^2 + \chi_3^2} + \frac{1}{2}\right] \tag{8.12}$$

In addition, we can estimate the errors in the fitting parameters a_j by varying each parameter about its minimum to increase χ^2 by 1 from the minimum value. The variation σ_j in the parameter a_j, which will increase χ^2 by 1 from its value at the minimum of the parabola, is given by

$$\sigma_j = \Delta a_j\sqrt{2(\chi_1^2 - 2\chi_2^2 + \chi_3^2)^{-1}} \tag{8.13}$$

Alternatively, we can attempt to calculate the second derivative of χ^2 at the minimum and find the standard deviation from Equation (8.11).

If the parameters are correlated, the method summarized in Equation (8.13) for determining uncertainties in the parameters is valid only under the condition that,

We provide a more detailed discussion in Section 11.5. When the covariant terms in the error matrix are important, it is best to obtain the full error matrix by the method described in Section 7.2.

8.2 SEARCHING PARAMETER SPACE

The method of least squares consists of determining the values of the parameters a_j of the function $y(x)$ that yield a minimum for the function χ^2 given in Equation (8.4). For nonlinear fitting problems, there are several ways of finding this minimum value. In Sections 8.3 and 8.4 we discuss approximation methods for finding solutions to the m coupled nonlinear equations in m unknowns that result from the minimization procedure of Equation (8.5).

Starting Values and Local Minima

Fitting nonlinear functions to data samples sometimes seems to be more of an art than a science. In part, this is in the nature of the approximation process, where the speed of convergence toward a solution may depend upon the choice of the method for finding solutions, the choice of starting values for the parameters, and possibly the choice of the step size. To use any of these methods, we must first determine starting values, estimates to be used by the fitting routine for initial calculations of the function and of chi square. For the pure search methods we must also define step sizes, the initial variations of the parameters. Neither starting values nor step sizes, of course, are needed in linear fitting.

Another problem in nonlinear fitting is the existence of multiple solutions or local minima. For an arbitrary function there may be more than one minimum of the χ^2 function within a reasonable range of values for the parameters, and thus, more than one set of solutions of the m coupled equations. An unfortunate choice of starting point may "drive" the solution toward a local minimum rather than to the absolute minimum that we seek. Before attempting a nonlinear least-squares fit, therefore, it is useful to search the parameter space to locate the main minima and identify the desired range of parameters over which to refine the search.

The first step is to find starting values for the parameters. A convenient approach, for which a computer graphics program is very useful, is to make plots of the data with curves calculated from trial values of the parameters. By visual inspection, one can often determine acceptable starting values with little or no further calculations. A basic requirement is that the area under the plotted curve be approximately the same as that under the data.

Another approach is to map the parameter space and search for values of the parameters that approximately minimize χ^2. In the simplest brute-force mapping procedure, the permissible range of each parameter a_j is divided into p equal increments Δa_j so that the m-parameter space is divided into $\Pi_{j=1}^{m} P_j$ hypercubes. The value of χ^2 is then evaluated at the vertices of each hypercube. This procedure yields a coarse map of the behavior of χ^2 as a function of all the parameters a_j. At the vertex for which χ^2 has its lowest value, the size of the grid can be reduced to obtain more precise values of the parameters. For a simple two- or three-parameter

fit, the parameters obtained by this procedure may be sufficiently precise that no further searching is required. For more than three parameters, the mapping is rather tedious and displaying the grid map is difficult.

A variation on the regular lattice method is a Monte Carlo search of the m-dimensional space. Trial values of the parameters are generated randomly from uniform distributions of the parameters, selected within predefined ranges, and a value of χ^2 determined for each trial. After several trials, the set of trial values that gives the lowest value of χ^2 can be used as starting values. The general Monte Carlo method was discussed in Chapter 5.

A more sophisticated method of locating the various minima of the χ^2 hypersurface involves traversing the surface from minimum to minimum by the path of lowest value in χ^2, as a river follows a ravine in travelling from lake to ocean. Starting at a point in the m-dimensional space, the search traverses the length of the local minimum, then continues in the same general direction but in a direction that minimizes the new values of χ^2. When a new local minimum is discovered, the search repeats the process until all local minimum have been located in the specified region of the space.

For relatively straightforward fitting problems, it should be sufficient to plot the data, make a reasonable estimate of the parameters to be used as starting values in the search procedure, and perform the fit by one or more of the methods described in the following sections. As a precaution, one should vary the starting values of the parameters to test whether or not the various fits converge to the same values of the parameters, within the expected uncertainties. If the dimensionality of the space is low enough, a grid of starting points may be used. For higher dimensionality, a Monte Carlo method may be used to select random starting points.

Bounds on the Parameters

From a particular set of starting values for the parameters, the search may converge toward solutions that are physically unreasonable. In Example 8.1 negative values for the parameters are not acceptable, and the current trial value of one of the parameters a_2, or a_3, may limit the possibility of determining values of the others. For example, if a_2 becomes very small or 0, a_4 cannot be determined at all. If it is not possible to find starting values for the parameters that prevent the search from wandering into these illegal regions, it may be necessary to place limits on them in the search procedure to keep them within physically allowable ranges. Simple *if then* statements in the routines may be sufficient. Care should be taken that the final value of any parameter is not at one of these artificially imposed limits.

Selection and Adjustment of Step Sizes

There are no hard and fast rules for selecting step sizes for the search methods. Clearly the steps will be different for different parameters and should be related to the slope of the χ^2 function. Very small step sizes result in slow convergence, whereas step sizes that are too large will overshoot the local minima and require constant readjustment to bracket the valleys. In the sample routines in Section 8.7,

we choose initial step sizes to be proportional to the starting values of the parameters and readjust them if necessary after each local minimum is found. In the simple grid-search calculation, we adjust the step sizes to be those values that increase χ^2 by approximately 2 from its value at the local minimum.

Condition for Convergence

A change in χ^2 per degree of freedom (χ^2/dof) of less than about 1% from one trial set of parameters to the next is probably not significant. However, because of the problems of local minima and very flat valleys in the parameter space, it may not be sufficient to set an arbitrary condition for convergence, start a search, and let it run to completion. If the starting parameters are not chosen very carefully, the search may stop in a flat valley with an inappropriately large value of χ^2. If this happens, there are several possible ways to proceed. We can choose different starting values and retry the fit, as suggested in the previous sections, or we can set tighter convergence requirements (e.g., $\Delta\chi^2/\text{dof} < 0.1\%$) and rerun the search in the hope that the program will escape from the valley and reach the appropriate minimum. A convenient approach for small problems is to observe the process of the search and to cut it off manually when it appears that a stable minimum has been found. If a suitable minimum cannot be found, then different starting values should be tried. When fitting curves to several similar samples of data, we may find it satisfactory to establish suitable starting parameters, step sizes, and a cutoff criterion for the first set, and employ an automatic method for the remaining sets.

