John Wolberg

Designing Quantitative Experiments

Prediction Analysis

Springer

Designing Quantitative Experiments

John Wolberg

Designing Quantitative Experiments

Prediction Analysis

Prof. Emeritus John Wolberg
Faculty of Mechanical Engineering
Technion – Israel Institute of Technology
Haifa, Israel
jwolber@attglobal.net

ISBN 978-3-642-11588-2 e-ISBN 978-3-642-11589-9
DOI 10.1007/978-3-642-11589-9
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010925034

Cover design: WMXDesign GmbH, Heidelberg

Cover illustration: Original painting (100 cm × 80 cm) by the author and currently in a private collection in Los Angeles

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

For my parents, Sidney and Beatrice Wolberg ז"ל

My wife Laurie

My children and their families:

Beth, Gilad, Yoni and Maya Sassoon
David, Pazit and Sheli Wolberg
Danny, Iris, Noa, Adi, Liat and Shani Wolberg
Tamar, Ronen, Avigail, Aviv and Hila Kimchi

Preface

Early in my career I was given the task of designing a sub-critical nuclear reactor facility that was to be used to perform basic research in the area of reactor physics. We planned to run a series of experiments to determine fundamental parameters related to the distribution of neutrons in such systems. I felt that it was extremely important to understand how the design would impact upon the accuracy of our results and as a result of this requirement I developed a design methodology that I subsequently called **prediction analysis**. After working with this method for several years and applying it to a variety of different experiments, I wrote a book on the subject. Not surprisingly, it was entitled ***Prediction Analysis*** and was published by Van Nostrand in 1967.

Since the book was published over 40 years ago science and technology have undergone massive changes due to the computer revolution. Not only has available computing power increased by many orders of magnitude, easily available and easy to use software has become almost ubiquitous. In the 1960's my emphasis was on the development of equations, tables and graphs to help researchers design experiments based upon some well-known mathematical models. When I reconsider this work in the light of today's world, the emphasis should shift towards applying current technology to facilitate the design process. The purpose of this book is to revisit prediction analysis with the emphasis upon application of available software to the design of quantitative experiments.

I should emphasize that quantitative experiments are performed in most branches of science and technology. Astronomers analyze data from asteroid sightings to predict orbits. Computer scientists develop models for improving network performance. Physicists measure properties of materials at low temperatures to understand superconductivity. Materials engineers study the reaction of materials to varying load levels to develop methods for prediction of failure. Chemical engineers consider reactions as functions of temperature and pressure. The list is endless. From the very small-scale work on DNA to the huge-scale study of black holes, quantitative experiments are performed and the data must be analyzed.

The common denominator in all this work is the similarity in the analysis phase of the experimental process. If one can assume that the measurement errors in the obtained data are normally distributed, the **method of least squares** is usually used to "fit" the data. The assumption of normality is usually reasonable so for this very broad class of experiments the method of least squares is the "best" method of analysis. The word "best" implies that the estimated parameters are determined with the smallest estimated uncertainty. Actually, the theoretically best solution to the minimization of estimated uncertainty is achieved by applying the **method of maximum likelihood**. This method was proposed as a general method of estimation by the renowned statistician R. A. Fisher in the early part of the 20th century. The method can be applied when the uncertainties associated with the observed or calculated data exhibit any type of distribution. However, when the uncertainties are normally distributed or when the normal distribution is a reasonable approximation, the method of maximum likelihood reduces to the method of least squares. The assumption of normally distributed random errors is reasonable for most situations and thus the method of least squares is applicable for analysis of most quantitative experiments. For problems in which the method of least squares will be applicable for analysis of the data, the method of prediction analysis is applicable for designing the proposed experiments.

Many of the examples of prediction analyses of experiments included in this book were done using the REGRESS program which is discussed in Section 3.10. The program is available free of charge and can be obtained through my website.

John Wolberg
www.technion.ac.il/wolberg
Haifa, Israel
Oct, 2009

Acknowledgments

I would like to thank a number of people who were kind enough to read sections of the book and offer corrections and suggestions for improvements: David Aronson, Yakov Ben Haim, Tomer Duman, Franco Elena, Michael Greenberg, Shimon Haber, Gad Hetsroni, Marshall Rafal and Honggang Zhao.

I've been teaching a graduate course at the Technion for over 20 years called ***Design and Analysis of Experiments***. The subject of Prediction Analysis has been included in the course from its inception and I have received many useful comments from the students throughout the years. A number of the students have used this technology to design real experiments as part of their graduate research and from their feedback I've been able to improve the design aspects of the REGRESS program.

As I mentioned in the Preface, I originally wrote a book called ***Prediction Analysis*** that was published early in my career. I just revisited the Acknowledgments section of that book and came across the names of those who helped me with that book many years ago: Sam Aronson, Yossi Etzion, Tsahi Gozani, Dan Ilberg, Mary Pozdena, Shlomo Shalev and Wolfgang Rothenstein. The original quote from that book is still quite perfect for this new version of this subject:

סוף מעשה במחשבה תחילה

The result of the deed comes from the thought at the beginning

Shlomo Halevi Elkavatz, *Circa* 1500

Table of Contents

Chapter 1 INTRODUCTION

1.1 The Experimental Method

We can consider the experimental method as consisting of four distinct phases:

1) Design
2) Execution
3) Data Analysis
4) Interpretation

The design phase is partially problem dependent. For example, we require answers to questions regarding the choice of equipment and the physical layout of the experimental setup. However, there are also questions of a more generic nature: for example, how many data points are needed and what are the accuracy requirements of the data (i.e., how accurately must we measure or compute or otherwise obtain the data)?

The execution phase is completely problem dependent. The performance of an experiment is usually a physical process although we often see computer experiments in which data is "obtained" as the result of computer calculations or simulations. The purpose of the execution phase is to run the experiment and obtain data.

The data analysis phase is typically independent of the details of the physical problem. Once data has been acquired and a mathematical model has been proposed, the actual analysis is no longer concerned with what the data represents. For example, assume that we have obtained values of a dependent variable ***Y*** as a function of time ***t***. We propose a mathematical model that relates ***Y*** to ***t*** and the ensuing analysis considers the values of ***Y*** and ***t*** as just a collection of numbers.

J. Wolberg, *Designing Quantitative Experiments*, DOI 10.1007/978-3-642-11589-9_1,

The interpretation phase is really a function of what one hopes to accomplish. There are a number of reasons why the experiment might have been performed. For example, the purpose might have been to prove the validity of the mathematical model. Alternatively, the purpose might have been to measure the parameters of the model. Yet another purpose might have been to develop the model so that it could be used to predict values of the dependent variable for combinations of the independent variables. To decide whether or not the experiment has been successful one usually considers the resulting accuracy. Is the accuracy of results sufficient to meet the criteria specified in the design phase? If not what can be done to improve the results?

1.2 Quantitative Experiments

The subject of this book is the design of quantitative experiments. We define quantitative experiments as experiments in which data is obtained for a dependent variable (or variables) as a function of an independent variable (or variables). The dependent and independent variables are then related through a mathematical model. These are experiments in which the variables are represented numerically.

One of the most famous quantitative experiments was performed by an Italian astronomer by the name of Giuseppe Piazzi of Palermo. In the late 1700's he set up an astronomical observatory which afforded him access to the southernmost sky in Europe at that time. His royal patron allowed him to travel to England where he supervised the construction of a telescope that permitted extremely accurate observations of heavenly bodies. Once the telescope was installed at the Royal Observatory in Palermo, Piazzi started work on a star catalog that was the most accurate that had been produced up to that time. In 1801 he stumbled across something in the sky that at first he thought was a comet. He took reading over a 42 day period when weather was permitting and then the object faded from view. Eventually it was recognized that the object was a planetoid in an orbit between Mars and Jupiter and he named it Ceres.

What made the discovery of Ceres such a memorable event in the history of science is the analysis that Gauss performed on the Piazzi data. To perform the analysis, Gauss developed the method of least squares which he

described in his famous book ***Theoria Motus***. A translation of this book was published in English in 1857 [GA57]. By applying the method to calculate the orbit of Ceres, Gauss was able to accurately predict the reappearance of the planetoid in the sky. The method of least squares has been described as "the most nontrivial technique of modern statistics" [St86]. To this day, analysis of quantitative experiments relies to an overwhelming extent on this method. A simulation of Piazzi's experiment is included in Section 6.6 of this book.

1.3 Dealing with Uncertainty

The estimation of uncertainty is an integral part of data analysis. Typically uncertainty is expressed quantitatively as a value such as σ (the standard deviation) of whatever is being measured or computed. (The definition of σ is discussed below.) With every measurement we should include some indication regarding accuracy of the measurement. Some measurements are limited by the accuracy of the measuring instrument. For example, digital thermometers typically measure temperature to accuracies of 0.1°C. However, there are instruments that measure temperature to much greater accuracies. Alternatively, some measurements are error free but are subject to probability distributions. For example, consider a measurement of the number of people affected by a particular genetic problem in a group of 10000 people. If we examine all 10000 people and observe that twenty people test positive, what is our uncertainty? Obviously for this particular group of 10000 people there is no uncertainty in the recorded number. However, if we test a different group of 10000 people, the number that will test positive will probably be different than twenty. Can we make a statement regarding the accuracy of the number 20?

One method for obtaining an estimate of uncertainty is to repeat the measurement n times and record the measured values x_i, $i = 1$ to n. We can estimate σ (the standard deviation of the measurements) as follows:

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{i=n} (x_i - x_{avg})^2 \qquad (1.3.1)$$

In this equation x_{avg} is the average value of the n measurements of x. A qualitative explanation for the need for n-1 in the denominator of this equ-

ation is best understood by considering the case in which only one measurement of x is made (i.e., $n = 1$). For this case we have no information regarding the "spread" in the measured values of x. A detailed derivation of this equation is included in most elementary books on statistics. The implication in Equation 1.3.1 is that we need to repeat measurements a number of times in order to obtain estimates of uncertainties. Fortunately this is rarely the case.

Often the instrument used to perform the measurement is provided with some estimation of the uncertainty of the measurements. Typically the estimation of σ is provided as a fixed percentage (e.g., $\sigma = 1\%$ of the value of x) or a fixed value (e.g., $\sigma = 0.1°C$). Sometimes the uncertainty is dependent upon the value of the quantity being measured in a more complex manner than just a fixed percentage or a constant value. For such cases the provider of the measuring instrument might supply this information in a graphical format or perhaps as an equation. For cases in which the data is calculated rather than measured, the calculation is incomplete unless it is accompanied by some estimate of uncertainty.

For measurements that are subject to statistical distributions like the example cited above regarding the genetic problem per 10000 people, often a knowledge of the distribution is sufficient to allow us to estimate the uncertainty. For that particular problem we could assume a Poisson distribution (discussed in Section 2.4) and the estimated value of σ is the square root of the number of people diagnosed as positive (i.e., $\sqrt{20} = 4.47$). We should note that our measurement is only accurate to a ratio of 4.47/20 which is approximately 22%. If we increase our sample size to 100,000 people and observe about 200 with the genetic problem, our measurement accuracy would be about $\sqrt{200} = 14.1$ which is approximately 7%. This is an improvement of more than a factor of 3 in fractional accuracy but at an increase in the cost for running the experiment by a factor of 10!

Once we have an estimation of σ, how do we interpret it? In addition to σ, we have a result (i.e., the value of whatever we are trying to determine) either from a measurement or from a calculation. Let us define the result as x and the true (but unknown value) of what we are trying to measure or compute as μ. Typically we assume that our best estimate of this true value of μ is x and that μ is located within a region around x. The size of the region is characterized by σ. In the preceding example our result x was 20 and the estimated value of σ was 4.47. This implies that if we were able to determine the true value of μ there is a "large" probability that it would be

somewhere in the range 15 to 25. How "large" is a question that is considered in the discussion of probability distributions in Section 2.4. A typical assumption is that the probability of μ being greater or less than x is the same. In other words, our measurement or calculation includes a random error characterized by σ. Unfortunately this assumption is not always valid!

Sometimes our measurements or calculations are corrupted by **systematic errors**. Systematic errors are errors that cause us to either systematically under-estimate or over-estimate our measurements or computations. One source of systematic errors is an unsuccessful calibration of a measuring instrument. Another source is failure to take into consideration external factors that might affect the measurement or calculation (e.g., temperature effects). The choice of the mathematical model can also lead to systematic errors if it is overly simplistic or if it includes erroneous constants. Data analysis of quantitative experiments is based upon the assumption that the measured or calculated variables are not subject to systematic errors and that the mathematical model is a true representation of the process being modeled. If these assumptions are not valid, then errors are introduced into the results that do not show up in the computed values of the σ's. One can modify the least squares analysis to study the sensitivity of the results to systematic errors but whether or not systematic errors exist is a fundamental issue in any work of an experimental nature.

1.4 Parametric Models

Quantitative experiments are usually based upon parametric models. In this discussion we define **parametric models** as models utilizing a mathematical equation that describes the phenomenon under observation. As an example of a quantitative experiment based upon a parametric model, consider an experiment shown schematically in Figure 1.4.1. In this experiment a radioactive source is present and a detector is used to monitor radiation emanating from the source. Each time a radioactive particle enters the detector an "event" is noted. The recorder counts the number of events observed within a series of user specified time windows and thus produces a record of the number of counts observed as a function of time. The purpose of this experiment is to measure the half-life of a radioactive isotope. There is a single dependent variable ***counts*** (number of counts per unit of time) and a single independent variable ***time***. The parametric model for this particular experiment is:

$$counts = a_1 \cdot e^{-a_2 \cdot time} + a_3 \tag{1.4.1}$$

This model has 3 parameters: the amplitude a_1, the decay constant a_2 and the background count rate a_3. The decay constant is related to the half-life (the time required for half of the isotope atoms to decay) as follows:

$$e^{-a_2 \cdot half\_life} = 1/2$$

$$half\_life = \frac{\ln(2)}{a_2} = \frac{0.69315}{a_2} \tag{1.4.2}$$

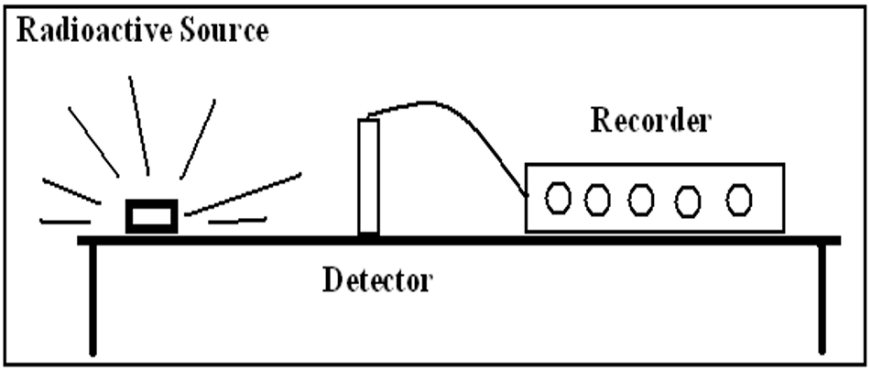

Figure 1.4.1 Experiment to Measure Half-life of a Radioisotope

The model equation (or equations) contains unknown parameters and the purpose of the experiment is often to determine the parameters including some indication regarding the accuracies (i.e., values of the σ's) of these parameters. There are many situations in which the values of the individual parameters are of no interest. All that is important for these cases is that the model can be used to predict values of the dependent variable (or variables) for other combinations of the independent variables. In addition, we are also interested in some measure of the accuracy (i.e.,σ) of the predictions.

We need to use mathematical terminology to define parametric models. Let us use the term y to denote the dependent variable and x to denote the independent variable. Usually y is a scalar, but when there is more than

one dependent variable, y can denote a vector. The parametric model is the mathematical equation that defines the relationship between the dependent and independent variables. For the case of a single dependent and a single independent variable we can denote the model as:

$$y = f(x; a_1, a_2 .., a_p) \tag{1.4.3}$$

The a_k's are the p unknown parameters of the model. The function f is based on either theoretical considerations or perhaps it is a function that seems to fit the measured values of y and x. Equation 1.4.1 is an example of a model in which p=3. The dependent variable y is ***counts*** and the independent variable x is ***time***.