Computer Illustration of Nonlinear Fitting Methods

In the following sections we discuss and illustrate with computer routines four methods of fitting Equation (8.2) to the data of Example 8.1.

Program 8.0. NONLINFT (Appendix E) Common calling routine to test the four different fitting methods. Repeats the calculations until a χ^2-minimum is found. Variables are defined in the program until FITVARS and data input and output are handled in the program unit FITUTIL as in the fitting programs of Chapters 6 and 7. FITFUNC8 calculates the fitting function.

Step sizes for the fit are set initially in the routine FETCHPARAMETERS to be a fraction of the starting values of the parameters. (The step sizes must not be scaled to the parameters throughout the calculation, however, lest they become 0 when a parameter is 0, which would halt the search in that parameter.)

Tables 8.2, 8.3, 8.4, and 8.5 show values of χ^2 and the parameters a_1 through a_5 for several stages of the calculation at the beginning, middle, and end of each of the four types of search. The tables include the time to find the solution relative to the time for the fastest procedure.

Program 8.1. GRIDSEAR (Appendix E) Routine GRIDLS illustrates the grid-

Program 8.3. EXPNDFIT (Appendix E) Routine CHIFIT illustrates fitting by expansion of the fitting function.

Program 8.4. MARQFIT (Appendix E) Routine MARQUARDT illustrates fitting by the gradient-expansion algorithm.

Program 8.5. FITFUN8 (Appendix E) Fitting function and χ^2-calculation for all fits called from Program 8.0.

Program 8.6. MAKEAB8 (Appendix E) Matrix set-up for non-linear fits.

Program 8.7. NUMDERIV (Website) Numerical derivatives.

Program B.1. MATRIX (Appendix E) Matrix products and inversion.

8.3 GRID-SEARCH METHOD

If the variation of χ^2 with each parameter a_j is not very sensitive to the values of the other parameters, then the optimum parameter values can be obtained most simply by minimizing χ^2 with respect to each of the parameters separately. This is the *grid-search* method. The procedure is simply to select starting values of the parameters, find the value of one of the parameters that minimizes χ^2 with respect to that parameter, set the parameter to that value, and repeat the procedure for each parameter in turn. The entire process is then repeated until a stable χ^2 minimum is obtained.

Grid search. The procedure for a grid search may be summarized as follows:

1. Select starting values a_j and step or increment sizes Δa_j for each parameter and calculate χ^2 with the starting parameters.
2. Increment one parameter a_j by $\pm\Delta a_j$ and calculate χ^2, where the sign is chosen so that χ^2 decreases.
3. Repeat step 2 until χ^2 stops decreasing and begins to increase. The increase in χ^2 indicates that the search has crossed a ravine and started up the other side.
4. Use the last three values of a_j (which bracket the minimum) and the associated values of χ^2 to determine the minimum of the parabola, which passes through the three points as illustrated in Figure 8.3. [See Equation (8.12).]
5. Repeat to minimize χ^2 with respect to each parameter in turn.
6. Continue to repeat the procedure until the last iteration yields a predefined negligibly small decrease in χ^2.

The main advantage of the grid-search method is its simplicity. With successive iterations of the search, the absolute minimum of the χ^2 function in parameter space can be located to any desired precision.

The main disadvantage is that, if the variations of χ^2 with the parameters are strongly correlated, then the approach to the minimum may be very slow. Consid-

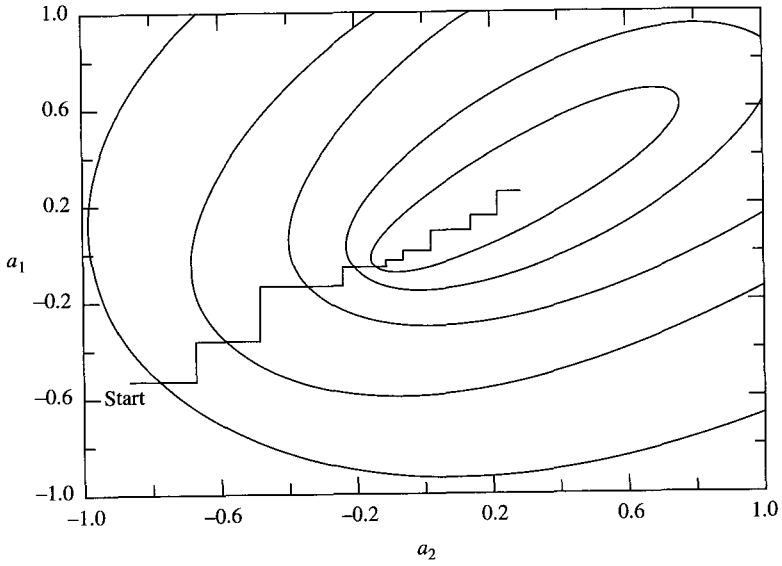


FIGURE 8.4
Contour plot of χ^2 as a function of two highly correlated variables. The zigzag line represents the search path approach to a local minimum by the grid-search method.

parameters are not correlated, so that the variation of χ^2 with each parameter is independent of the variation with the other, then the axes of the ellipse will be parallel to the coordinate axes. Thus, if a grid search is initiated near one end of a tilted ellipse, the search may follow a zigzag path as indicated by the solid line in Figure 8.4 and the search will be very inefficient. Nevertheless, the simplicity of the calculations involved in a grid search often compensates for this inefficiency.

Program 8.1. GRIDSEAR (Appendix E) Routine GRIDLS illustrates the grid-search method.

The main search routine, GRIDLS, is entered with the value of χ^2 (CHISQR) as argument. In a loop over each of the m parameters in turn, the value of the parameter is varied until χ^2 has passed through a local minimum in the parameter. The three most recent values of χ^2 that bracket the minimum are stored in the variables CHISQ1, CHISQ2, and CHISQ3. The best estimate of the parameter at this stage of the calculation is determined from the minimum of the parabola that passes through the three points. The step size (DELTAA(J)) is then adjusted to be that value that increases χ^2 by 2 from its value at the local minimum.

One pass through GRIDLS corresponds to a single zigzag along the path of Figure 8.4. The search is repeated until χ^2 does not change by more than the preset level, CHICUT.

A call to the function SIGPARAB in the program unit FITUTIL at the end

TABLE 8.2
Two exponentials plus constant background: grid-search method

Trial	χ^2	a_1	a_2	a_3	a_4	a_5
0	406.6	10.0	900.0	80.0	27.0	225.0
1	143.0	14.5	1332.3	106.8	27.7	207.2
2	96.9	12.6	1233.9	127.9	28.2	198.4
3	79.4	11.6	1155.1	140.2	28.8	192.2
4	72.9	11.2	1100.3	147.0	29.3	189.2
⋮						
16	66.7	11.3	963.5	148.8	32.3	185.3
17	66.7	11.3	962.5	148.2	32.4	185.8
⋮						
39	66.3	10.9	959.3	139.1	33.3	195.4
40	66.2	10.8	959.2	138.9	33.3	195.7
Uncertainties		0.6	28.3	4.5	0.8	5.0

χ^2/dof = 1.23; probability = 12.1%; relative time = 9.1

Note: Stages in the fit to counts from the decay of excited states of silver. The values of χ^2 and the parameters are listed at the beginning, middle, and end of the search. The uncertainties in the parameters correspond to a change of 1 in χ^2 from its value at the end of the search.

Table 8.2 shows values of χ^2 and the parameters a_1 through a_5 for several stages of the calculation at the beginning, middle, and end of the search. The search is relatively slow, but eventually a satisfactory solution is found. Note that the calculated uncertainties correspond to the diagonal terms in the error matrix for uncorrelated parameters. If correlations are considered to be important, the matrix inversion methods discussed in the following sections could be used to find better approximations to the uncertainties.

8.4 GRADIENT-SEARCH METHOD

The search could be improved if the zigzagging direction of travel in Figure 8.4 were replaced by a more direct vector toward the appropriate minimum. In the *gradient-search* method of least squares, all the parameters a_j are incremented simultaneously, with relative magnitudes adjusted so that the resultant direction of travel in parameter space is along the gradient (or direction of maximum variation) of χ^2.

The gradient $\nabla\chi^2$ is a vector that points in the direction in which χ^2 increases most rapidly and has components in parameter space equal to the rate of change of χ^2 along each axis:

$$\nabla\chi^2 = \sum_{j=1}^{n} \left[\frac{\partial \chi^2}{\partial a_j} \hat{a}_j \right] \tag{8.14}$$

where $\hat{a}_j$ indicates a unit vector in the direction of the a_j coordinate axis. In order to

$$(\nabla\chi^2)_j = \frac{\partial\chi^2}{\partial a_j} \simeq \frac{\chi^2(a_j + f\Delta a_j) - \chi^2(a_j)}{f\Delta a_j} \tag{8.15}$$

where f is a fraction of the step size Δa_j by which a_j is changed in order to determine the derivative.

The gradient has both magnitude and dimensions and, if the dimensions of the various parameters a_j are not all the same (which is usually the case), the components of the gradient do not even have the same dimensions. Let us define dimensionless parameters b_j by rescaling each of the parameters a_j to a size that characterizes the variation of χ^2 with a_j rather roughly. We shall use the step sizes Δa_j as the scaling constants, so that

$$b_j = \frac{a_j}{\Delta a_j} \tag{8.16}$$

The derivative with respect to b_j then becomes

$$\frac{\partial\chi^2}{\partial b_j} = \frac{\partial\chi^2}{\partial a_j}\Delta a_j \tag{8.17}$$

which may be calculated numerically as

$$\frac{\partial\chi^2}{\partial b_j} \simeq \frac{\chi^2(a_j + f\Delta a_j) - \chi^2(a_j)}{f\Delta a_j}\Delta a_j = \frac{\chi^2(a_j + f\Delta a_j) - \chi^2(a_j)}{f} \tag{8.18}$$

We can then define a dimensionless gradient γ, with unit magnitude and components

$$\gamma_j = \frac{\partial\chi^2/\partial b_j}{\sqrt{\sum_{j=1}^{m}(\partial\chi^2/\partial b_j)^2}} \tag{8.19}$$

In the numerical calculation of Equation (8.18), the quantities Δa_j and f occur only in the argument of χ^2 and not as scale factors.

The direction that the gradient-search method follows is the *direction of steepest descent,* which is opposite of the gradient γ. The search begins by incrementing all parameters simultaneously by an amount Δa_j, with relative value given by the corresponding component γ_j of the dimensionless gradient and absolute magnitude given by the size constant Δa_j:

$$\delta a_j = -\gamma_j \Delta a_j \tag{8.20}$$

The minus sign ensures that the value of χ^2 decreases. The size constant Δa_j of Equation (8.20) is the same as that of Equation (8.16).

There are several possible methods of continuing the gradient search after a first step. The most straightforward is to recompute the gradient after each change in the parameters. One disadvantage of this method is that it is difficult to approach the bottom of the minimum asymptotically because the gradient tends to 0 at the minimum. Another disadvantage is that recomputation of the gradient at each step

rise again. At this point, the gradient is recomputed and the search continues in the new direction. Whenever the search straddles a minimum, a parabolic interpretation of χ^2 is used to improve the determination of the minimum.

A more sophisticated approach would be to use second partial derivatives of χ^2 to determine changes in the gradient along the search path:

$$\left.\frac{\partial\chi^2}{\partial a_j}\right|_{a_j+\delta a_j} \simeq \left.\frac{\partial\chi^2}{\partial a_j}\right|_{a_j} + \sum_{k=1}^{m}\left(\frac{\partial^2\chi^2}{\partial a_j \partial a_k}\delta a_k\right) \tag{8.21}$$

If the search is already fairly near the minimum, this method does decrease the number of steps needed, but at the expense of more elaborate computation. If the search is not near enough to the minimum, this method can actually increase the number of steps required when first-order perturbations on the gradient are not valid.

The efficiency of the gradient search decreases markedly as the search approaches a minimum because the evaluation of the derivative according to the method of Equation (8.18) involves taking differences between nearly equal numbers. In fact, at the minimum of χ^2, these differences should vanish. For this reason, one of the methods discussed in the following sections may be used to locate the actual minimum once the gradient search has approached it fairly closely.

Program 8.2. GRADSEAR (Appendix E) Routine GRADLS illustrates the gradient-search method.

On each entry to the main search routine, GRADLS, the components of the gradient GRADLS(J) are calculated numerically from Equation (8.18) in the procedure CALCGRAD. The argument FRACT of this routine, corresponding to the variable f of Equation (8.18), determines the fraction of the step size (DELTAA) used in the numerical calculation of the partial derivative. Each parameter A(J) is then changed by the amount STEPDOWN*DELTAA(J)*GRAD(J), where STEPDOWN is a scaling factor that is set initially in the main program and readjusted after each stage to the size needed to locate the minimum.