When there is more than one independent variable, we can use the following equation to denote the model:

$$y = f(x_1, x_2 .., x_m; a_1, a_2 .., a_p) \tag{1.4.4}$$

The x_j's are the m independent variables. As an example of an experiment in which there is more than one independent variable, consider an experiment based upon the layout shown in Figure 1.4.2. In this experiment a grid of detectors is embedded in a block of material. A source of radioactivity is placed near the block of material and the level of radioactivity is measured at each of the detectors. The count rate at each detector is a function of the position (x_1, x_2) within the block of material. The unknown parameters of the model are related to the radiation attenuation properties of the material.

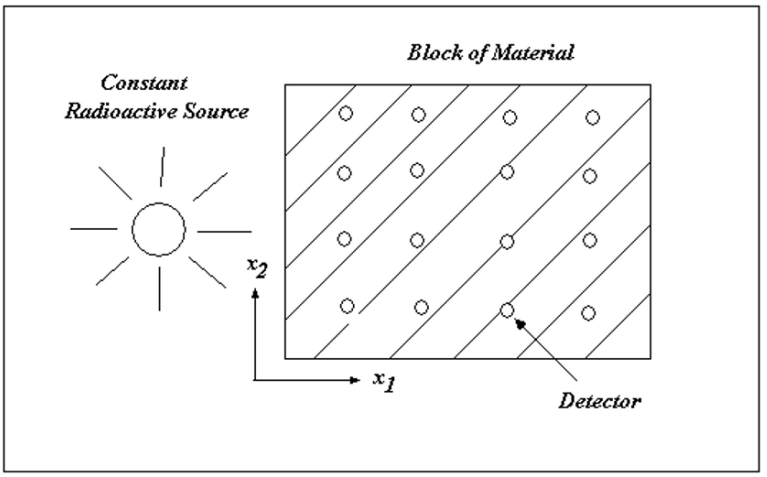

Figure 1.4.2 - Experiment to measure radioactive attenuation

If there is more than one dependent variable, we require a separate function for each element of the ***y*** vector:

$$y_l = f_l(x_1, x_2 .., x_m; a_1, a_2 .., a_p) \quad l = 1 \text{ to } d \tag{1.4.5}$$

For cases of this type, ***y*** is a ***d*** dimensional vector and the subscript ***l*** refers to the l^{th} term of the ***y*** vector. It should be noted that some or all of the x_j's and the a_k's may be included in each of the ***d*** equations. The notation for the i^{th} data point for this l^{th} term of the ***y*** vector would be:

$$y_{l_i} = f_l(x_{1_i}, x_{2_i} .., x_{m_i}; a_1, a_2 .., a_p)$$

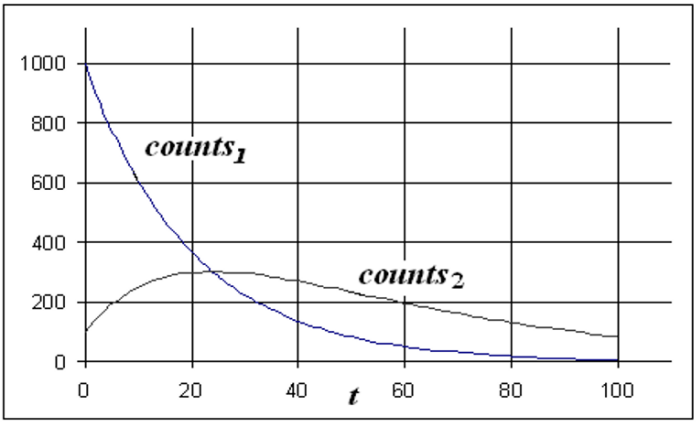

Figure 1.4.3 ***Counts*** **versus Time for Equations 1.4.6 and 1.4.7**
a_1**=*1000*,** a_2**=*100*,** c_1**=*0.05*,** c_2**=*0.025***

An example of an experiment requiring a model of the form of Equation 1.4.5 can also be based upon the layout shown in Figure 1.4.1. If we assume that there are two radioactive species in the source and that species 2 is the daughter product of the species 1, we can measure the number of counts emanating from each species by discriminating the counts based upon the differing energies of the particles reaching the detector from the two species. Assuming that the two count rates (***counts*****$_1$** and ***counts*****$_2$**) are corrected for their background count rates, they are related to time ***t*** as follows:

$$\boldsymbol{counts_1 = a_1 \cdot e^{-c_1 \cdot t}} \tag{1.4.6}$$

$$\boldsymbol{counts_2 = a_2 \cdot e^{-c_2 \cdot t} + a_1 \frac{c_2}{c_2 - c_1}\left(e^{-c_1 \cdot t} - e^{-c_2 \cdot t}\right)} \tag{1.4.7}$$

For this example, there are four unknown parameters: the initial amplitudes $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$, and the decay constants $\boldsymbol{c_1}$ and $\boldsymbol{c_2}$. A typical plot of ***counts*****$_1$** and ***counts*****$_2$** versus ***t*** is shown in Figure 1.4.3. A least squares analysis of the data will determine the values of these four parameters and estimates of their standard deviations.

A model is recursive if the functions defining the dependent variables $\boldsymbol{y_l}$ are interdependent. The form for the elements of recursive models is as follows:

$$y_l = f_l(x_1, x_2 .., x_m; y_1, y_2 .., y_d; a_1, a_2 .., a_p) \tag{1.4.8}$$

An example of a recursive model is the well-known prey-predator model of Kolmogorov [HO06, BR01, FR80]:

$$\frac{dy_1}{dt} = y_1 f_1(y_1, y_2) \tag{1.4.9}$$

$$\frac{dy_2}{dt} = y_2 f_2(y_1, y_2) \tag{1.4.10}$$

where y_1 is the prey population and y_2 is the predator population. The famous Italian mathematician Vito Volterra proposed a simple model to represent predator-prey interactions:

$$f_1 = a_1 - a_2 y_2 \tag{1.4.11}$$

$$f_2 = a_3 y_1 - a_4 \tag{1.4.12}$$

The parameter a_1 is the prey growth rate in the absence of predators and a_4 is the predator death rate in the absence of prey. The parameters a_2 and a_3 are the interaction coefficients. Increasing the predator population (i.e., y_2) causes a decrease in the prey population (i.e., y_1) and visa versa. Both of these equations are recursive: there is one independent variable t, four unknown parameters (a_1 to a_4) and two dependent variables (y_1 and y_2). We see that y_1 is dependent upon y_2 and y_2 is dependent upon y_1. The solution of Equations 1.4.9 and 1.4.10 introduces 2 new parameters: the initial values of y_1 and y_2 (i.e., y_{10} and y_{20}):

$$y_1 = y_{10} + \int y_1 f_1(y_1, y_2) dt \tag{1.4.13}$$

$$y_2 = y_{20} + \int y_2 f_2(y_1, y_2) dt \tag{1.4.14}$$

These two parameters can be treated as known constants or unknown parameters that are determined as part of the analysis of the data. Once a parametric model has been proposed and data is available, the task of data analysis must be performed. There are several possible objectives of interest to the analyst:

1) Compute the values of the p unknown parameters $a_1, a_2, \ldots a_p$.

2) Compute estimates of the standard deviations of the p unknown parameters.
3) Use the p unknown parameters to compute values of y for desired combinations of the independent variables $x_1, x_2, \ldots x_m$.
4) Compute estimates of the standard deviations σ_f for the values of $y = f(x)$ computed in 3.

It should be mentioned that the theoretically best solution to all of these objectives is achieved by applying the **method of maximum likelihood**. This method was proposed as a general method of estimation by the renowned statistician R. A. Fisher in the early part of the 20th century [e.g., FR92]. The method can be applied when the uncertainties associated with the observed or calculated data exhibit any type of distribution. However, when these uncertainties are normally distributed or when the normal distribution is a reasonable approximation, the method of maximum likelihood reduces to the **method of least squares** [WO06, HA01]. Fortunately, the assumption of normally distributed random errors is reasonable for most situations and thus the method of least squares is applicable for analysis of most quantitative experiments.

1.5 Basic Assumptions

The method of least squares can be applied to a wide variety of analyses of experimental data. The common denominator for this broad class of problems is the applicability of several basic assumptions. Before discussing these assumptions let us consider the measurement of a dependent variable Y_i. For the sake of simplicity, let us assume that the model describing the behavior of this dependent variable includes only a single independent variable. Using Equation 1.4.3 as the model that describes the relationship between x and y then y_i is the computed value of y at x_i. We define the difference between the measured and computed values as the residual R_i:

$$Y_i = y_i + R_i = f(x_i; a_1, a_2, \ldots a_p) + R_i \tag{1.5.1}$$

It should be understood that neither Y_i nor y_i are necessarily equal to the true value η_i. In fact there might not be a single true value if the dependent variable can only be characterized by a distribution. However, for the sake of simplicity let us assume that for every value of x_i there is a unique true value (or a unique mean value) of the dependent variable that is η_i. The difference between Y_i and η_i is the error (or uncertainty) ε_i:

$$Y_i = \eta_i + \varepsilon_i \tag{1.5.2}$$

The method of least squares is based upon the following assumptions:

1) If the measurement at x_i were to be repeated many times, then the values of error ε_i would be normally distributed with an average value of zero. Alternatively, if the errors are not normally distributed, the approximation of a normal distribution is reasonable.

2) The errors are uncorrelated. This is particularly important for time-dependent problems and implies that if a value measured at time t_i includes an error ε_i and at time t_{i+k} includes an error ε_{i+k} these errors are not related (i.e., uncorrelated). Similarly, if the independent variable is a measure of location, then the errors at nearby points are uncorrelated.

3) The standard deviations σ_i of the errors can vary from point to point. This assumption implies that σ_i is not necessarily equal to σ_j.

The implication of the first assumption is that if the measurement of Y_i is repeated many times, the average value of Y_i would be the true (i.e., errorless) value η_i. Furthermore, if the model is a true representation of the connection between y and x and if we knew the true values of the unknown parameters the residuals R_i would equal the errors ε_i:

$$Y_i = \eta_i + \varepsilon_i = f(x_i; \alpha_1, \alpha_2, \ldots \alpha_p) + \varepsilon_i \tag{1.5.3}$$

In this equation the true value of the a_k is represented as α_k. However, even if the measurements are perfect (i.e., $\varepsilon_i = 0$), if f does not truly describe the dependency of y upon x, then there will certainly be a difference between the measured and computed values of y.

The first assumption of normally distributed errors is usually reasonable. Even if the data is characterized by other distributions (e.g., the binomial or Poisson distributions), the normal distribution is often a reasonable approximation. But there are problems where an assumption of normality causes improper conclusions. For example, in financial risk analysis the probability of catastrophic events (for example, the financial meltdown in

the mortgage market in 2008) might be considerably greater than one might predict using normal distributions. To cite another area, earthquake predictions require analyses in which normal distributions cannot be assumed. Yet another area that is subject to similar problems is the modeling of insurance claims. Most of the data represents relatively small claims but there are usually a small fraction of claims that are much larger, negating the assumption of normality. Problems in which the assumption of normal error distributions is invalid are beyond the scope of this book. However, there is a large body of literature devoted to this subject. An extensive review of the subject is included in a book by Yakov Ben Haim [BE06].

One might ask when the second assumption (i.e., uncorrelated errors) is invalid. There are areas of science and engineering where this assumption is not really reasonable and therefore the method of least squares must be modified to take error correlation into consideration. Davidian and Giltinan discuss problem in the biostatistics field in which repeated data measurements are taken [DA95]. For example, in clinical trials, data might be taken for many different patients over a fixed time period. For such problems we can use the term Y_{ij} to represent the measurement at time t_i for patient j. Clearly it is reasonable to assume that ε_{ij} is correlated with the error at time t_{i+1} for the same patient. In this book, no attempt is made to treat such problems.

Many statistical textbooks include discussions of the method of least squares but use the assumption that all the σ_i's are equal. This assumption is really not necessary as the additional complexity of using varying σ_i's is minimal. Another simplifying assumption often used is that the models are linear with respect to the a_k's. This assumption allows a very simple mathematical solution but is too limiting for the analysis of many real-world experiments. This book treats the more general case in which the function f (or functions f_l) can be nonlinear.

1.6 Treatment of Systematic Errors

I did my graduate research in nuclear science at MIT. My thesis was a study of the fast fission effect in heavy water nuclear reactors and I was reviewing previous measurements [WO62]. The fast fission effect had been measured at two national laboratories and the numbers were curiously different. Based upon the values and quoted σ's, the numbers were

many σ's apart. I discussed this with my thesis advisors and we agreed that one or both of the experiments was plagued by systematic errors that biased the results in a particular direction. We were proposing a new method which we felt was much less prone to systematic errors.

Results of experiments can often be misleading. When one sees a result stated as 17.3 ± 0.5 the reasonable assumption is that the true value should be somewhere within the range 16.8 to 17.8. Actually, if one can assume that the error is normally distributed about 17.3, and if 0.5 is the estimated standard deviation of the error, then the probability of the true value falling within the specified range is about 68%. The assumption of a normally distributed error centered at the measured value is based upon an assumption that the measurement is not effected by systematic errors. If, however, one can place an upper limit on all sources of systematic errors, then the estimated standard deviation can be modified to include treatment of systematic errors. By including systematic errors in the estimated standard deviation of the results, a more realistic estimate of the uncertainty of a measurement can be made.

As an example of an experiment based upon a single independent variable and a single dependent variable, consider the first experiment discussed in Section 1.4: the measurement of the half-life of a radioactive species. The mathematical model for the experiment is Equation 1.4.1 but let us assume that the background radiation is negligible (i.e., a_3 is close to zero). We could then use the following mathematical model:

$$counts = a_1 \cdot e^{-a_2 \cdot time} \tag{1.6.1}$$

The fact that we have neglected to include a background term is a potential source of a systematic error. For this example, the source of the systematic error is the choice of the mathematical model itself. The magnitude of this systematic error can be estimated using the same computer program that is used to compute the values of the unknown parameters (i.e., a_1 and a_2).

Probably the greatest sources of systematic errors are measurement errors that introduce a bias in one direction or the other for the independent and dependent variables. One of the basic assumptions mentioned in the previous section is that the errors in the data are random about the true values. In other words, if a measurement is repeated n times, the average value would approach the true value as n becomes large. However, what happens if this assumption is not valid? When one can establish maximum

possible values for such errors, the effect of these errors on the computed values of the parameters can be estimated by simple computer simulations. An example of this problem is included in Section 5.4.

Another source of systematic errors is due to errors in the values of parameters treated as known constants in the analysis of the data. As an example of this type of error consider the constants y_{10} and y_{20} in Equations 1.4.13 and 1.4.14. If these constants are treated as input quantities, then if there is uncertainty associated with their values, these uncertainties are a potential source of systematic errors. Estimates for the limits of these systematic errors can be made using the same software used for analysis of the data. All that needs to be done is to repeat the analysis for the range of possible values of the constants (e.g., y_{10} and y_{20}). This procedure provides a direct measurement of the effect of these sources of uncertainty on the resulting values of the unknown parameters (e.g., a_1 through a_4).

We can make some statements about combining estimates of systematic errors. Let us assume that we have identified ***nsys*** sources of systematic errors and that we can estimate the maximum size of each of these error sources. Let us define ε_{jk} as the systematic error in the measurement of a_j caused by the k^{th} source of systematic errors. The magnitude of the value of ε_j (the magnitude of the systematic error in the measurement of a_j caused by all sources) could range from zero to the sum of the absolute values of all the ε_{jk} 's. However, a more realistic estimate of ε_j is the following:

$$\varepsilon_j^2 = \sum_{k=1}^{k=nsys} \varepsilon_{jk}^2 \tag{1.6.2}$$

This equation is based upon the assumption that the systematic errors are uncorrelated. The total estimated uncertainty σ_j for the variable a_j should include the computed estimated standard deviation from the least squares analysis plus the estimated systematic error computed using Equation 1.6.2.

$$\sigma_j^2 = \sigma_{aj}^2 + \varepsilon_j^2 \tag{1.6.3}$$

Once, again this equation is based upon an assumption that the least squares estimated standard deviation and the estimated systematic error are uncorrelated.