The initial values of DELTAA(J) determines to some extent the execution speed of each pass through the routine GRADLS, and the value of CHICUT determines when the search will stop. Because of the small gradient near the χ^2 minimum, it may take many steps to reach a reasonable value of χ^2, and the cutoff, CHICUT, may have to be set to a very low value. For such cases, user intervention can be provided as an alternate method of stopping the search.

At the conclusion of the search, the uncertainties in the parameters are estimated in the function SIGPARAB as in the routine GRADLS.

Table 8.3 shows values of χ^2 and the parameters a_1 through a_5 for several stages of the calculation at the beginning, middle, and end of the search. For Example 8.1, the gradient search is considerably faster than the grid-search approach because all the parameters are varied together at each step. However, the gradient-search method has one disadvantage that is not illustrated. If the starting values of

TABLE 8.3
Two exponentials plus constant background: gradient-search method

Trial	χ^2	a_1	a_2	a_3	a_4	a_5
0	406.6	10.0	900.0	80.0	27.0	225.0
1	82.3	10.6	1061.0	94.0	34.4	254.2
2	72.6	9.8	984.0	98.8	36.8	237.4
3	69.8	9.9	966.9	100.9	36.8	244.6
4	69.3	9.8	953.7	101.6	36.7	242.1
⋮						
19	66.6	8.9	952.2	114.7	35.5	233.6
20	66.5	8.9	954.8	114.9	35.6	233.9
Uncertainties		0.6	26.5	3.8	0.8	7.0

χ^2/dof = 1.23; probability = 11.8%; relative time = 4.0

Note: Stages in the fit to counts from the decay of excited states of silver. The values of χ^2 and the parameters are listed at the beginning, middle, and end of the search. The uncertainties in the parameters corresponding to a change of 1 in χ^2 from its value at the end of the search.

8.5 EXPANSION METHODS

Instead of searching the χ^2 hypersurface to map the variation of χ^2 with parameters, we should be able to find an approximate analytical function that describes the χ^2 hypersurface and use this function to locate the minimum, with methods developed for linear least-squares fitting. The approximations will introduce errors into the calculated values of the parameters, but successive iterations of the analytical method should approach the χ^2 minimum with increasing accuracy. The main advantage of such an approach is that the number of points on the χ^2 hypersurface at which computations must be made will be fewer than for a grid or gradient search. This advantage is somewhat offset by the fact that the computations at each point are considerably more complicated. However, the analytical solution essentially chooses its own step size and, thus, the user is spared the problem of trying to optimize the step size for speed and precision.

Parabolic Expansion of χ^2

In Equation (8.9) we expanded χ^2 to second order in the parameters about a local minimum χ_0^2 where $a_j = a_j'$:

$$\chi^2 \simeq \chi_0^2 + \sum_{j=1}^{m}\left\{\frac{\partial\chi_0^2}{\partial a_j}\delta a_j\right\} + \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m}\left\{\frac{\partial^2\chi_0^2}{\partial a_j\partial a_k}\delta a_j\delta a_k\right\} \tag{8.22}$$

which is equivalent to approximating the χ^2 hypersurface by a parabolic surface. Here we define $\delta a_j \equiv a_j - a_j'$, and χ_0^2 is given by

$$\chi^2 = \sum\left\{\frac{1}{[\quad]}[\quad - y'(x)]^2\right\} \tag{8.23}$$

Applying the method of least squares, we minimize χ^2 as expressed in Equation (8.22) with respect to the *increments* (δa_j) in the parameters, and solve for the optimum values of these increments to obtain

$$\frac{\partial\chi^2}{\partial(\delta a_k)} = \frac{\partial\chi_0^2}{\partial a_k} + \sum_{j=1}^{m}\left\{\frac{\partial^2\chi_0^2}{\partial a_k\partial a_j}\delta a_j\right\} = 0 \quad k = 1, m \tag{8.24}$$

The result is a set of m linear equations in δa_j that we can write as

$$\beta_k - \sum_{j=1}^{m}(\delta a_j\alpha_{jk}) = 0 \qquad k = 1, m \tag{8.25}$$

with

$$\beta_k \equiv -\frac{1}{2}\frac{\partial\chi_0^2}{\partial a_k} \qquad \text{and} \qquad \alpha_{jk} \equiv \frac{1}{2}\frac{\partial^2\chi_0^2}{\partial a_j\partial a_k} \tag{8.26}$$

The factors $\pm\frac{1}{2}$ are included for agreement with the conventional definitions of these quantities.

As in Chapter 7, we can treat Equation (8.25) as a matrix equation:

$$\boldsymbol{\beta} = \boldsymbol{\delta a}\,\boldsymbol{\alpha} \tag{8.27}$$

where $\boldsymbol{\beta}$ and $\boldsymbol{\delta a}$ are row matrices and $\boldsymbol{\alpha}$ is a symmetric matrix of order m. We shall find that $\boldsymbol{\alpha}$ is the *curvature matrix* discussed in Section 7.2, so named because it measures the curvature of the χ^2 hypersurface.

Method of Computation

The solution of Equation (8.27) can be obtained by matrix inversion as in Section 7.2:

$$\boldsymbol{\delta a} = \boldsymbol{\beta\epsilon} \qquad \delta a_k = \sum_{j=1}^{m}(\epsilon_{kj}\beta_j) \tag{8.28}$$

where the error matrix $\boldsymbol{\epsilon} = \boldsymbol{\alpha}^{-1}$ is the inverse of the curvature matrix.

If the parameters are independent of one another, that is, if the variation of χ^2 with respect to each parameter is independent of the values of the other parameters, then the cross-partial derivatives a_{jk} ($j \neq k$) will be 0 in the limit of a very large data sample and the matrix $\boldsymbol{\alpha}$ will be diagonal. The inverse matrix $\boldsymbol{\epsilon}$ will also be diagonal and Equation (8.27) will degenerate into m separate equations:

$$\delta\alpha_j \simeq \frac{\beta_j}{\alpha_{jj}} = \frac{\partial x_0^2}{\partial a_j} \div \frac{\partial^2 x_0^2}{\partial a_j^2} \tag{8.29}$$