1.7 Nonparametric Models

There are situations in which attempts to describe the phenomenon under observation by a single equation is extremely difficult if not impossible. For example, consider a dependent variable that is the future percentage return on stocks traded on the NYSE (New York Stock Exchange). One might be interested in trying to find a relationship between the future returns and several indicators that can be computed using currently available data. For this problem there is no underlying theory upon which a parametric model can be constructed. A typical approach to this problem is to use the historic data to define a surface and then use some sort of smoothing technique to make future predictions regarding the dependent variable [WO00]. The data plus the algorithm used to make the predictions are the major elements in what we define as a **nonparametric model.**

Nonparametric methods of data modeling predate the modern computer era [WO00]. In the 1920's two of the most well-known statisticians (Sir R. A. Fisher and E. S. Pearson) debated the value of such methods [HA90]. Fisher correctly pointed out that a parametric approach is inherently more efficient. Pearson was also correct in stating that if the true relationship between $\boldsymbol{X}$ and $\boldsymbol{Y}$ is unknown, then an erroneous specification in the function $\boldsymbol{f(X)}$ introduces a model bias that might be disastrous.

Hardle includes a number of examples of successful nonparametric models [HA90]. The most impressive is the relationship between change in height (cm/year) and age of women (Figure 1.7.1). A previously undetected growth spurt at around age 8 was noted when the data was modeled using a nonparametric smoother [GA84]. To measure such an effect using parametric techniques, one would have to anticipate this result and include a suitable term in $\boldsymbol{f(X)}$.

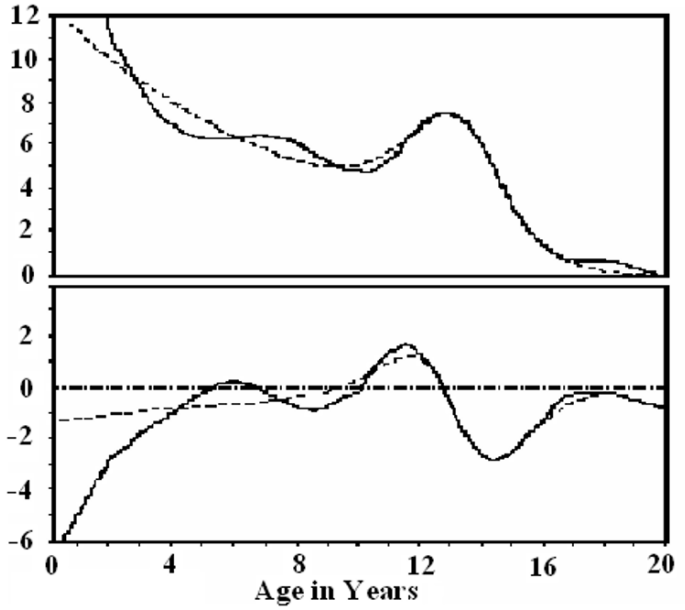

Figure 1.7.1 Human growth in women versus Age. The top graph is in cm/year. The bottom graph is acceleration in cm/year^2. The solid lines are from a model based upon nonparametric smoothing and the dashed lines are from a parametric fit [GA84, HA90].

The point at which one decides to give up attempts to develop a parametric model and cross over to nonparametric modeling is not obvious. For problems that are characterized by a large set of candidate predictors (i.e., predictors that might or might not be included in the final model) nonparametric modeling techniques can be used in an effort to seek out information rich subsets of the candidate predictor space. For example, when trying to model financial markets, one may consider hundreds of candidate predictors [WO00]. Financial market predictions are, of course, an area of intense world-wide interest. As a result there is considerable interest in applying nonparametric methods to the development of tools for making financial market predictions. A number of books devoted to this subject have been written in recent years (e.g., [AZ94, BA94, GA95, RE95, WO00, HO04, MC05]). If the nonparametric methods can successfully reduce the set of candidate predictors to a much smaller subset then a parametric approach to modeling might be possible. For such cases the design techniques considered in this book are applicable. However, if the experimen-

tal data will only be modeled using nonparametric techniques, the tion analysis approach to design is not applicable.

1.8 Statistical Learning

The term **statistical learning** is used to cover a broad class of methods and problems that have become feasible as the power of the computer has grown. An in-depth survey of this field is covered in an excellent book by Hastie, Tibshirani and Friedman entitled ***The Elements of Statistical Learning: Data Mining, Inference and Prediction*** [HA01]. Their book covers both supervised and unsupervised learning. The goal of supervised learning is to predict an output variable as a function of a number of input variables (or as they are sometimes called: indicators or predictors). In unsupervised learning there is no particular output variable and one is interested in finding associations and patterns among the variables. The cornerstone of statistical learning is to *learn from the data.* The analyst has access to data and his or her goal is to make sense out of the available information.

Supervised learning problems can be subdivided into **regression** and **classification** problems. The goal in regression problems is to develop quantitative predictions for the dependent variable. The goal in classification problems is to develop methods for predicting to which class a particular data point belongs. An example of a regression problem is the development of a model for predicting the unemployment rate as a function of economic indictors. An example of a classification problem is the development of a model for predicting whether or not a particular email message is a spam message or a real message. In this book, the emphasis is on regression rather than classification problems. The design methods discussed in this book are not applicable for classification problems.

Chapter 2 STATISTICAL BACKGROUND

2.1 Experimental Variables

An experimental variable is measured, observed or calculated (for example, in a computer experiment). In statistics the terms **population** and **sample** are used when discussing variables. For example, consider all boys in a particular city aged 12 years old. There is a mean (or average) height of these boys which is called the **population mean value**. To estimate this value, an experiment can be performed in which a random sample of boys of this age is selected and a **sample mean value** of their heights is determined. Clearly, the greater number of boys included in the sample, the "better" the estimate of the population mean. The word "better" implies that the estimated uncertainty in the calculation is smaller. The heights of the boys in the population and in the sample are **distributed** about their mean values. For this particular variable, the population mean value is really only an instantaneous value because the population is continually changing. Thus this variable is slightly time dependent.

An experimental variable can be classified as either **discrete** or **continuous**. An example of a discrete variable is the number of people observed to have a certain genetic profile in a sample group of 1000. The number of people with the genetic profile is a discrete number and does not change for the sample group, however, if we select a new sample group, the number observed will also be a discrete number but not necessarily the same number as in the original group. We can say that the number observed in any group of 1000 people is a **distributed discrete variable**. This discrete variable can be used to estimate a **continuous variable**: the fraction of all people that would test positive for the genetic profile.

Defining a continuous variable is not as clear-cut. Theoretically a continuous variable is defined as a variable that can have any value within a particular range. For example, consider the temperature of water which

J. Wolberg, *Designing Quantitative Experiments*, DOI 10.1007/978-3-642-11589-9_2,

can vary from 0 to 100 °C. If we measure the temperature of the water using a thermometer that is accurate to 1 °C then there are 101 possible values that might be observed and we can consider temperature measured in this way as a discrete variable. However, if we had a more accurate thermometer, then we could observe many more values of temperature. For example, if the accuracy of the thermometer is increased to 0.1 °C then 1001 values are possible. Typically, such variables are treated as **continuous distributed variables** that can assume any value within a specified range. Repeated measurements of continuous variables typically contain random variations due to accuracy limitations of the measuring instrument and random processes affecting the measurement (e.g., noise in the electronics associated with the measurement, temperature effects, etc.).

To characterize the distribution of a variable we use the notation $\Phi(x)$. If x is a discrete variable then the following condition must be satisfied:

$$\sum_{xmin}^{xmax} \Phi(x) = 1 \tag{2.1.1}$$

If x is a continuous variable:

$$\int_{xmin}^{xmax} \Phi(x)dx = 1 \tag{2.1.2}$$

There are several parameters used to characterize all distributions. The most well known and useful parameters are the **mean** μ and the **variance** σ^2. The **standard deviation** σ is the square root of the variance. For discrete distributions they are defined as follows:

$$\mu = \sum_{xmin}^{xmax} x\Phi(x) \tag{2.1.3}$$

$$\sigma^2 = \sum_{xmin}^{xmax} (x-\mu)^2\Phi(x) \tag{2.1.4}$$

For continuous distributions:

$$\mu = \int_{xmin}^{xmax} x\Phi(x)dx \tag{2.1.5}$$

$$\sigma^2 = \int_{xmin}^{xmax} (x-\mu)^2\Phi(x)dx \tag{2.1.6}$$

2.2 Measures of Location

The **mean** is a **measure of location** and there are several other measures of location that are sometimes used. The **median** is defined as the value of x in which the probability of x being less than the median is equal to the probability of x being greater than the median:

$$\int_{xmin}^{xmed} \Phi(x)dx = \int_{xmed}^{xmax} \Phi(x)dx = 0.5 \tag{2.2.1}$$

The median ***xmed*** can be defined in a similar manner for discrete variables. An example of the computation of the median for a discrete variable is included in the discussion associated with Figure 2.2.2 below.

Another measure of location is the **mode** which is defined as the value of x for which $\Phi(x)$ is maximized. As an example, consider the distribution:

$$\Phi(x) = \frac{x}{c^2} e^{-x/c} \tag{2.2.2}$$

This is a distribution with range $0 \le x \le \infty$ as it satisfies Equation 2.1.2. This can be show from the following integral:

$$\int_0^\infty x^n e^{-x/c} dx = n!c^{n+1} \tag{2.2.3}$$

For n=1 this equation yields a value of c^2. Substituting this value into Equation 2.1.2 we see that the integral over the range of possible values (i.e., 0 to ∞) equals one and thus Equation 2.2.2 is a continuous distribution in this range. We can get the mean of the distribution by substituting 2.2.2 into 2.1.5:

$$\mu = \frac{1}{c^2}\int_0^\infty x^2 e^{-x/c} dx = \frac{1}{c^2} 2! c^{2+1} = 2c \tag{2.2.4}$$

It can be shown that the median for this distribution is 1.68c and the mode is c. (For this particular distribution, to determine the mode, set the derivative of 2.2.2 with respect to x to zero and solve for x.) The distribution and its measures of location are shown in Figure 2.2.1. (Note: This distribution is a special case of a class called **Gamma Distributions** [JO70] which are discussed in Section 2.4.)

Why do we consider measures of location other than the mean? For some variables the median is more meaningful than the mean value. For example, when we consider the salary paid to a particular group of people, the median salary is usually more meaningful than the mean salary. As an example, consider the mean salary paid at most companies. Often the salary of the top executives is very large compared to the vast majority of the other employees. To judge whether or not a particular company adequately compensates its employees, the median salary is a more meaningful number than the mean salary. The mode is the preferred measure of location when we are interested in the most probable value of a variable.

We often use the term **average** interchangeably with mean. To be more exact, the term average is sometimes called the **unbiased estimate** of the mean. If, for example, a variable x is measured n times the average x_{avg} is computed as:

$$x_{avg} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.2.5}$$

If the measurements of x are unbiased, then the average value approaches the population mean value μ as n approaches infinity.

The concept of a mean, median and mode of a distribution can also be applied to discrete distributions. As an example, we can again use the num-

ber of people testing positive for a genetic profile in a sample group of 1000. Let us assume that 10 different groups are tested and the number of positives in each group are 20, 23, 16, 20, 19, 20, 18, 19, 25, and 18. The data can be shown graphically using a **histogram** as seen in Figure 2.2.2. The estimate of the mean of this distribution is the average of the 10 measurements which is 19.8. The estimated mode for this distribution is 20 and the estimated median can be determined from Figure 2.2.2. Half the area of the histogram is in the region $x <= 19$ and half is in the region $x >= 20$ so the estimated median is 19.5. We use the term **estimated** for these three measures of location because they are not the true values of the mean, median and mode. In fact, for this particular discrete variable there is no true value! There is a true mean value for the fraction of all people on earth testing positive for this genetic profile, but when translated to the number in a sample group of 1000 people, this would no longer be a discrete variable. For the sake of brevity, we will use the terms mean, median and mode to imply the unbiased estimates of the parameters when applicable.

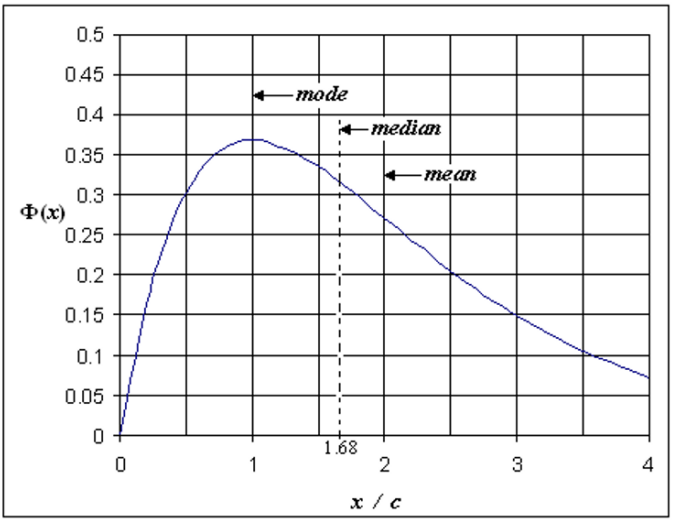

Figure 2.2.1 $\Phi(x)$ **versus** x/c **for** $\Phi(x) = xe^{-x/c}/c^2$

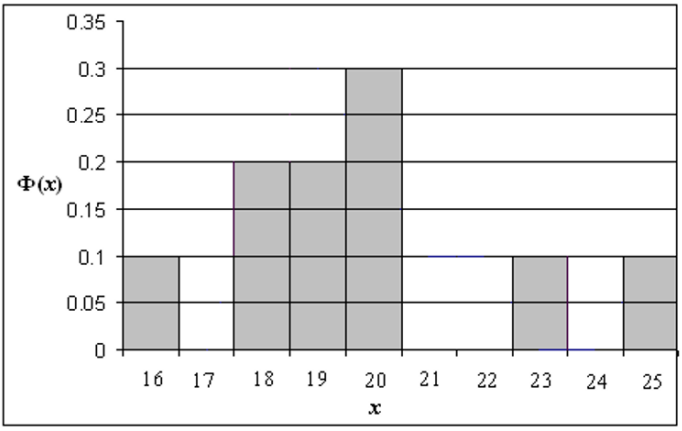

Figure 2.2.2 $\Phi(x)$ versus x for 10 measurement of the number of positives in genetic profile testing of groups of 1000 people.

2.3 Measures of Variation

The variance and standard deviation are **measures of variation**. Equations 2.1.4 and 2.1.6 are used to compute these parameters for discrete and continuous distributions. These equations imply that μ (the population mean of the distribution) is known. For most cases all we have is a sample mean so the best we can do is determine an estimate of the variance. Using the notation s_x^2 as the unbiased estimate of the variance of x, we compute it as follows for a discrete variable:

$$s_x^2 = \frac{1}{n-1} \sum_{x=xmin}^{xmax} n_x (x - x_{avg})^2 \tag{2.3.1}$$

In this equation n_x is the number of values of x and n is the total number of all x's from $xmin$ to $xmax$. For a continuous variable:

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - x_{avg})^2 \tag{2.3.2}$$

The need for the term n-1 in the denominator of these equations is necessary because at least two measurements or observations of x are required to determine an estimate of the variance. A single value gives no information regarding the "spread" in the data.

Knowledge of μ and σ do not completely describe a distribution, although for most cases they provide a fairly accurate description. In Figure 2.3.1 we see three distributions each with the same value of μ and σ and yet they are quite different. The middle distribution is symmetrical about its mean but the other two are skewed – one to the left and the other to the right. Another parameter of distributions is know as the **skewness** and is based upon the 3rd moment about the mean:

$$skewness = \frac{1}{\sigma^3}\sum_{xmin}^{xmax}(x-\mu)^3\Phi(x) \quad \text{(discrete distribution)}$$

$$skewness = \frac{1}{\sigma^3}\int_{xmin}^{xmax}(x-\mu)^3\Phi(x)dx \quad \text{(continuous dist.)} \quad (2.3.3)$$

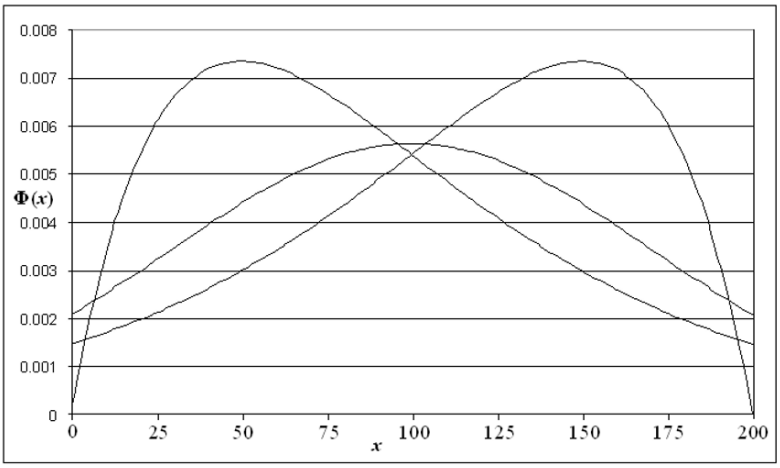

Figure 2.3.1 $\Phi(x)$ versus x for 3 different distributions: each with μ equal to 100 and σ equal to $50\sqrt{2}$.

The distribution to the left of Figure 2.3.1 exhibits positive skewness while the distribution to the right exhibits negative skewness. The symmetric distribution in the middle has zero skewness.

We can use Equation 2.2.3 to compute the variance and skewness of the distribution characterized by Equation 2.2.2. We have already shown that the mean of this distribution is $2c$. Substituting this value of the mean and 2.2.2 into 2.1.6 we get the following expression for the variance:

$$\sigma^2 = \frac{1}{c^2}\int_0^\infty (x-2c)^2 x e^{-x/c} dx = \int_0^\infty (\frac{x^3}{c^2} - \frac{4x^2}{c} + 4x) e^{-x/c} dx \qquad (2.3.4)$$

Using Equation 2.2.3 we get:

$$\sigma^2 = \frac{6c^4}{c^2} - \frac{8c^3}{c} + 4c^2 = 2c^2 \quad \text{and therefore} \quad \sigma = \sqrt{2}c$$

Similarly we can show that the ***skewness*** is $4c^3/\sigma^3=\sqrt{2}$. In Figure 2.3.1, the left distribution was generated using Equation 2.2.2 and c = 50. The distribution on the right is the mirror image of the distribution on the left. All three distributions have means of 100 and σ's of 70.71 (i.e., $50\sqrt{2}$). The mode for the distribution on the left is 50 and the mode for the distribution on the right is 150. The mode for the symmetric distribution is 100.

Another parameter called ***kurtosis*** is defined in a manner similar to skewness but is based upon the 4th moment of the distribution about the mean:

$$kurtosis = \frac{1}{\sigma^4}\sum_{xmin}^{xmax} (x-\mu)^4 \Phi(x) \qquad \text{(discrete distribution)}$$

$$kurtosis = \frac{1}{\sigma^4}\int_{xmin}^{xmax} (x-\mu)^4 \Phi(x) dx \qquad \text{(continuous dist.)} \quad (2.3.5)$$

We can again use Equation 2.2.3 to compute the kurtosis of the distribution characterized by Equation 2.2.2 and derive a value equal to $24c^4/\sigma^4=6$.

We define the n^{th} central moment of a distribution as:

$$\mu_n = \int_{xmin}^{xmax} (x-\mu)^n dx \qquad (2.3.6)$$

By definition, for any distribution, μ_1 is 0 and μ_2 is σ^2. For normal distributions the following recursion relationship can be shown:

$$\mu_n = (n-1)\sigma^2 \mu_{n-2} \tag{2.3.7}$$

We see that for normal distributions $\mu_3 = 2\sigma^2\mu_1 = 0$ and $\mu_4 = 3\sigma^2\mu_2 = 3\sigma^4$. Thus for normal distributions ***skewness*** = 0 and ***kurtosis*** =3. The value of ***kurtosis*** = 6 for the distribution characterized by Equation 2.2.2 is greater than for a normal distribution with the same value of variance. Comparing a distribution with a normal distribution of the same variance (i.e., σ^2), kurtosis measures whether the distribution is tall and skinny (less than 3) or short and fat (greater than 3).

In finance, the term **kurtosis risk** is sometimes used to warn against assuming normality when in fact a distribution exhibits significant kurtosis. Mandelbrot and Hudson suggest that the failure to include kurtosis in their pricing model caused the famous Long-Term Capital Management collapse in 1998 [MA04].

To estimate skewness and kurtosis from a set of data we use the following equations:

$$skewness = \frac{1}{s_x{}^3 (n-1)} \sum_{i=1}^{n} (x_i - x_{avg})^3 \tag{2.3.8}$$

$$kurtosis = \frac{1}{s_x{}^4 (n-1)} \sum_{i=1}^{n} (x_i - x_{avg})^4 \tag{2.3.9}$$

2.4 Statistical Distributions

In nature most quantities that are observed are subject to a statistical distribution. The distribution is often inherent in the quantity being observed but might also be the result of errors introduced in the method of observation. In the previous section, we considered the testing of people for a genetic profile. The number of people testing positive for a given sample size is a discrete number and if the experiment is repeated many times the

number will vary from sample to sample. In other words, the number is subject to a distribution. Statistics helps us make statements regarding the distributions that we encounter while analyzing experimental data.

The distribution of the number of people testing positive for a given genetic profile is an example of an experimental variable with an inherent distribution. Some experimental variables are subject to errors due to the measuring process. Thus if the measurement were repeated a number of times, the results would be distributed about some mean value. As an example of a distribution caused by a measuring instrument, consider the measurement of temperature using a thermometer. Uncertainty can be introduced in several ways:

1) The persons observing the result of the thermometer can introduce uncertainty. If, for example, a nurse observes a temperature of a patient as 37.4°C, a second nurse might record the same measurement as 37.5°C. (Modern thermometers with digital outputs can eliminate this source of uncertainty.)

2) If two measurements are made but the time taken to allow the temperature to reach equilibrium is different, the results might be different. (Modern digital thermometers eliminate this problem by emitting a beep when the temperature has reached its steady state.)

3) If two different thermometers are used, the instruments themselves might be the source of a difference in the results. This source of uncertainty is inherent in the quality of the thermometers. Clearly, the greater the accuracy, the higher the quality of the instrument and usually, the greater is the cost. It is far more expensive to measure a temperature to 0.001°C than 0.1°C!

4) The temperature might be slightly time dependent.

In this section some of the most important distributions are discussed. These distributions cover most of the distributions that are encountered in work of an experimental nature.

The normal distribution

When x is a continuous variable the normal distribution is often applicable. The normal distribution assumes that the range of x is from $-\infty$ to ∞ and

that the distribution is symmetric about the mean value μ. These assumptions are often reasonable even for distributions of discrete variables, and thus the normal distribution can be used for some distributions of discrete variables as well as continuous variables. The equation for a normal distribution is:

$$\Phi(x) = \frac{1}{\sigma(2\pi)^{1/2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2.4.1}$$

Equation 2.4.1 satisfies the condition expressed by Equation 2.1.2 for the range -∞ to ∞. The normal distribution is shown in Figure 2.4.1 for various values of the standard deviation σ. We often use the term **standard normal distribution** to characterize one particular distribution: a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$. The symbol u is usually used to denote this distribution. Any normal distribution can be transformed into a standard normal distribution by subtracting μ from the values of x and then dividing this difference by σ.

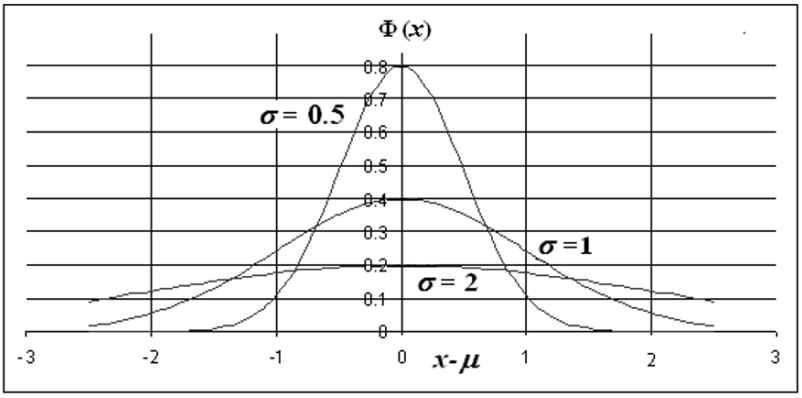

Figure 2.4.1 $\Phi(x)$ vs $x-\mu$ for Normal Distribution (σ=0.5, 1 and 2).

We can define the *effective range* of the distribution as the range in which a specified percentage of the data can be expected to fall. If we specify the effective range of the distribution as the range between $\mu \pm \sigma$, then as n becomes large approximately 68.3% of all measurements would fall within this range. Extending the range to $\mu \pm 2\sigma$, 95.4% would fall within this range and 99.7% would fall within the range $\mu \pm 3\sigma$. The true range of any normal distribution is always -∞ to ∞. Values of the percentage that

fall within the range 0 to u (i.e., $(x-\mu)/\sigma$) are included in tables in many sources [e.g., AB64, FR92]. The standard normal table is also available online [ST03]. Approximate equations corresponding to a given value of probability are also available (e.g., AB64).

The normal distribution is not applicable for all distributions of continuous variables. In particular, if the variable x can only assume positive values and if the mean of the distribution μ is close to zero, then the normal distribution might lead to erroneous conclusions. If however, the value of μ is large (i.e., $\mu/\sigma >> 1$) then the normal distribution is usually a good approximation even if negative values of x are impossible.

We are often interested in understanding how the mean of a sample of n values of x (i.e., x_{avg}) is distributed. It can be shown that the standard deviation of the value of x_{avg} has a standard deviation of $\sigma / \sqrt{n}$. Thus the quantity $(x_{avg}-\mu) / (\sigma/\sqrt{n})$ follows the standard normal distribution u. This result is extremely important and is one of the fundamental concepts of statistics. Stated in words, the standard deviation of the average value of x (i.e., σ_{xavg}) is equal to the value of σ of the distribution divided by $\sqrt{n}$. Thus as n (the number of measurements of x) increases, the value of σ_{xavg} decreases and for large n it approaches zero. For example, let us consider a population with a mean value of 50 and a standard deviation of 10. Thus for large n we would expect about 68% of the x values to fall within the range 40 to 60. If we take a sample of n = 100 observations and then compute the mean of this sample, we would expect that this mean would fall in the range 49 to 51 with a probability of about 68%. In other words, even though the population σ is 10, the standard deviation of an average of 100 observations is only $10/\sqrt{100}$ = 1.

The binomial distribution

When x is a discrete variable of values 0 to n (where n is a relatively small number), the binomial distribution is usually applicable. The variable x is used to characterize the number of **successes** in n trials where p is the probability of a single success for a single trial. The symbol $\Phi(x)$ is thus the probability of obtaining exactly x successes. The number of successes can theoretically range from 0 to n. The equation for this distribution is:

$$\Phi(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \tag{2.4.2}$$

As an example, consider the following problem: what is the probability of drawing at least one ace from a deck of cards if the total number of trials is 3. A deck of 52 cards includes a total of 4 aces so the probability of drawing an ace in any trial is 4/52 = 1/13. After each trial the card drawn is reinserted into the deck and the deck is randomly shuffled. Since the card is returned to the deck the value of p remains unchanged. For this problem the possible values of x are 0, 1, 2 and 3. The value of p (i.e., drawing an ace) is 1/13 and the probability of not drawing an ace in a single trial is 12/13. The probability of drawing no aces in 3 trials is:

$$\Phi(0) = \frac{3!}{0!(3)!} p^0 (1-p)^3 = (12/13)^3 = 0.78652$$

The probability of drawing an ace once is:

$$\Phi(1) = \frac{3!}{1!(2)!} p^1 (1-p)^2 = \frac{6}{2}(1/13)^1 (12/13)^2 = 0.19663$$

The probability of drawing two aces is:

$$\Phi(2) = \frac{3!}{2!(1)!} p^2 (1-p)^1 = \frac{6}{2}(1/13)^2 (12/13)^1 = 0.01639$$

The probability of drawing an ace all three times is:

$$\Phi(3) = \frac{3!}{3!(0)!} p^3 (1-p)^0 = (1/13)^3 = 0.00046$$

The sum of all 4 of these probable outcomes is one. The probability of drawing an ace at least once is 1 – 0.78652 = 0.21348.

The mean value μ and standard deviation σ of the binomial distribution can be computed from the values of n and p:

$$\mu = np \tag{2.4.3}$$

$$\sigma = (np(1-p))^{1/2} \tag{2.4.4}$$

Equation 2.4.3 is quite obvious. If, for example, we flip a coin 100 times, what is the average value of the number of heads we would observe? For this problem, $p = ½$, so we would expect to see on average 100 * 1/2 = 50 heads. The equation for the standard deviation is not obvious, however the proof of this equation can be found in many elementary textbooks on statistics. For this example we compute σ as $(100*1/2*1/2)^{1/2} = 5$. If μ and n are large compared to 1 we can use the normal approximation to compute the probability of seeing exactly x heads. (For large n the number of significant digits required to use Equation 2.4.2 exceeds the calculational capabilities of most computer programming languages.) The area under the normal approximation for x heads would be the area in the range from x - ½ to x + ½. Converting this to the standard normal distribution we use the range $(x-\mu-½)/\sigma$ to $(x-\mu+½)/\sigma$. For example, the probability of seeing exactly 50 heads would correspond to the area under the standard normal distribution from -0.5/5 = -0.1 to 0.5/5 = 0.1. Using the standard normal table, the probability of falling between 0.0 and 0.1 is 0.0398 so the probability of seeing 50 heads is approximately 2 * 0.0398 = 0.0796. The probability of falling in the range 45 to 55 (i.e., 44.5 to 55.5) corresponds to the standard normal range of -1.1 to 1.1 and equals 0.7286. We could have tried to compute the exact probability of seeing 50 heads by using Equation 2.4.2 directly:

$$\Phi(50) = \frac{100!}{50!(50)!}0.5^{50}(1-0.5)^{50} = \frac{100!}{(50!)^2}0.5^{100}$$

This seemingly simply equation is a calculational nightmare!

The Poisson distribution

The binomial distribution (i.e., Equation 2.4.2) becomes unwieldy for large values of n. The Poisson distribution is used for a discrete variable x that can vary from 0 to ∞. If we assume that we know the mean value μ of the distribution, then $\Phi(x)$ is computed as:

$$\Phi(x) = \frac{e^{-\mu}\mu^x}{x!} \tag{2.4.5}$$

It can be shown that the standard deviation σ of the Poisson distribution is:

$$\sigma = \mu^{1/2} \tag{2.4.6}$$

If μ is a relatively small number and n is large, then the Poisson distribution is an excellent approximation of the binomial distribution. If μ is a large value, the normal distribution is an excellent approximation of a Poisson distribution.