Computation of the matrix elements by Equation (8.26) requires knowledge of the first and second derivatives of χ^2 evaluated at the current values of the parameters. Analytic forms of the derivatives are generally quickest to compute, but may be difficult or cumbersome to derive. If it is not convenient or possible to provide

for efficient calculations. The intervals Δa_j should be chosen to be large enough to avoid roundoff errors but small enough to furnish reasonably accurate values of the derivatives near the minimum:

$$\begin{aligned}
\frac{\partial \chi_0^2}{\partial a_j} &\simeq \frac{\chi_0^2(a_j + \Delta a_j, a_k) - \chi_0^2(a_j, a_k)}{\Delta a_j} \\
\frac{\partial^2 \chi_0^2}{\partial^2 a_j} &\simeq 4\left[\frac{\chi_0^2(a_j, a_k) - 2\chi_0^2(a_j + \delta a_j/2, a_k) + \chi_0^2(a_j + \Delta a_j, a_k)}{(\Delta a_j)^2}\right] \\
\frac{\partial^2 \chi_0^2}{\partial a_j \partial a_k} &\simeq [\chi_0^2(a_j, a_k) \\
&\quad - \chi_0^2(a_j + \Delta a_j, a_k) - \chi_0^2(a_j, a_k + \Delta a_k) \\
&\quad + \chi_0^2(a_j + \Delta a_j, a_k + \Delta a_k)]/[\Delta a_j \Delta a_k]
\end{aligned} \tag{8.30}$$

In actual practice, calculations are faster and, in general, more accurate if the elements of the matrix $\boldsymbol{\alpha}$ are determined from the first-order expansion (to be discussed in the following text), which involves only first derivatives of $y(x)$ with respect to the parameters, rather than the second derivatives of χ^2 as expressed in Equation (8.30).

Fitting Procedure

Within the limits of the approximation of the χ^2 hypersurface by a parabolic extrapolation, we can solve Equation (8.27) directly to yield parameter increments δa_j such that χ^2 should be minimized for $a_j' + \delta a_j$. If the starting point is close enough to the minimum so that higher-order terms in the expansion can be neglected, this becomes an accurate and precise method. But if the starting point is not near enough, the parabolic approximation of the χ^2 hypersurface is not valid and the results will be in error. In fact, if the starting point is so far from the minimum that the curvature of χ^2 is negative, the solution will tend toward a maximum rather than a minimum. During computation, therefore, the diagonal elements α_{jj} of the matrix $\boldsymbol{\alpha}$ must be set positive whether they are or not. The resulting magnitude for δa_j will be incorrect, but the sign will be correct.

Expansion of the Fitting Function

An alternative to expanding the χ^2 function to develop an analytic description for the hypersurface is to expand the fitting function $y(x)$ in the parameters a_j and to use the method of linear least squares to determine the optimum value for the parameter increments δa_j. If we carry out the derivation rigorously and drop higher-order terms, we should achieve the same result as for the expansion of χ^2 to first and second order.

First-Order Expansion

Let us expand the fitting function $y(x)$ in a Taylor series about the point a_j', to first order in the *parameter increments* $\delta a_j = a_j - a_j'$:

$$y(x) \simeq y'(x) + \sum_{j=1}^{m}\left[\frac{\partial y'(x)}{\partial a_j}\delta a_j\right] \tag{8.31}$$

where $y'(x)$ is the value of the fitting function when the parameters have starting point values a_j' and the derivatives are evaluated at the starting point. The result is a linear function in the parameter increments δa_j to which we can apply the method of linear least squares developed in Chapter 7.

In this approximation, χ^2 can be expressed explicitly as a function of the parameter increments δa_j:

$$\chi^2 = \sum\left(\frac{1}{\sigma_i^2}\left\{y_i - y'(x_i) - \sum_{j=1}^{m}\left[\frac{\partial y'(x_i)}{\partial a_j}\delta a_j\right]\right\}^2\right) \tag{8.32}$$

Following the method of least squares, we minimize χ^2 with respect to each of the parameter increments δa_j by setting the derivatives equal to 0:

$$\frac{\partial \chi^2}{\partial \delta a_k} = -2\sum\left(\frac{1}{\sigma_i^2}\left\{y_i - y'(x_i) - \sum_{j=1}^{m}\left[\frac{\partial y'(x_i)}{\partial a_j}\delta a_j\right]\right\}\frac{\partial y'(x_i)}{\partial a_k}\right) = 0 \tag{8.33}$$

As before, this yields the set of m simultaneous Equations (8.25), which can be expressed as the matrix Equation (8.27):

$$\boldsymbol{\beta} = \boldsymbol{\delta a}\,\boldsymbol{\alpha} \tag{8.34}$$

where β_k is defined as in Equation (8.26) and α_{jk} is given by

$$\alpha_{jk} \simeq \sum\left[\frac{1}{\sigma_i^2}\frac{\partial y'(x_i)}{\partial a_j}\frac{\partial y'(x_i)}{\partial a_k}\right] \tag{8.35}$$

Second-Order Expansion

Suppose we make a Taylor expansion of the fitting function $y(x)$ to second order in the parameter increments δa_j:

$$y(x) \simeq y'(x) + \sum_{j=1}^{m}\left[\frac{\partial y'(x)}{\partial a_j}\delta a_j\right] + \frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left[\frac{\partial^2 y'(x)}{\partial a_j \partial a_k} da_j da_k\right] \tag{8.36}$$

If we include the last term of Equation (8.36) in the expression for χ^2 of Equation (8.32) and again minimize χ^2 by setting to 0 the derivatives with respect to the increments δa_j, we again obtain Equation (8.25), this time with

$$\begin{aligned}
\beta_k &\equiv \sum\left\{\frac{1}{\sigma_i^2}[y_i - y'(x_i)]\frac{\partial y'(x_i)}{\partial a_k}\right\} = -\frac{1}{2}\frac{\partial \chi_0^2}{\partial a_k} \\
\alpha_{jk} &\equiv \sum\frac{1}{\sigma_i^2}\left\{\frac{\partial y'(x_i)}{\partial a_j}\frac{\partial y'(x_i)}{\partial a_k} - [y_i - y'(x_i)]\frac{\partial^2 y'(x_i)}{\partial a_j \partial a_k}\right\} \\
&= \frac{1}{2}\frac{\partial^2 \chi_0^2}{\partial a_j \partial a_k}
\end{aligned} \tag{8.37}$$

The resulting definitions for β_k and a_{jk} are identical to those of Equation (8.26) obtained by expanding the χ^2 function, and the χ^2-expansion method is therefore equivalent to a second-order expansion of the fitting function.