As an example of a Poisson distribution, consider the observation of a rare genetic problem. Let us assume that the problem is observed on average 2.3 times per 1000 people. For practical purposes n is close to ∞ so we can assume that the Poisson distribution is applicable. We can compute the probability of observing x people with the genetic problem out of a sample population of 1000 people. The probability of observing no one with the problem is:

$$\Phi(0) = e^{-2.3}\, 2.3^0 / 0! = e^{-2.3} = 0.1003$$

The probability of observing one person with the problem is:

$$\Phi(1) = e^{-2.3}\, 2.3^1 / 1! = 2.3e^{-2.3} = 0.2306$$

The probability of observing two people with the problem is:

$$\Phi(2) = e^{-2.3}\, 2.3^2 / 2! = 2.3^2 e^{-2.3} / 2 = 0.2652$$

The probability of observing three people with the problem is:

$$\Phi(3) = e^{-2.3}\, 2.3^3 / 3! = 2.3^3 e^{-2.3} / 6 = 0.2136$$

From this point on, the probability $\Phi(x)$ decreases more and more rapidly and for all intents and purposes approaches zero for large values of x. (The probability of observing 10 or more is only about 0.00015.)

Another application of Poisson statistics is for counting experiments in which the number of counts is large. For example, consider observation of a radioisotope by an instrument that counts the number of signals emanating from the radioactive source per unit of time. Let us say that 10000

counts are observed. Our first assumption is that 10000 is our best estimate of the mean μ of the distribution. From equation 2.4.6 we can then estimate the standard deviation σ of the distribution as $10000^{1/2} = 100$. In other words, in a counting experiment in which 10000 counts are observed, the accuracy (i.e., the estimated standard deviation) of this observed count rate is approximately 1% (i.e., 100/10000 = 0.01). To achieve an accuracy of 0.5% we can compute the required number of counts:

$$0.005 = \sigma/\mu = \mu^{1/2} / \mu = \mu^{-1/2}$$

Solving this equation we get a value of μ = 40000. In other words to double our accuracy (i.e., halve the value of σ) we must increase the observed number of counts by a factor of 4.

The χ^2 distribution

The χ^2 (**chi-squared**) distribution is defined using a variable ***u*** that is normally distributed with a mean of 0 and a standard deviation of 1. This ***u*** distribution is called the standard normal distribution. The variable $\chi^2(k)$ is called the χ^2 value with ***k*** degrees of freedom and is defined as follows:

$$\chi^2(k) = \sum_{i=1}^{i=k} u_i^2 \tag{2.4.7}$$

In other words, if ***k*** samples are extracted from a standard normal distribution, the value of $\chi^2(k)$ is the sum of the squares of the ***u*** values. The distribution of these values of $\chi^2(k)$ is a complicated function:

$$\Phi(\chi^2(k)) = \frac{(\chi^2)^{k/2-1}}{2^{k/2}\Gamma(k/2)} \exp(-\chi^2/2) \tag{2.4.8}$$

In this equation Γ is called the gamma function and is defined as follows:

$$\Gamma(k/2) = (k/2-1)(k/2-2)...3*2*1 \quad \textit{for k even}$$
$$\Gamma(k/2) = (k/2-1)(k/2-2)...3/2*1/2*\pi^{1/2} \quad \textit{for k odd} \tag{2.4.9}$$

Equation 2.4.8 is complicated and rarely used. Of much greater interest is determination of a range of values from this distribution. What we are more interested in knowing is the probability of observing a value of χ^2 from 0 to some specified value. This probability can be computed from the following equation [AB64]:

$$P(\chi^2/k) = \frac{1}{2^{k/2}\Gamma(k/2)} \int_0^{\chi^2} t^{k/2-1} e^{-t/2} dt \qquad (2.4.10)$$

For small values of k (typically up to k=30) values of χ^2 are presented in a tabular format [e.g., AB64, FR92, ST03] but for larger values of k, approximate values can be computed (using the normal distribution approximation described below). The tables are usually presented in an inverse format (i.e., for a given value of k, the values of χ^2 corresponding to various probability levels are tabulated).

i	*x*	*u* = (*x*-108)/2	u^2
1	110.0	1.00	1.0000
2	105.5	1.25	1.5625
3	108.6	0.30	0.0900
4	103.8	-2.10	4.4100
5	105.2	-1.40	1.9600
6	113.7	2.85	8.1225
7	106.9	-1.05	1.1025
8	109.7	0.85	0.7225
sum	**862.4**	**-0.80**	**18.9700**
avg	**107.8**	**-0.10**	**2.3712**

Table 2.4.1 Measured values of x after a change in the process. The historical values of μ and σ are 108 and 2.

As an example of the use of this distribution, consider the data in Table 2.4.1. This data is from an experiment in which we are testing a process to check if something has changed. Some variable x characterizes the process. We know from experience that the historical mean of the distribution of x is μ=108 and the standard deviation is σ=2. The experiment consists of measuring 8 values of x. The computed average for the 8 values of x is 107.8 which is close to the historical value of μ but can we make a statement regarding the variance in the data? If there was no

change in the process, we would expect that the following variable would be distributed as a standard normal distribution (μ=0, σ=1):

$$u = \frac{(x - \mu)}{\sigma} = \frac{(x - 108)}{2} \tag{2.4.11}$$

Using Equation 2.4.11, 2.4.7 and the 8 values of x we can compute a value for χ^2. From Table 2.4.1 we see that the value obtained is 18.97. The question that we would like to answer is what is the probability of obtaining this value or a greater value by chance? From [ST03] it can be seen that for k = 8, there is a probability of 2.5% that the value of χ^2 will exceed 17.53 and 1% that the value will exceed 20.09. If there was no change in the process, this test indicates that the probability of obtaining a value of 18.97 or greater is between 1% and 2.5%. This is a fairly low probability so we might suspect that there has been a significant change in the process.

An important use for the χ^2 distribution is analysis of variance. The **variance** is defined as the standard deviation squared. We can get an unbiased estimate of the variance of a variable x by using n observations of the variable. Calling this unbiased estimate as s^2, we compute it as follows:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x - x_{avg})^2 \tag{2.4.12}$$

We can use this equation to estimate the variance in the data in Table 2.4.1 and this value can then be tested using the χ^2 distribution. For this data the value of s is 3.29. The quantity $(n\text{-}1)s^2/\sigma^2$ is distributed as χ^2 with n-1 degrees of freedom. We can use the χ^2 distribution to test if this value is significantly different than the historic value of 2. The quantity $(n\text{-}1)s^2/\sigma^2$ is equal to $7*3.29^2/2^2 = 18.94$. If there was no change in the process, for n-1=7 degrees of freedom, the probability of exceeding 18.47 is 1%. The probability of obtaining a value of 18.94 or greater is thus less than 1% which leads us to suspect that there might have been a change in the process. Notice that this test yields an even lower probability than the usage of 2.4.11 (i.e., between 1% and 2.5%). The reason for this increased estimate of the probability of a change in the process is due to the fact that we have included additional information in the analysis (i.e., the estimated value of the x_{avg} which is included in Equation 2.4.12).

Two very useful properties of the χ^2 distribution are the mean and standard deviation of the distribution. For k degrees of freedom, the mean is k and the standard deviation is $\sqrt{2k}$. For large values of k, we can use the fact that this distribution approaches a normal distribution and thus we can easily compute ranges. For example, if k = 100, what is the value of χ^2 for which only 1% of all samples would exceed it by chance? For a standard normal distribution, the 1% limit is 2.326. The value for the χ^2 distribution would thus be $\mu + 2.326*\sigma = k + 2.326*(2k)^{1/2} = 100 + 31.2 = 131.2$.

The *t* distribution

The t distribution (sometimes called the student-t distribution) is used for samples in which the standard deviation is not known. Using n observations of a variable x, the mean value x_{avg} and the unbiased estimate s of the standard deviation can be computed. The variable t is defined as:

$$t = (x_{avg} - \mu)/(s/\sqrt{n}) \tag{2.4.13}$$

The t distribution was derived to explain how this quantity is distributed. In our discussion of the normal distribution, it was noted that the quantity $(x_{avg}-\mu) / (\sigma/\sqrt{n})$ follows the standard normal distribution u. When σ of the distribution is not known, the best that we can do is use s instead. For large values of n the value of s approaches the true value of σ of the distribution and thus t approaches a standard normal distribution. The mathematical form for the t distribution is based upon the observation that Equation 2.4.13 can be rewritten as:

$$t = \frac{x_{avg} - \mu}{\sigma/\sqrt{n}}(\sigma/s) = u\left(\frac{n-1}{\chi^2_{n-1}}\right)^{1/2} \tag{2.4.14}$$

The term σ/s is distributed as $((n\text{-}1) / \chi^2)^{1/2}$ where χ^2 has n-1 degrees of freedom. Thus the mathematical form of the t distribution is derived from the product of u (the standard normal distribution) and $((n\text{-}1)/ \chi^2_{n\text{-}1})^{1/2}$. Values of t for various percentage levels for n-1 up to 30 are included in

tables in many sources [e.g., AB64, FR92]. The ***t*** table is also available online [ST03]. For values of $n > 30$, the ***t*** distribution is very close to the standard normal distribution.

For small values of ***n*** the use of the ***t*** distribution instead of the standard normal distribution is necessary to get realistic estimates of ranges. For example, consider the case of 4 observations of ***x*** in which x_{avg} and ***s*** of the measurements are 50 and 10. The value of $s/\sqrt{n}$ is 5. The value of ***t*** for $n - 1 = 3$ degrees of freedom and 1% is 4.541 (i.e., the probability of exceeding 4.541 is 1%). We can use these numbers to determine a range for the true (but unknown value) of μ:

$$\mathbf{27.30 = 50 - 4.541*5 <= \mu <= 50 + 4.541*5 = 77.72}$$

In other words, the probability of μ being below 27.30 is 1%, above 77.71 is 1% and within this range is 98%. Note that the value of 4.541 is considerably larger than the equivalent value of 2.326 for the standard normal distribution. It should be noted, however, that the ***t*** distribution approaches the standard normal rather rapidly. For example, the 1% limit is 2.764 for 10 degrees of freedom and 2.485 for 25 degrees of freedom. These values are only 19% and 7% above the standard normal 1% limit of 2.326.

The *F* distribution

The ***F*** distribution plays an important role in data analysis. This distribution was named to honor R.A. Fisher, one of the great statisticians of the early 20th century. The ***F*** distribution is defined as the ratio of two χ^2 distributions divided by their degrees of freedom:

$$F = \frac{\chi^2(k_1)/k_1}{\chi^2(k_2)/k_2} \tag{2.4.15}$$

The resulting distribution is complicated but tables of values of ***F*** for various percentage levels and degrees of freedom are available in many sources (e.g., [AB64, FR92, ST03]). Simple equations for the mean and standard deviation of the ***F*** distribution are as follows:

$$\mu = \frac{k_2}{k_2 - 2} \qquad \textbf{for } k_2 > 2 \tag{2.4.16}$$

$$\sigma^2 = \frac{2k_2^2(k_2 + k_1 - 2)}{k_1(k_2 - 2)^2(k_2 - 4)} \quad \textbf{for } k_2 > 4 \tag{2.4.17}$$

From these equations we see that for large values of k_2 the mean μ approaches 1 and σ^2 approaches $2(1/k_1 + 1/k_2)$. If k_1 is also large, we see that σ^2 approaches zero. Thus if both k_1 and k_2 are large, we would expect the value of F to be very close to one.

The Gamma distributions

The gamma distributions are characterized by three constants: α, β and γ. The values of α and β must be positive. The range of values of the variable x must be greater than γ. The general form of all gamma distributions is:

$$\Phi(x) = \frac{(x - \gamma)^{\alpha - 1}}{\beta^{\alpha}\Gamma(\alpha)} e^{-(x-\gamma)/\beta} \tag{2.4.18}$$

Equation 2.2.2 is an example of a gamma distribution with α=2, β=c and γ=0. (From 2.4.9 we see that for even integers, $\Gamma(n)$ is $(n\text{-}1)!$ so $\Gamma(2)$ is one.) Gamma distributions for α=1, 1.5, 2, 2.5, 4 and for β=1 and γ=0 are shown in Figure 2.4.2. The χ^2 distribution with k degrees of freedom is a special case of the Gamma distribution with α=k/2, β=2 and γ=0.

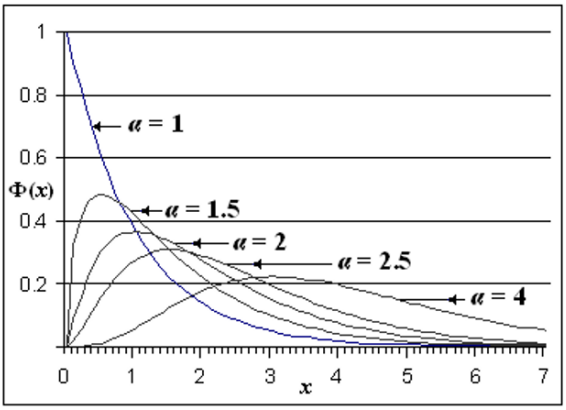

Figure 2.4.2 $\Phi(x)$ vs x for Gamma Distributions ($\alpha = 1, 1.5, 2, 2.5 \& 4,\ \beta = 1, \gamma = 0$).

2.5 Functions of Several Variables

In many experiments, data is often combined to create the independent and/or the dependent variables. For example, in Section 1.4 experiments to measure half lives of radioisotopes were discussed. In Equation 1.4.1 the background term (i.e., a_3) was included as an unknown but in Equations 1.4.6 and 1.4.7 the dependent variables $counts_1$ and $counts_2$ were assumed to be corrected for their background count rates. Thus the variables $counts_1$ and $counts_2$ are computed as the differences between the measured rates and the background rates.

In general we can consider the problem as v (the variable) being a function computed using r measured or computed variables u_1, u_2 .. u_r :

$$v = f(u_1, u_2 ... u_r) \tag{2.5.1}$$

Defining Δv and Δu_i as the variations from the true values of v and u_i, if these variations are small we obtain:

$$\Delta v = \frac{\partial f}{\partial u_1}\Delta u_1 + \frac{\partial f}{\partial u_2}\Delta u_2 + ... \frac{\partial f}{\partial u_r}\Delta u_r \tag{2.5.2}$$

Squaring Equation 2.5.2:

$$(\Delta v)^2 = \left(\frac{\partial f}{\partial u_1}\Delta u_1\right)^2 + \left(\frac{\partial f}{\partial u_2}\Delta u_2\right)^2 + ... \left(\frac{\partial f}{\partial u_r}\Delta u_r\right)^2 +$$

$$\left(2\frac{\partial f}{\partial u_1}\frac{\partial f}{\partial u_2}\Delta u_1 \Delta u_2\right) + ... \left(2\frac{\partial f}{\partial u_{r-1}}\frac{\partial f}{\partial u_r}\Delta u_{r-1}\Delta u_r\right) \qquad (2.5.3)$$

If we repeat the measurements n times, we can then say that the average value of $(\Delta v)^2$ is the sum of the average values of all the terms in Equation 2.5.3. If the variables u_1, u_2 .. u_r are uncorrelated (i.e., not related) then the average values of the products $\Delta u_i \Delta u_j$ approach zero as n approaches infinity and Equation 2.5.3 approaches:

$$(\Delta v)^2_{avg} = \left(\frac{\partial f}{\partial u_1}\Delta u_1\right)^2_{avg} + \left(\frac{\partial f}{\partial u_2}\Delta u_2\right)^2_{avg} + ... \left(\frac{\partial f}{\partial u_r}\Delta u_r\right)^2_{avg}$$

These average values can be replaced by σ^2 terms for large n:

$$\sigma_v^{\,2} = \sum_{i=1}^{r}\left(\frac{\partial f}{\partial u_i}\sigma_{u_i}\right)^2 \qquad (2.5.4)$$

For large n if the average values of the products $\Delta u_i \Delta u_j$ do not approach zero then the variables are correlated and we must include these terms in the determination of $\sigma_v^{\,2}$:

$$\sigma_v^{\,2} = \sum_{i=1}^{r}\left(\frac{\partial f}{\partial u_i}\sigma_{u_i}\right)^2 + \sum_{i=1}^{r-1}\sum_{j=i+1}^{r} 2\frac{\partial f}{\partial u_i}\frac{\partial f}{\partial u_j}\sigma_{ij} \qquad (2.5.5)$$

The terms σ_{ij} are called the covariances of the variables u_i and u_j.