Let us compare Equations (8.37) with the analogous Equations (7.14) and (7.15) for linear least-squares fitting. The definitions of α_{jk} in Equations (8.37) and (7.15) are equivalent in the linear approximation [See Equation (7.22)] and thus $\boldsymbol{\alpha}$ corresponds to the curvature matrix. The definition of β_k in Equation (8.37) is equivalent, in the linear approximation, to the definition of β_k in Equations (7.14) except for the substitution of $y_i - y'(x_i)$ for y_i. We can justify this substitution by noting that the solutions of Equation (8.34) are the parameter increments δa_j, whereas those of Equation (7.14) are the parameters themselves. In essence, we are applying linear least-squares methods to fit the parameter increments to difference data Δy_i between the actual data and the starting values of the fitting $y'(x_i)$:

$$\Delta y_i = y_i - y'(x_i) \tag{8.38}$$

Thus, the expression given in Equation (8.35) for α_{jk} is a first-order approximation to the curvature matrix that is given to second order in Equation (8.37). For linear functions, the second-order term vanishes. It is convenient to use the first-order approximation for fitting nonlinear functions and thus avoid the necessity of calculating the second derivatives in Equation (8.37). We note that this procedure can be somewhat justified on the grounds that, in the vicinity of the χ^2 minimum, we should expect the factor of $y_i - y'(x_i)$ in the expression for α of Equation (8.37) to be close to 0 so that the first term in the expression will dominate.[1]

Program 8.3. EXPNDFIT (Appendix E) Routine CHIFIT illustrates non-linear fitting by expansion of the fitting function. The program is called repeatedly from the main program NONLINFIT, until χ^2 passes through a minimum. EXPNDFIT calls the following routines to set up and manipulate the matrices.

Program 8.6. MAKEAB8 (Appendix E) Sets up the $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ matrices.
The routine uses the first-order approximation of Equation (8.35) to calculate the components α_{jk} of the curvature matrix. This is equivalent to neglecting terms in the second derivatives of the fitting function $y(x)$ in the expression for α_{jk} in Equation (8.37). The routines in this program unit use numerical derivatives and therefore differ from those with the same names in Chapter 7, which use analytic derivatives.

Program 8.7. NUMDERIV (website) Numerical derivatives.
Derivatives of χ^2 (XISQ) are calculated numerically by the functions DXISQ_DA, D2XISQ_DA2, and D2XISQ_DAJK in this program unit. To avoid repetitive calculations, the values of the derivatives at each value of x and for the variation of each of the m parameters are calculated once for each trial and stored in arrays. If available, analytic expressions for the derivatives could be substituted directly for the functions to increase the speed and accuracy of the calculation.

Program B.1. MATRIX (Appendix E) Matrix multiplication and inversion.

[1]See Press et al. (1986), page 523.

TABLE 8.4
Two exponentials plus constant background: χ^2 expansion method

Trial	χ^2	a_1	a_2	a_3	a_4	a_5
0	406.6	10.0	900.0	80.0	27.0	225.0
1	86.2	11.1	933.8	140.4	33.8	170.5
2	66.6	10.8	861.2	128.9	33.9	201.7
3	66.1	10.4	958.2	131.2	34.0	205.4
Uncertainties		1.8	49.9	21.7	2.5	30.5

χ^2/dof = 1.22; probability = 12.4%; relative time = 1.0

Note: All stages in the fit to counts from the decay of excited states of silver. The uncertainties in the parameters correspond to the square roots of the diagonal terms in the error matrix.

At the conclusion of the search, the inverse $\boldsymbol{\epsilon}$ of the final value of the curvature matrix $\boldsymbol{\alpha}$ is treated as the error matrix, and the errors in the parameters are obtained from the square roots of the diagonal terms by calls to the function SIGMATRX in the unit FitFunc8. Table 8.4 shows values of χ^2 and the parameters a_1 through a_5 for all stages of the calculation.

8.6 THE MARQUARDT METHOD

Convergence

One disadvantage inherent in the analytical methods of expanding either the fitting function $y(x)$ or χ^2 is that although they converge quite rapidly to the point of minimum χ^2 from points nearby, they cannot be relied on to approach the minimum with any accuracy from a point outside the region where the χ^2 hypersurface is approximately parabolic. In particular, if the curvature of the χ^2 hypersurface is used, as in Equation (8.37) or (8.26), the analytical solution is clearly unreliable whenever the curvature becomes negative. Symptomatic of this problem is the need to set positive the diagonal elements α_{jj} of the matrix α so that all curvatures are treated as if they were positive.

In contrast, the gradient search of Section 8.4 is ideally suited for approaching the minimum from far away, but does not converge rapidly near the minimum. Therefore, we need an algorithm that behaves like a gradient search for the first portion of a search and behaves more like an analytical solution as the search converges. In fact, it can be shown (see Marquardt 1963) that the path directions for gradient and analytical searches are nearly perpendicular to each other, and that the optimum direction is somewhere between these two vectors.

One advantage of combining these two methods into one algorithm is that the simpler first-order expansion of the analytical method will certainly suffice because the expansion need only be valid in the immediate neighborhood of the minimum. Thus, to calculate the curvature matrix α, we can use the approximation of Equation (8.35) and ignore the second derivatives of Equation (8.37).

Gradient-Expansion Algorithm

A convenient algorithm (see Marquardt 1963), which combines the best features of the gradient search with the method of linearizing the fitting function, can be obtained by increasing the diagonal terms of the curvature matrix $\boldsymbol{\alpha}$ by a factor $1 + \lambda$ that controls the interpolation of the algorithm between the two extremes. Equation (8.34) becomes

$$\boldsymbol{\beta} = \boldsymbol{\delta a}\,\boldsymbol{\alpha}' \quad \text{with} \quad \alpha'_{jk} = \begin{cases} \alpha_{jk}(1+\lambda) & \text{for } j = k \\ \alpha_{jk} & \text{for } j \neq k \end{cases} \tag{8.39}$$

If λ is very small, Equations (8.39) are similar to the solution of Equation (8.34) developed from the Taylor expansion. If λ is very large, the diagonal terms of the curvature matrix dominate and the matrix equation degenerates into m separate equations

$$\beta_j \simeq \lambda \delta a_j \alpha_{jj} \tag{8.40}$$

which yield the vector increment $\boldsymbol{\delta a}$ in the same direction as the vector $\boldsymbol{\beta}$ of Equation (8.37) (or opposite to the gradient of χ^2).

The solution for the parameter increments δa_j follows from Equations (8.39) after matrix inversion

$$\delta a_j = \sum_{k=1}^{m} (\beta_k \epsilon'_{jk}) \tag{8.41}$$

where the β_k are given by Equation (8.37) and the matrix $\boldsymbol{\epsilon}'$ is the inverse of the matrix $\boldsymbol{\alpha}'$ with elements given by Equations (8.39).

The initial value of the constant factor λ should be chosen small enough to take advantage of the analytical solution, but large enough that χ^2 decreases. Because this algorithm approaches the gradient-search method with small steps for large λ, there should exist a value of λ such that $\chi^2(a + \delta a) < \chi^2(a)$. The recipe given by Marquardt is:

1. Compute $\chi^2(a)$.
2. Start initially with $\lambda = 0.001$.
3. Compute δa and $\chi^2(a + \delta a)$ with this choice of λ.
4. If $\chi^2(a + \delta a) > \chi^2(a)$, increase λ by a factor of 10 and repeat step 3.
5. If $\chi^2(a + \delta a) < \chi^2(a)$, decrease λ by a factor of 10, consider $a' = a + \delta a$ to be the new starting point, and return to step 3, substituting a' for a.

For each iteration it may be necessary to recompute the parameter increments δa_j from Equation (8.41), and the elements α_{jk} and β_j of the matrices, several times to optimize λ. As the solution approaches the minimum, the value of λ will decrease

TABLE 8.5
Two exponentials plus constant background: Marquardt method

Trial	χ^2	a_1	a_2	a_3	a_4	a_5
0	406.6	10.0	900.0	80.0	27.0	225.0
1	82.9	11.0	933.5	139.3	33.9	173.9
2	66.4	10.8	960.1	130.6	33.8	201.2
3	66.1	10.4	958.3	131.4	33.9	205.0
Uncertainties		1.8	49.9	21.7	2.5	30.5

χ^2/dof = 1.22; probability = 12.4%; relative time = 1.0

Note: All stages in the fit to counts from the decay of excited states of silver. The uncertainties in the parameters correspond to the square roots of the diagonal terms in the error matrix.

TABLE 8.6
Elements of the error matrix (Marquardt method)

1/k	1	2	3	4	5
1	3.38	−3.69	27.98	−2.34	−49.24
2	−3.69	2492.26	81.89	−69.21	−3.90
3	27.98	81.89	468.99	−44.22	−615.44
4	−2.34	−69.21	−44.22	6.39	53.80
5	−49.24	−3.90	−615.44	53.80	929.45

Note: Error matrix from a fit to the radioactive silver data. The diagonal terms are the variances σ_k^2 and the off-diagonal terms are the covariances σ_{kl}^2 of the parameters a_k.

Program 8.4. MARQFIT (Appendix E) Routine MARQUARDT illustrates fitting by the gradient-expansion algorithm.
The procedure uses the same program units as those in Program 8.3, and is identical to that program except for the adjustment of the diagonal elements α_{jj} of the matrix $\boldsymbol{\alpha}$ by the variable LAMBDA according to Equation (8.39).

At the conclusion of the search, the inverse $\boldsymbol{\epsilon}$ of the final value of the curvature matrix $\boldsymbol{\alpha}$ is treated as the error matrix, and the errors in the parameters are obtained from the square roots of the diagonal terms by calls to the function SIGMATRX in the unit FitFunc8. Table 8.5 shows values of χ^2 and the parameters a_1 through a_5 for all stages of the calculation. Table 8.6 shows the error matrix from the fit.

8.7 COMMENTS

Although the Marquardt method is the most complex of the four fitting routines, it is also the clear winner for finding fits most directly and efficiently. It has the strong advantage of being reasonably insensitive to the starting values of the parameters,

ranges. The Marquardt method also has the advantage over the grid- and gradient-search methods of providing an estimate of the full error matrix and better calculation of the diagonal errors.

The routines of Programs 8.3 and 8.4 were tested with both numerical and analytical derivatives. Typical search paths with numerical derivatives are shown in Tables 8.4 and 8.5. For the sample problem with the assumed starting conditions, the minimum χ^2 was found in only a few steps by either method with essentially no time difference. Both methods are reasonably insensitive to starting values of parameters in which the fit is linear, but can be sensitive to starting values of the nonlinear parameters. Program 8.4 had remarkable success over a broad range of starting values, whereas Program 8.3 required better definition of the starting values of the parameters and generally required many more iterations.

The uncertainties in the parameters for these fits were calculated from the diagonal terms in the error matrices and are, in general, considerably larger than the uncertainties obtained in the grid- and gradient-search methods. Because the latter errors were obtained by finding the change in each parameter to produce as change of χ^2 of 1 from the minimum values, without reoptimizing the fit, there is a strong suggestion that correlations among the parameters play an important role in fitting Figure 8.1. This point of view is supported by examination of the error matrix from the method 4 fit (Table 8.6), which shows large off-diagonal elements.

With poorly selected starting values, the searches may terminate in local minima with unacceptably high values of χ^2 and, therefore, with unacceptable final values for the parameters. Termination in the sample programs is controlled simply by considering the reduction in χ^2 from one iteration to the next and stopping at a preselected difference. With this method, it is essential to check the results carefully to be sure that the absolute minimum has indeed been found.

SUMMARY

Nonlinear function: One that cannot be expressed as a sum of terms with the coefficients of the terms.
Minimum of χ^2 (parabolic approximation):

$$a_j' = a_{j3} - \Delta a_j \left[\frac{\chi_3^2 - \chi_2^2}{\chi_1^2 - 2\chi_2^2 + \chi_3^2} + \frac{1}{2} \right]$$

Estimate of standard deviation from $\Delta\chi^2 = 1$:

$$\sigma_j = \Delta a_j \sqrt{2(\chi_1^2 - 2\chi_2^2 + \chi_3^2)^{-1}}$$

Grid search: Vary each parameter in turn, minimizing χ^2 with respect to each parameter independently. Many successive iterations are required to locate the minimum of χ^2 unless the parameters are independent; that is, unless the variation of χ^2 with respect to one parameter is independent of the values of the other parameters.
Gradient search: Vary all the parameters simultaneously, adjusting relative magnitudes of the variations so that the direction of propagation in parameter space is along the direction of steepest descent of χ^2.