Example 2.5.1: A very simple but useful example of a function of several variables is the case of v being the sum or difference of two other variables. We often encounter these relationships when analyzing data. Assume we have data obtained for two separate

groups (for example, males and females). We are interested in developing models for the males and females separately and then together. We start with the simple relationship: $v = u_1 + u_2$ where u_1 represents the males and u_2 represents females. Since there is no correlation between the values of u_1 and u_2 we use Equation 2.5.4 to compute the value of σ_v. If, for examples u_1=25±3 and u_2=35±4 then v=60±5:

$$\sigma_v^2 = \left(\frac{\partial f}{\partial u_1}\sigma_{u_1}\right)^2 + \left(\frac{\partial f}{\partial u_2}\sigma_{u_2}\right)^2 = 3^2 + 4^2 = 5^2$$

Note that for this simple relationship the partial derivatives are one.

Example 2.5.2: When the relationship is a difference and not a sum we run into a problem that can sometimes be quite crucial in work of an experimental nature. In Section 1.4 we discussed the problem of measuring the count rate from a radioisotope and then correcting it using the measured background. For this case the relationship is $v = u_1 - u_2$. The measured count rates are independent so once again we can use Equation 2.5.4. Since the u variables are numbers of counts, from Poisson statistics the relevant σ's are the square roots of the u's. For example if u_1=400 and u_2=256 then v=144. We compute σ_v as follows:

$$\sigma_v^2 = \left(\frac{\partial f}{\partial u_1}\sigma_{u_1}\right)^2 + \left(\frac{\partial f}{\partial u_2}\sigma_{u_2}\right)^2 = 20^2 + 16^2 = 25.6^2$$

We started with count rates that were accurate to about 5% (20/400) and 6% (16/256) and ended with a corrected count rate accurate to 17.8% (25.6/144) which is an erosion in accuracy of about a factor of three. Summarizing, for sums and differences:

$$v = u_1 \pm u_2 \pm \ldots u_r \tag{2.5.6}$$

The value of σ_v^2 is computed as follows:

$$\sigma_v^2 = \sigma_{u_1}^2 + \sigma_{u_2}^2 + \ldots \sigma_{u_r}^2 \tag{2.5.7}$$

For products and ratios ($v = u_1 * u_2$ and $v = u_1 / u_2$) Equation 2.5.4 yields the following after dividing by v^2:

$$\left(\frac{\sigma_v}{v}\right)^2 = \left(\frac{\sigma_{u_1}}{u_1}\right)^2 + \left(\frac{\sigma_{u_2}}{u_2}\right)^2 \tag{2.5.8}$$

In general if v is a series of products and ratios:

$$\left(\frac{\sigma_v}{v}\right)^2 = \left(\frac{\sigma_{u_1}}{u_1}\right)^2 + \left(\frac{\sigma_{u_2}}{u_2}\right)^2 + \ldots \left(\frac{\sigma_{u_r}}{u_r}\right)^2 \tag{2.5.9}$$

Another very useful relationship is $v = u^n$. Using Equation 2.5.4 we can derive the following simple relationship:

$$\frac{\sigma_v}{v} = n\frac{\sigma_u}{u} \tag{2.5.10}$$

Example 2.5.3: As an example, consider the measurement of a radioisotope using two different specimens. For the first specimen 1600 counts were recorded and for the second, 900 counts were recorded. The uncertainties computed for these two specimens are:

$$\sigma_{u_1} = (u_1)^{1/2} = 40 \text{ counts}$$
$$\sigma_{u_2} = (u_2)^{1/2} = 30 \text{ counts}$$

and these uncertainties are uncorrelated. The sum, difference, product, ratio and squares and their σ's are tabulated in Table 2.5.1.

Variable	Value	Units	σ	% σ
u_1	1600	counts	40	2.50
u_2	900	counts	30	3.33
$u_1 + u_2$	2500	counts	50	2.00
$u_1 - u_2$	700	counts	50	7.14
$u_1 * u_2$	1,440,000	counts^2	60,000	4.17
u_1 / u_2	1.778	---------	0.0741	4.17
u_1^2	2,560,000	counts^2	128,000	5.00
u_2^2	810,000	counts^2	54,000	6.67

Table 2.5.1 Elementary functions of variables u_1 and u_1

When we can no longer assume that the variables are independent, we must use Equation 2.5.5 to compute σ_v. We see from this equation that the covariance terms are included. The covariance between two variables (lets say u and v) is defined as in a manner similar to Equation 2.1.6:

$$\sigma_{uv} = \int_{umin}^{umax} \int_{vmin}^{vmax} (u - \mu_u)(v - \mu_v)\Phi(u,v)dvdu \qquad (2.5.11)$$

The term $\phi(u,v)$ is called the **joint probability distribution** and is only equal to $\phi(u)\phi(v)$ when u and v are uncorrelated (i.e., the **correlation coefficient** ρ is zero). The parameter ρ is defined as:

$$\rho = \frac{\sigma_{uv}}{\sigma_u \sigma_v} \qquad (2.5.12)$$

To estimate σ_{uv} we use an equation similar to 2.3.2:

$$s_{uv} = \frac{1}{n-1} \sum_{i=1}^{n} (u_i - u_{avg})(v_i - v_{avg}) \qquad (2.5.13)$$

To estimate ρ we use the following:

$$r = \frac{s_{uv}}{s_u s_v} \qquad (2.5.14)$$

Example 2.5.4: The data in Table 2.5.2 include values of u_i and v_i which are paired data measured at the same time. We see that the estimated value of the covariance s_{uv} = -34/7 = -4.857 and the estimated correlation coefficient r = -4.857/(3.808*2.138) = -0.597. Looking at the data we see that as u increases, v tends to decrease.

i	U	v	$u+v$	$u-u_{avg}$	$v-v_{avg}$	$(u-u_{avg})(v-v_{avg})$
1	10	8	18	1.75	-0.50	-0.875
2	6	10	16	-2.25	1.50	-3.375
3	8	10	18	-0.25	1.50	-0.375
4	12	8	20	3.75	-0.50	-1.875
5	3	12	15	-5.25	3.50	-18.375
6	9	5	14	0.75	-3.50	-2.625
7	14	7	21	5.75	-1.50	-8.625
8	4	8	12	-4.25	-0.50	2.125
Sum	**66**	**68**	**134**	**0.00**	**0.00**	**-34.000**
Avg	**8.250**	**8.500**	**16.75**	**0.00**	**0.00**	
Est σ	**3.808**	**2.138**	**3.059**	**3.808**	**2.138**	s_{uv} **= -4.857**

Table 2.5.2 Values of paired data u_i and v_i measured at times i= 1 to 8

Example 2.5.5: We next compute the values of $x_1 = u + v$ and $x_2 = u*v$ using Equation 2.5.5 and the estimated value of covariance σ_{uv} computed in Example 2.5.4. When there are only two variables, Equation 2.5.2 simplifies to:

$$\sigma_x^2 = \left(\frac{\partial f}{\partial u}\sigma_u\right)^2 + \left(\frac{\partial f}{\partial v}\sigma_v\right)^2 + 2\frac{\partial f}{\partial u}\frac{\partial f}{\partial v}\sigma_{uv} \qquad (2.5.15)$$

For x_1 Equation 2.5.15 reduces to:

$$\sigma_{x_1}^2 = (\sigma_u)^2 + (\sigma_v)^2 + 2\sigma_{uv} \qquad (2.5.16)$$

Equation 2.5.16 is not a function of either u or v so one value of σ_{x1} is applicable to all the points. The value computed using this equation is 3.059 which we see in Table 2.5.2 is exactly the same as what we get if we just use the values of $u + v$ directly to compute σ_{x1}.

For $x_2 = u*v$ Equation 2.5.15 reduces to:

$$\sigma_{x_2}^{\ 2} = (v\sigma_u)^2 + (u\sigma_v)^2 + 2uv\sigma_{uv}$$

Dividing this equation by $x_2^2 = (uv)^2$ we get the more useful form:

$$\left(\frac{\sigma_{x_2}}{x_2}\right)^2 = \left(\frac{\sigma_u}{u}\right)^2 + \left(\frac{\sigma_v}{v}\right)^2 + 2\frac{\sigma_{uv}}{uv} \qquad (2.5.17)$$

Equation 2.5.17 is a function of both u and v so there is a different value of σ_{x2} for each point. For example, for point 5:

$$\left(\frac{\sigma_{x_2}}{x_2}\right)^2 = \left(\frac{3.808}{3}\right)^2 + \left(\frac{2.138}{12}\right)^2 - 2\frac{4.857}{36} = 1.373$$

For this point $\sigma_{x2} = x_2 * \sqrt{1.373} = 3*12*1.172=42.19$. Note that for this particular point the value of σ_{x2} is actually larger than x_2 due to the large fractional uncertainty in the value of u.

Chapter 3 THE METHOD OF LEAST SQUARES

3.1 Introduction

The method of least squares was the cause of a famous dispute between two giants of the scientific world of the early 19th century: Adrien Marie Legendre and Carl Friedrich Gauss. The first published treatment of the method of least squares was included in an appendix to Legendre's book ***Nouvelles methods pour la determination des orbites des cometes***. The 9 page appendix was entitled ***Sur la methode des moindres quarres***. The book and appendix was published in 1805 and included only 80 pages but gained a 55 page supplement in 1806 and a second 80 page supplement in 1820 [ST86]. It has been said that the method of least squares was to statistics what calculus had been to mathematics. The method became a standard tool in astronomy and geodesy throughout Europe within a decade of its publication. Gauss in 1809 in his famous ***Theoria Motus*** claimed that he had been using the method since 1795. (Gauss's book was first translated into English in 1857 under the authority of the United States Navy by the Nautical Almanac and Smithsonian Institute [GA57]). Gauss applied the method to the calculation of the orbit of Ceres: a planetoid discovered by Giuseppe Piazzi of Palermo in 1801 [BU81]. (A prediction analysis of Piazzi's discovery is included in Section 6.6.) Another interesting aspect of the method of least squares is that it was rediscovered in a slightly different form by Sir Francis Galton. In 1885 Galton introduced the concept of regression in his work on heredity. But as Stigler says: "Is there more than one way a sum of squared deviations can be made small? Even though the method of least squares was discovered more than 200 years ago, it is still "the most widely used nontrivial technique of modern statistics" [ST86].

The least squares method is discussed in many books but the treatment is usually limited to linear least squares problems. In particular, the emphasis is often on fitting straight lines or polynomials to data. The multiple li-

J. Wolberg, *Designing Quantitative Experiments*, DOI 10.1007/978-3-642-11589-9_3,

near regression problem (described below) is also discussed extensively (e.g., [FR92, WA93]). Treatment of the general nonlinear least squares problem is included in a much smaller number of books. One of the earliest books on this subject was written by W. E. Deming and published in the pre-computer era in 1943 [DE43]. An early paper by R. Moore and R. Zeigler discussing one of the first general purpose computer programs for solving nonlinear least squares problems was published in 1960 [MO60]. The program described in the paper was developed at the Los Alamos Laboratories in New Mexico. Since then general least squares has been covered in varying degrees and with varying emphases by a number of authors (e.g., DR66, WO67, BA74, GA92, VE02, WO06).

For most quantitative experiments, the method of least squares is the "best" analytical technique for extracting information from a set of data. The method is best in the sense that the parameters determined by the least squares analysis are normally distributed about the true parameters with the least possible standard deviations. This statement is based upon the assumption that the uncertainties in the data are uncorrelated and normally distributed. For most quantitative experiments this is usually true or is a reasonable approximation. When the curve being fitted to the data is a straight line, the term **linear regression** is often used. For the more general case in which a plane based upon several independent variables is used instead of a simple straight line, the term **multiple linear regression** is often used. Prior to the advent of the digital computer, curve fitting was usually limited to linear relationships. For the simplest problem (i.e., a straight line), the assumed relationship between the dependent variable $\boldsymbol{y}$ and the independent variable $\boldsymbol{x}$ is:

$$\boldsymbol{y = a_1 + a_2 x} \tag{3.1.1}$$

For the case of more than one independent variable (multiple linear regression), the assumed relationship is:

$$\boldsymbol{y = a_1 x_1 + a_2 x_2 + \ldots + a_m x_m + a_{m+1}} \tag{3.1.2}$$

For this more general case each data point includes $\boldsymbol{m}+1$ values: $\boldsymbol{y_i, x_{1i}, x_{2i}, \ldots, x_{mi}}$.

The least squares solutions for problems in which Equations 3.1.1 and 3.1.2 are valid fall within the much broader class of **linear least squares**

problems. In general, all linear least squares problems are based upon an equation of the following form:

$$y = f(\mathbf{X}) = \sum_{k=1}^{k=p} a_k g_k(\mathbf{X}) = \sum_{k=1}^{k=p} a_k g_k(x_1, x_2 .... x_m) \qquad (3.1.3)$$

In other words, y is a function of $\mathbf{X}$ (a vector with m terms). Any equation in which the p unknown parameters (i.e., the a_k's) are coefficients of functions of only the independent variables (i.e., the m terms of the vector $\mathbf{X}$) can be treated as a linear problem. For example in the following equation, the values of a_1, a_2, and a_3 can be determined using linear least squares:

$$y = a_1 + a_2 \sin(x) + a_3 e^x$$

This equation is nonlinear with respect to x but the equation is linear with respect to the a_k's. In this example, the $\mathbf{X}$ vector contains only one term so we use the notation x rather than x_1. The following example is a linear equation in which $\mathbf{X}$ is a vector containing 2 terms:

$$y = a_1 + a_2 \sin(x_1) + a_3 e^{x_2}$$

The following example is a nonlinear function:

$$y = a_1 + a_2 \sin(a_4 x) + a_3 e^x$$

The fact that a_4 is embedded within the second term makes this function incompatible with Equation 3.1.3 and therefore it is nonlinear with respect to the a_k's.

For both linear and nonlinear least squares, a set of p equations and p unknowns is developed. If Equation 3.1.3 is applicable then this set of equations is linear and can be solved directly. However, for nonlinear equations, the p equations require estimates of the a_k's and therefore iterations are required to achieve a solution. For each iteration, the a_k's are updated, the terms in the p equations are recomputed and the process is continued until some convergence criterion is achieved. Unfortunately, achieving convergence is not a simple matter for some nonlinear problems.

For some problems our only interest is to compute $y = f(x)$ and perhaps some measure of the uncertainty associated with these values (e.g., σ_f) for various values of x. This is what is often called the **prediction** problem. We use measured or computed values of x and y to determine the parameters of the equation (i.e., the a_k's) and then apply the equation to calculate values of y for any value of x. For cases where there are several (let us say m) independent variables, the resulting equation allows us to predict y for any combination of $x_1, x_2, .. x_m$. The least squares formulation developed in this chapter also includes the methodology for prediction problems.

3.2 The Objective Function

The method of least squares is based upon the minimization of an **objective function**. The term "least squares" comes from the fact that the objective function includes a squared term for each data point. The simplest problems are those in which y (a scalar quantity) is related to an independent variable x (or variables x_j's) and it can be assumed that there is no (or negligible) errors in the independent variable (or variables). The objective function for these cases is:

$$S = \sum_{i=1}^{i=n} w_i R_i^2 = \sum_{i=1}^{i=n} w_i (Y_i - y_i)^2 = \sum_{i=1}^{i=n} w_i (Y_i - f(\mathbf{X}_i))^2 \qquad (3.2.1)$$

In this equation n is the number of data points, Y_i is the i^{th} input value of the dependent variable and y_i is the i^{th} computed value of the dependent variable. The variable R_i is called the i^{th} residual and is the difference between the input and computed values of y for the i^{th} data point. The variable $\mathbf{X}_i$ (unitalicized) represents the independent variables and is either a scalar if there is only one independent variable or a vector if there is more than one independent variable. The function f is the equation used to express the relationship between $\mathbf{X}$ and y. The variable w_i is called the "weight" associated with the i^{th} data point and is discussed in the next section.