Direction of steepest descent: Opposite the gradient $\nabla\chi^2$:

$$(\nabla\chi^2)_j = \frac{\partial\chi^2}{\partial a_j} \simeq \frac{\chi^2(a_j + f\Delta a_j) - x(a_j)}{f\Delta a_j}$$

$$\delta a_j = \frac{-((\partial\chi^2/\partial a_j)\Delta a_j^2)}{\sqrt{\sum_{j=1}^{m}((\partial\chi^2/\partial a_j)\Delta a_j)^2}}$$

Parabolic expansion of χ^2:

$$\boldsymbol{\delta\alpha = \beta\epsilon} \qquad \delta a_k = \sum_{j=1}^{m} (\epsilon_{kj}\beta_j)$$

with

$$\beta_k \equiv -\frac{1}{2}\frac{\partial\chi^2}{\partial a_k} \quad \text{and} \quad \alpha_{jk} \equiv \frac{1}{2}\frac{\partial^2\chi^2}{\partial a_j \partial a_k}$$

Linearization of the fitting function:

$$\beta_k \equiv \sum \left\{ \frac{1}{\sigma_i^2}[y_i - y(x_i)]\frac{\partial y(x_i)}{\partial a_k} \right\} = -\frac{1}{2}\frac{\partial\chi_0^2}{\partial a_k}$$

$$\alpha_{jk} \equiv \sum \frac{1}{\sigma_i^2}\left\{ \frac{\partial y(x_i)}{\partial a_j}\frac{\partial y(x_i)}{\partial a_k} - [y_i - y'(x_i)\frac{\partial^2 y(x_i)}{\partial a_j \partial a_k} \right\}$$

$$= \frac{1}{2}\frac{\partial^2\chi^2}{\partial a_j \partial a_k}$$

Gradient-expansion algorithm—the Marquardt method: Make λ just large enough to insure that χ^2 decreases:

$$\alpha_{jk}' = \begin{cases} \alpha_{jk}(1+\lambda) & \text{for } j = k \\ \alpha_{jk} & \text{for } j \neq k \end{cases}$$

$$\alpha_{jk} \simeq \sum \left[\frac{1}{\sigma_i^2}\frac{\partial y(x_i)}{\partial a_j}\frac{\partial y(x_i)}{\partial a_k} \right] \qquad \beta_k = -\frac{1}{2}\frac{\partial\chi^2}{\partial a_k}$$

$$\delta a_j = \sum_{k=1}^{m} (\beta_k \epsilon_k')$$

Uncertainty in parameter a_j: $\alpha_{aj} = e_{jj}$ corresponds to $\Delta\chi^2 = 1$.

EXERCISES

8.1. Use an interpolation method (see Appendix A) to find the equation of the parabola that passes through the three points (x_1, y_1), (x_2, y_2), and (x_3, y_3). Find the value of x at the minimum of the parabola and thus verify Equation (8.12).

8.2. From the results of Exercise 8.1, verify Equation (8.13).

8.3. The following data represent histogram bin counts across a Lorentzian peak:

x_i	1.824	1.828	1.832	1.836	1.840	1.844	1.848	1.852	1.856	1.860
y_i	558	679	696	736	834	812	899	817	767	657

(*a*) Use the grid-search method to fit the equation $y(x) = AP_L(x; \mu, \Gamma)$ to the data and find the maximum-likelihood value of μ, where $P_L(x; \mu, \Gamma)$ is the Lorentzian function of Equation (2.32) and the known parameters are $A = 75$ and $\Gamma = 0.055$. Assume that x is given at the lower edge of each histogram bin and that the errors in y are statistical. Find the uncertainty in μ.

Suggested procedure: (i) Calculate χ^2 at the peak of the distribution and at a value on each side. (ii) Find the minimum of a parabola that passes through the three points. (iii) Repeat the procedure with three points centered on the minimum χ^2 until the value of μ has been determined to ± 0.001.

(*b*) Repeat the procedure for a two-parameter fit, with Γ as the second unknown.

8.4. Consider the histogram of measured time intervals displayed in Figure 1.2. The numbers of events in the bins bounded by $t = 0.59$ to 0.70s.

2, 2, 11, 6, 12, 8, 4, 3, 1, 1, 0

Fit a Gaussian curve [Equation (2.23)] to these data by the least-squares method to find μ, σ, and the amplitude of the curve A. Bins with fewer than seven events should be merged to improve the reliance on Gaussian statistics. Compare the parameters obtained from the fit with those determined by taking the mean and standard deviation of the data.

8.5. The following data correspond to counts recorded in Example 6.2 with the addition of an unknown randomly fluctuating background term a_1. Use the Marquardt method to fit the equation $C = a_1 + a_2/d^2$ to these data to find the parameters a_1 and a_2 and the full error matrix. Assume statistical uncertainties.

i	1	2	3	4	5	6	7	8	9	10
d_i(m)	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.60	0.75	1.00
C_i	944	688	467	366	316	317	264	251	214	184

8.6. Use the method of least squares to fit the five-parameter equation $y(x) = a_1 + a_2x + a_3G(x; a_4, a_5)$ to the following data where $a_4 = \mu$, $a_5 = \sigma$, and $G(x; \mu, \sigma)$ is the Gaussian curve of Equation (2.23).

i	1	2	3	4	5	6	7	8	9	10
x_i	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
y_i	31	25	24	30	34	37	31	30	64	54
i	11	12	13	14	15	16	17	18	19	20
x_i	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
y_i	95	94	78	79	43	54	58	52	46	41

Use the Marquardt method and find an estimate of the error matrix. The value of x is given at the lower edge of each bin. Assume statistical uncertainties.

8.7. To check the inverse-square relationship expressed in Coulomb's law,

$$F = kQ_1Q_2/r^2$$

Students in an undergraduate laboratory measured the force of electrostatic repulsion between two charged conducting spheres as a function of the distance between the centers of the spheres.

They applied the same potential to each sphere so that each carried the same charge. Because of the mutual repulsion of the charges on the conducting spheres, the effective separation of the two charge distributions is not simply the separation of the centers of the spheres. The resulting reduction in the repulsive force is a function of the separation r of the spheres and their radii a, given approximately by the correction factor

$$f = 1 - 4(a/r)^3$$

where $a = 1.9$ cm in this experiment. Thus, the relation between the mutual force on the spheres and their separation, including the correction factor, can be expressed as

$$F_{coulomb} = \left[1 - 4\left(\frac{a}{r}\right)^3\right]\frac{kQ_0Q_1}{r^2}$$

The students used a torsion balance to study the variation of the repulsive force, so that the force was proportional to the measured torsion angle. The relation between the torsion angle θ and the separation r of the centers of the spheres, including the correction factor, can be rewritten as a "fitting equation"

$$\theta = A[1 - 4(a/r)^e]$$

with unknown parameters, the scale factor A and the exponent e.

The students obtained the following measurements of the torsion angle (θ in degrees) as a function of the separation between the centers of the spheres (r in cm).

r_i	5.0	6.0	7.0	8.0	9.0	10.0	12.0	14.0	16.0	18.0	20.0
θ_i	264	233	179	136	111	84	63	53	33	30	27

Assume that the uncertainty in the angle is $\pm 1°$.

(*a*) Use one of the nonlinear fitting methods to determine the two parameters e and A of the fitting equation, and their uncertainties.

(*b*) Make a better estimate of the uncertainty in θ by considering the uncertainty required to give χ^2 = number of degrees of freedom.

(*c*) What effect does this change have on the uncertainties in the fitted parameters?