A schematic diagram of the variables for point i is shown in Figure 3.2.1. In this diagram there is only a single independent variable so the notation x is used instead of $\mathbf{X}$. The variable E_i is the true but unknown error in the i^{th} value of y. Note that neither the value of Y_i nor y_i is exactly equal to the unknown η_i (the true value of y) at this value of x_i. However, a fundamen-

tal assumption of the method of least squares is that if $\boldsymbol{Y_i}$ is determined many times, the average value would approach this true value.

The next level of complexity is when the uncertainties in the measured or calculated values of $\boldsymbol{x}$ are not negligible. The relevant schematic diagram is shown in Figure 3.2.2. For the simple case of a single independent variable, the objective function must also include residuals in the $\boldsymbol{x}$ as well as the $\boldsymbol{y}$ direction:

$$\begin{aligned} \boldsymbol{S} &= \sum_{\boldsymbol{i}=1}^{\boldsymbol{i}=n} \left(\boldsymbol{w}_{y_i} \boldsymbol{R}_{y_i}^2 + \boldsymbol{w}_{x_i} \boldsymbol{R}_{x_i}^2 \right) \\ &= \sum_{\boldsymbol{i}=1}^{\boldsymbol{i}=n} \left(\boldsymbol{w}_{y_i} \left(\boldsymbol{Y}_i - \boldsymbol{y}_i \right)^2 + \boldsymbol{w}_{x_i} \left(\boldsymbol{X}_i - \boldsymbol{x}_i \right)^2 \right) \end{aligned} \tag{3.2.2}$$

In this equation, $\boldsymbol{X_i}$ (italicized) is the measured value of the $\boldsymbol{i}^{\mathbf{th}}$ independent variable and $\boldsymbol{x_i}$ is the computed value. Note that $\boldsymbol{X_i}$ is not the same as $\mathbf{X}_i$ in Equation 3.2.1. In that equation capital $\mathbf{X}$ (unitalicized) represents the vector of independent variables.

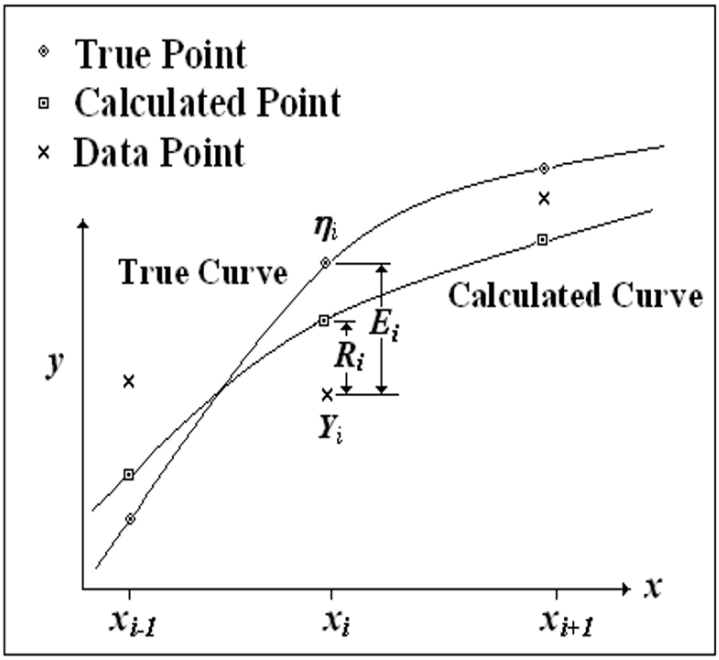

Figure 3.2.1 The True, Calculated and Measured Data Points with no Uncertainties in *x*.

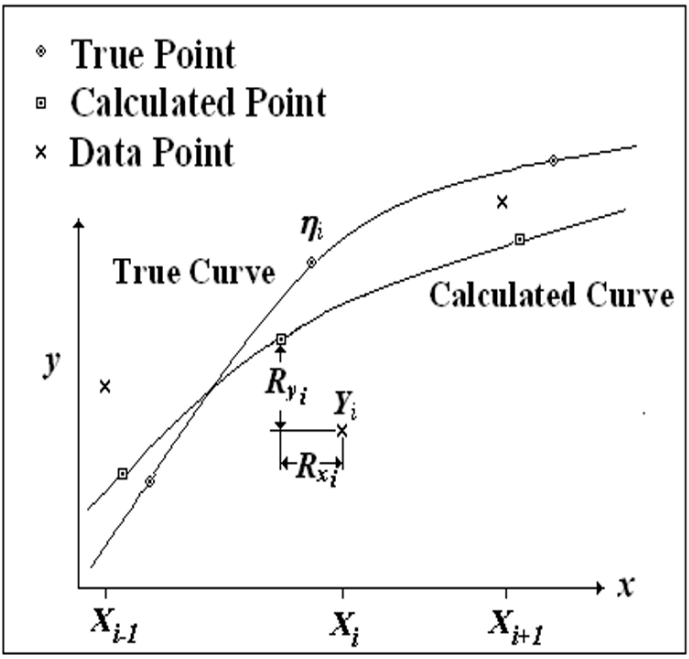

Figure 3.2.2 The True, Calculated and Measured Data Points with Uncertainties in *X*.

It can be shown [WO67] that we can create a modified form of the weights so that the objective function reduces to the following simple form:

$$\begin{aligned} S &= \sum_{i=1}^{i=n} \left(w_{y_i} R_{y_i}^2 + w_{x_i} R_{x_i}^2 \right) \\ &= \sum_{i=1}^{i=n} w_i \left(Y_i - y_i \right)^2 = \sum_{i=1}^{i=n} w_i \left(Y_i - f\left(X_i \right) \right)^2 \end{aligned} \tag{3.2.3}$$

In other words, Equation 3.2.1 is valid even if the uncertainties in the $\boldsymbol{x}$ variables are not negligible. All that is required is a modified form of the weighting function used to determine the values of w_i. Note that if there is more than one independent variable, an additional summation is required:

$$S = \sum_{i=1}^{i=n} \left(w_{y_i} R_{y_i}^2 + \sum_{j=1}^{j=m} w_{x_i} R_{x_{ji}}^2 \right) = \sum_{i=1}^{i=n} w_i \left(Y_i - f\left(\mathbf{X}_i \right) \right)^2 \tag{3.2.4}$$

Note that in Eq 3.2.4 $\mathbf{X}_i$ is unitalicized because it represents the vector of the independent variables. The italicized X_i used in Eq 3.2.3 represents measured value of the scalar independent variable. If y is a vector quantity, then we must further modify the objective function by including a sum over all the y terms. Assuming that there are d terms in the y vector (i.e., y_i is a d dimensional vector), the objective function is:

$$S = \sum_{l=1}^{l=d} S_l = \sum_{l=1}^{d} \sum_{i=1}^{n} w_{l_i} (Y_{l_i} - y_{l_i})^2 = \sum_{l=1}^{d} \sum_{i=1}^{i=n} w_{l_i} (Y_{l_i} - f_l(\mathbf{X}_i))^2 \quad (3.2.5)$$

In a later section we discuss treatment of prior estimates of the unknown a_k parameters. To take these prior estimates into consideration we merely make an additional modification of the objective function. For example, assume that for each of the p unknown a_k's there is a prior estimate of the value. Let us use the notation b_k as the prior estimate of a_k and σ_{b_k} as the uncertainty associated with this prior estimate. In the statistical literature these prior estimates are sometimes called ***Bayesian*** estimators. (This terminology stems from the work of the Reverend Thomas Bayes who was a little known statistician born in 1701. Some of his papers eventually reached the Royal Society but made little impact until the great French mathematician Pierre Laplace discovered them.) The modified form of Equation 3.2.1 is:

$$S = \sum_{i=1}^{i=n} \sum_{i=1}^{i=n} w_i (Y_i - f(\mathbf{X}_i))^2 + \sum_{k=1}^{k=p} (a_k - b_k)^2 / \sigma_{b_k}^2 \quad (3.2.6)$$

If there is no Bayesian estimator for a particular a_k the value of σ_{b_k} is set to infinity.

Regardless of the choice of objective function and scheme used to determine the weights w_i, one must then determine the values of the p unknown parameters a_k that minimize S. To accomplish this task, the most common procedure is to differentiate S with respect to all the a_k's and the resulting expressions are set to zero. This yields p equations that can then be solved to determine the p unknown values of the a_k's. A detailed discussion of this process is included in Section 3.4.

An alternative class of methods to find a "best" set of a_k's is to use an intelligent search within a limited range of the unknown parameter space. A number of such stochastic algorithms are discussed in the literature (e.g., TV04). The only problem with this approach is that the solution does not include estimates of the uncertainties associated with the parameters nor with the uncertainties associated with the fitted surface. One solution to this lack of information is to use stochastic algorithms to find the best set of a_k's and then apply the standard least squares method (see Section 3.4) to compute the required uncertainty estimates. Clearly this approach only makes sense when convergence using the standard method is extremely difficult.

3.3 Data Weighting

In Section 3.2, we noted that regardless of the choice of the objective function, a weight w_i is specified for each point. The "weight" associated with a point is based upon the relative uncertainties associated with the different points. Clearly, we must place greater weight upon points that have smaller uncertainties, and less weight upon the points that have greater uncertainties. In other words the weight w_i must be related in some way to the uncertainties σ_{y_i} and $\sigma_{x_{ji}}$.

In this book the emphasis is on design of experiments and thus we must be able to attach estimates of σ_y and σ_x to the data points. The method of **prediction analysis** allows us to estimate the accuracy of the results that should be obtained from a proposed experiment but as a starting point we need to be able to estimate the uncertainties associated with the data points. By assuming unit weights (i.e., equal weight for each data point), the method of least squares can proceed without these estimated uncertainties. However, this assumption does not allow us to predict the uncertainties of the results that will be obtained from the proposed experiment.

The alternative to using w_i's associated with the σ's of the i^{th} data point is to simply use **unit weighting** (i.e., w_i=1) for all points. This is a reasonable choice for w_i if the σ's for all points are approximately the same or if we have no idea regarding the values (actual or even relative) of σ for the different points. However, when the differences in the σ's are significant, then use of unit weighting can lead to poor results. This point is illustrated

in Figure 3.3.1. In this example, we fit a straight line to a set of data. Note that the line obtained when all points are equally weighted is very different than the line obtained when the points are "weighted" properly. Also note how far the unit weighting line is from the first few points.

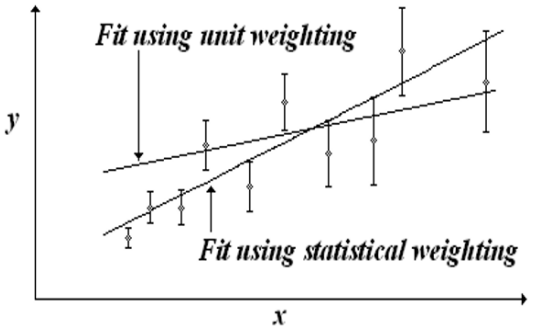

Figure 3.3.1 Two Least Squares Lines using Different Weighting Schemes.

The question that must be answered is how do we relate w_i to the σ's associated with the i^{th} data point? In Section 3.2 we noted that the objective function is of the form:

$$\begin{aligned} S &= \sum_{i=1}^{i=n} S_i = \sum_{i=1}^{i=n} w_i R_i^2 \\ &= \sum_{i=1}^{i=n} w_i (Y_i - y_i)^2 = \sum_{i=1}^{i=n} w_i (Y_i - f(\mathrm{X}_i))^2 \end{aligned} \tag{3.3.1}$$

We will see that the least squares solution is based upon adjusting the unknown values of the a_k's that are included in the function f such that S is minimized. **If the function f is representative of the data**, this minimization process yields values of S_i that tend to be distributed around an average value with some random error ε_i:

$$S_i = w_i R_i^2 = S_{\mathrm{avg}} + \varepsilon_i \tag{3.3.2}$$

For cases in which the uncertainties associated with the values of x_i are negligible our objective should be that the residuals R_i are more or less proportional to the values of σ_{y_i}. (We would like to see the larger resi-

duals associated with points with larger values of σ_{y_i} and visa versa.) If we define the relative error at point i as R_i / σ_{y_i}, our objective should be to have relative errors randomly distributed about 0. To accomplish this, we choose the following weighting function:

$$w_i = 1 / \sigma_{y_i}^2 \tag{3.3.3}$$

We call this type of weighting **statistical weighting** and we will see that it has many attractive properties. Substituting Equation 3.3.3 into 3.3.1 and 3.3.2 we obtain the following expression:

$$(R_i / \sigma_{y_i})^2 = S_i = S_{avg} + \varepsilon_i \tag{3.3.4}$$

We define ***RMS*(*R*)** as the "root mean square" error:

$$RMS(R) = \left(\sum_{i=1}^{i=n} R_i^2 \right)^{1/2} \tag{3.3.5}$$

What we expect is that ***RMS*(*R*)** approaches zero as the noise component of the y values approaches zero. In reality, can we expect this from the least squares analysis? The answer to this question is yes but only if several conditions are met:

1) The function f is representative of the data. In other words, the data falls on the curve described by the function f with only a random "noise" component associated with each data point. For the case where the data is truly represented by the function f (i.e., there is no "noise" component to the data), then all the values of R_i will be zero and thus the values of S_i will be zero.

2) There are no erroneous data points or if data errors do exist, they are not significantly greater than the expected noise component. For example, if a measuring instrument should be accurate to 1%, then errors several times larger than 1% would be suspicious and perhaps problematical. If some points are in error far exceeding 1% then the results of the analysis will probably lead to significant errors in the final results. Clearly, we would hope that there are methods for detect-

ing erroneous data points. This subject is discussed in detail in Section 3.9 in the section dealing with Goodness-of-Fit.

We can explore the effect of **not** using Equation 3.3.3 when the values of σ_y vary significantly from point to point. The data used to generate Figure 3.3.1 are shown in Table 3.3.1. Note the large relative errors for the first few points when unit weighting (i.e., $w_i = 1$) is used.

X	Y	σ_y	$(Y-y)/\sigma_y$ $(w_i=1)$	$(Y-y)/\sigma_y$ $(w_i=1/\sigma_y^2)$
1	6.90	0.05	15.86	0.19
2	11.95	0.10	3.14	-0.39
3	16.800	0.20	-1.82	-0.44
4	22.500	0.50	-0.38	1.23
5	26.200	0.80	-2.53	-0.85
6	33.500	1.50	-0.17	1.08
7	41.000	4.00	0.43	1.03

Table 3.3.1 Fitting Fig 3.3.1 data using $y = a_1 + a_2x$ with 2 different weighting schemes.

The effect of statistical weighting can be further illustrated using a simple computer simulation of a counting experiment. Data were generated using the model $y = a_1 e^{-a_2 x}$ for 10 values of x (0.5, 1.5,…9.5) using a_1=10000 and a_2=0.5. We call this data set the "pure" data and assume that the uncertainties in each data point is $\sigma_y = y^{1/2}$ (i.e., Poisson statistics). The weight for each point is therefore $1/\sigma_y^2 = 1/y$. The experiment proceeds as follows:

1) The value of y_1 is increased by σ (i.e., $y_1^{1/2}$). All other points remain exactly on the generated curve.
2) The method of least squares is then used to analyze the data with two different weighting schemes: w_i=1 and Equation 3.3.3.
3) The experiment is repeated but only y_{10} is increased by σ (i.e., $y_{10}^{1/2}$).

The results are summarized in Table 3.3.2. For all the data sets, using statistical weighting, the weight w_1 (i.e., $1/y_1$) of point 1 is the smallest and the weight w_{10} is the largest. For unit weighting all points are equally weighted. Thus when y_1 is increased by σ_{y1} this has less of an effect upon the results using statistical as compared to unit weighting. When y_{10} is increased by σ_{y10} this has the opposite effect. For the case of a corrupted

value of y_{10} we see that when unit weighting is used, the smallest effect (minimum values of Δa_1 and Δa_2) is observed. This happens because the relative weight attached to this point is less than if statistical weighting is used. Also, since this point is far from the y axis, we would expect a very small effect upon the value of a_1. Note that when the data is "pure" (i.e., all 10 points fall exactly on the exponential curve) both weighting schemes yield the exact values of both a_1 and a_2.

Data	w_i	a_1	σ_{a1}	Δa_1	a_2	σ_{a2}	Δa_2
Pure	Both	10000.0	0.0	0.0	0.5000	0.0000	0.0000
Δy_1	$1/\sigma_y^2$	10085.4	21.8	85.4	0.5021	0.0008	0.0021
Δy_1	1	10121.4	17.6	121.4	0.5046	0.0012	0.0046
Δy_{10}	$1/\sigma_y^2$	9984.6	34.2	-15.4	0.4990	0.0013	-0.0010
Δy_{10}	1	9999.3	5.0	-0.7	0.4999	0.0003	-0.0001

Table 3.3.2 Fitting data using $y = a_1 e^{-a_2 x}$ with 2 different weighting schemes. For 2 cases y_1 is corrupted and for 2 cases y_{10} is corrupted.

In the discussion preceding Equation 3.3.3 it was assumed that the errors in the independent variable (or variables) were negligible. If this assumption cannot be made, then if we assume that the noise component in the data is relatively small, it can be shown [WO67] that the following equation can be used instead of 3.3.3 for the weights w_i:

$$w_i = \frac{1}{\sigma_{y_i}^2 + \sum_{j=1}^{j=m} \left(\frac{\partial f}{\partial x_j} \sigma_{x_{j_i}} \right)^2} \tag{3.3.7}$$

This equation is a more generalized form of **statistical weighting** than Equation 3.3.3. The derivation of this equation is based upon the assumption that higher order terms can be neglected in a Taylor expansion in the region near the minimum value of S. As an example of the application of 3.3.7 to a specific problem, the weighting function for an exponential experiment (Table 3.3.2) would be the following if the σ_x's are included in the analysis:

$$w_i = \frac{1}{\sigma_{y_i}^2 + (-a_1 a_2 e^{-a_2 t} \sigma_{t_i})^2} = \frac{1}{\sigma_{y_i}^2 + a_1^2 a_2^2 e^{-2a_2 t} \sigma_{t_i}^2}$$

3.4 Obtaining the Least Squares Solution

The least squares solution is defined as the point in the "unknown parameter" space at which the objective function $\boldsymbol{S}$ is minimized. Thus, if there are $\boldsymbol{p}$ unknown parameters ($\boldsymbol{a_k}$, $\boldsymbol{k} = \mathbf{1}$ to $\boldsymbol{p}$), the solution yields the values of the $\boldsymbol{a_k}$'s that minimize $\boldsymbol{S}$. When the errors associated with the data points are normally distributed or are reasonably close to normal distributions, the least squares criterion leads to the "best" set of the unknown parameters. The word "best" implies that these $\boldsymbol{p}$ unknown parameters are determined with the smallest possible uncertainty. The most complete proof of this statement that I could find is in a book written by Merriman over 120 years ago [ME84]. To find this minimum point we set the $\boldsymbol{p}$ partial derivatives of $\boldsymbol{S}$ to zero yielding $\boldsymbol{p}$ equations for the $\boldsymbol{p}$ unknown values of $\boldsymbol{a_k}$:

$$\frac{\partial \boldsymbol{S}}{\partial \boldsymbol{a_k}} = \mathbf{0} \qquad \boldsymbol{k} = \mathbf{1} \text{ to } \boldsymbol{p} \tag{3.4.1}$$

In Section 3.2 the following expression (Equation 3.2.3) for the objective function $\boldsymbol{S}$ was developed:

$$\boldsymbol{S} = \sum_{i=1}^{i=n} \boldsymbol{w_i}(\boldsymbol{Y_i} - \boldsymbol{f}(\mathbf{X}_i))^2$$

In this expression the independent variable $\mathbf{X}_i$ can be either a scalar or a vector. The variable $\boldsymbol{Y_i}$ can also be a vector but is usually a scalar. Using this expression and Equation 3.4.1, we get the following $\boldsymbol{p}$ equations:

$$-2\sum_{i=1}^{i=n} \boldsymbol{w_i}(\boldsymbol{Y_i} - \boldsymbol{f}(\mathbf{X}_i))\frac{\partial \boldsymbol{f}(\mathbf{X}_i)}{\partial \boldsymbol{a_k}} = \mathbf{0} \qquad \boldsymbol{k} = \mathbf{1} \text{ to } \boldsymbol{p}$$

$$\sum_{i=1}^{i=n} \boldsymbol{w_i}\boldsymbol{f}(\mathbf{X}_i)\frac{\partial \boldsymbol{f}(\mathbf{X}_i)}{\partial \boldsymbol{a_k}} = \sum_{i=1}^{i=n} \boldsymbol{w_i}\boldsymbol{Y_i}\frac{\partial \boldsymbol{f}(\mathbf{X}_i)}{\partial \boldsymbol{a_k}} \qquad \boldsymbol{k} = \mathbf{1} \text{ to } \boldsymbol{p} \tag{3.4.2}$$

For problems in which the function $\boldsymbol{f}$ is linear, Equation 3.4.2 can be solved directly. In Section 3.1 Equation 3.1.3 was used to specify linear equations:

$$y = f(\mathbf{X}) = \sum_{i=1}^{i=p} a_k g_k(\mathbf{X}) = \sum_{i=1}^{i=p} a_k g_k(x_1, x_2 .... x_m) \tag{3.1.3}$$

The derivatives of f are simply:

$$\frac{\partial f(\mathbf{X})}{\partial a_k} = g_k(\mathbf{X}) \qquad k = 1 \text{ to } p \tag{3.4.3}$$

Substituting 3.4.3 into 3.4.2 we get the following set of equations:

$$\sum_{j=1}^{j=p} a_j \sum_{i=1}^{i=n} w_i g_j(\mathbf{X}_i) g_k(\mathbf{X}_i) = \sum_{i=1}^{i=n} w_i Y_i g_k(\mathbf{X}_i) \quad k = 1 \text{ to } p \tag{3.4.4}$$

Simplifying the notation by using g_k instead of $g_k(X_i)$ we get the following set of equations:

$$a_1 \sum w_i g_1 g_k + a_2 \sum w_i g_2 g_k + ... + a_p \sum w_i g_p g_k = \sum w_i Y_i g_k \qquad k = 1 \text{ to } p \tag{3.4.5}$$

We can rewrite these equations using matrix notation:

$$CA = V \tag{3.4.6}$$

In this equation C is a p by p matrix and A and V are vectors of length p. The terms C_{jk} and V_k are computed as follows:

$$C_{jk} = \sum_{i=1}^{i=n} w_i g_j g_k \tag{3.4.7}$$

$$V_k = \sum_{i=1}^{i=n} w_i Y_i g_k \tag{3.4.8}$$

The terms of the A vector (i.e., the unknown parameters a_k) are computed by solving the matrix equation 3.4.6:

$$\boldsymbol{A} = \boldsymbol{C}^{-1}\boldsymbol{V} \tag{3.4.9}$$

In this equation, $\boldsymbol{C}^{-1}$ is the inverse matrix of $\boldsymbol{C}$. As an example, let us consider problems in which a straight line is fit to the data:

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{a}_1 + \boldsymbol{a}_2\boldsymbol{x} \tag{3.4.10}$$

For this equation $\boldsymbol{g_1} = 1$ and $\boldsymbol{g_2} = \boldsymbol{x}$ so the $\boldsymbol{C}$ matrix is:

$$\boldsymbol{C} = \begin{bmatrix} \sum_{i=1}^{i=n} \boldsymbol{w}_i & \sum_{i=1}^{i=n} \boldsymbol{w}_i\boldsymbol{x}_i \\ \sum_{i=1}^{i=n} \boldsymbol{w}_i\boldsymbol{x}_i & \sum_{i=1}^{i=n} \boldsymbol{w}_i\boldsymbol{x}_i^2 \end{bmatrix} \tag{3.4.11}$$

The $\boldsymbol{V}$ vector is:

$$\boldsymbol{V} = \begin{bmatrix} \sum_{i=1}^{i=n} \boldsymbol{w}_i\boldsymbol{Y}_i \\ \sum_{i=1}^{i=n} \boldsymbol{w}_i\boldsymbol{Y}_i\boldsymbol{x}_i \end{bmatrix} \tag{3.4.12}$$

To apply these equations to a real set of data, let us use the 7 points included in Table 3.3.1 and let us use the case in which all the values of $\boldsymbol{w_i}$ are set to 1 (i.e., unit weighting). For this case, the $\boldsymbol{C}$ and $\boldsymbol{C}^{-1}$ matrices and the $\boldsymbol{V}$ vector are:

$$\boldsymbol{C} = \begin{bmatrix} 7 & 28 \\ 28 & 140 \end{bmatrix} \qquad \boldsymbol{C}^{-1} = \frac{1}{196}\begin{bmatrix} 140 & -28 \\ -28 & 7 \end{bmatrix} \qquad \boldsymbol{V} = \begin{bmatrix} 158.85 \\ 790.20 \end{bmatrix}$$

Solving Equation 3.4.9:

$$\boldsymbol{A} = \boldsymbol{C}^{-1}\boldsymbol{V} = \begin{bmatrix} \boldsymbol{C}_{11}^{-1}\boldsymbol{V}_1 + \boldsymbol{C}_{12}^{-1}\boldsymbol{V}_2 \\ \boldsymbol{C}_{21}^{-1}\boldsymbol{V}_1 + \boldsymbol{C}_{22}^{-1}\boldsymbol{V}_2 \end{bmatrix} = \begin{bmatrix} 0.5786 \\ 5.5286 \end{bmatrix} \tag{3.4.13}$$

For problems in which the $\boldsymbol{f}$ function is nonlinear, the procedure is similar but is iterative. One starts with initial guesses $\boldsymbol{a0_k}$ for the unknown values

of a_k. Simplifying the notation used for Equation 3.4.2, we see that the equations for terms of the C matrix and V vector are:

$$C_{jk} = \sum_{i=1}^{i=n} w_i \frac{\partial f}{\partial a_j} \frac{\partial f}{\partial a_k} \tag{3.4.14}$$

$$V_k = \sum_{i=1}^{i=n} w_i Y_i \frac{\partial f}{\partial a_k} \tag{3.4.15}$$

In the equation for V_k the parameter Y_i is no longer the value of the dependent variable. It is value of the dependent variable minus the computed values using the initial guesses (i.e., the $a0_k$'s). For linear problems we don't need to make this distinction because the initial guesses are zero and thus the computed values are zero. The A vector is determined using Equation 3.4.9 but for nonlinear problems, this vector is no longer the solution vector. It is the vector of computed changes in values of the initial guesses $a0_k$:

$$a_k = a0_k + A_k \qquad k = 1 \text{ to } p \tag{3.4.16}$$

The values of a_k are then used as initial guesses $a0_k$ for the next iteration. This process is continued until a convergence criterion is met or the process does not achieve convergence. Typically the convergence criterion requires that the fractional changes are all less than some specified value of ε:

$$\left| A_k / a0_k \right| \le \varepsilon \qquad k = 1 \text{ to } p \tag{3.4.17}$$

Clearly this convergence criterion must be modified if a value of $a0_k$ is zero or very close to zero. For such terms, one would only test the absolute value of A_k and not the relative value. This method of converging towards a solution is called the Gauss-Newton algorithm and will lead to convergence for many nonlinear problems. It is not, however, the only search algorithm and a number of alternatives to Gauss-Newton are used in least squares software [WO06].

As an example, of a nonlinear problem, let us once again use the data in Table 3.3.1 but choose the following nonlinear exponential function for f:

$$y = f(x) = a_1 e^{a_2 x} \tag{3.4.18}$$

The two derivatives of this function are:

$$f_1' \equiv \frac{\partial f}{\partial a_1} = e^{a_2 x} \qquad \text{and} \qquad f_2' \equiv \frac{\partial f}{\partial a_2} = x a_1 e^{a_2 x}$$

Let us choose as initial guesses $a0_1$=1 and $a0_2$=0.1 and weights w_i=1. (Note that if an initial guess of $a0_1$=0 is chosen, all values of the derivative of f with respect to a_2 will be zero. Thus all the terms of the C matrix except C_{11} will be zero. The C matrix would then be singular and no solution could be obtained for the A vector.) Using Equations 3.4.14 and 3.4.15 and the expressions for the derivatives, we can compute the terms of the C matrix and V vector and then using 3.4.9 we can solve for the values of A_1 and A_2. The computed values are 4.3750 and 2.1706 therefore the initial values for the next iteration are 5.3750 and 2.2706. Using Equation 3.4.17 as the convergence criterion and a value of ε = 0.001, final values of a_1 = 7.7453 and a_2 = 0.2416 are obtained. The value of S obtained using the initial guesses is approximately 1,260,000. The value obtained using the final values of a_1 and a_2 is 17.61. Details of the calculation for the first iteration are included in Tables 3.4.1.and 3.4.2.

X	y	$f=1.0e^{0.1x}$	$Y = y - f$
1	6.900	1.105	5.795
2	11.950	1.221	10.729
3	16.800	1.350	15.450
4	22.500	1.492	21.008
5	26.200	1.649	24.551
6	33.500	1.822	31.678
7	41.000	2.014	38.986

Table 3.4.1 Fitting Data using $f(x) = a_1 e^{a_2 x}$ with initial guesses a_1=1, a_2=0.1

Point	$(f_1')^2$	$(f_2')^2$	$f_1'f_2'$	$f_1'Y$	$f_2'Y$
1	1.221	1.221	1.221	6.404	6.404
2	1.492	5.967	2.984	13.104	26.209
3	1.822	16.399	5.466	20.855	62.565
4	2.226	35.609	8.902	31.340	125.360
5	2.718	67.957	13.591	40.478	202.390
6	3.320	119.524	19.921	57.721	346.325
7	4.055	198.705	28.386	78.508	549.556
Sum	**16.854**	**445.382**	**80.472**	**248.412**	**1318.820**
	C_{11}	C_{22}	C_{12}	V_1	V_2

Table 3.4.2 Computing terms of the *C* matrix and *V* vector

Using the values of the terms of the $\boldsymbol{C}$ matrix and $\boldsymbol{V}$ vector from Table 3.4.2, and solving for the terms of the $\boldsymbol{A}$ vector using Equation 3.4.9, we get values of $a_1 = 4.3750$ and $a_2 = 2.1706$. Using Equation 3.4.16, the values of the initial guesses for the next iteration are therefore 5.3750 and 2.2706. This process is repeated until convergence is obtained. As the initial guesses improve from iteration to iteration, the computed values of the dependent variable (i.e., $\boldsymbol{f}$) become closer to the actual values of the dependent variable (i.e., $\boldsymbol{y}$) and therefore the differences (i.e., $\boldsymbol{Y}$) become closer to zero. From Equation 3.4.15 we see that the values of V_k become smaller as the process progresses towards convergence and thus the terms of the $\boldsymbol{A}$ vector become smaller until the convergence criterion (Equation 3.4.17) is achieved.

A question sometimes asked is: if we increase or decrease the weights how does this affect the results? For example, for unit weighting what happens if we use a value of w other than 1? The answer is that it makes no difference. The values of the terms of the $\boldsymbol{V}$ vector will be proportional to w and all the terms of the $\boldsymbol{C}$ matrix will also be proportional to w. The $\boldsymbol{C}^{-1}$ matrix, however, will be inversely proportional to w and therefore the terms of the $\boldsymbol{A}$ vector (i.e., the product of $\boldsymbol{C}^{-1}\boldsymbol{V}$) will be independent of w. A similar argument can be made for statistical weighting. For example, if all the values of σ_y are increased by a factor of 10, the values of w_i will be decreased by a factor of 100. Thus all the terms of $\boldsymbol{V}$ and $\boldsymbol{C}$ will be decreased by a factor of 100, the terms of $\boldsymbol{C}^{-1}$ will be increased by a factor of 100 and the terms of $\boldsymbol{A}$ will remain unchanged. What makes a difference are the relative values of the weights and not the absolute values. We will see, however, when Goodness-of-Fit is discussed in Section 3.9, that an estimate of the amplitude of the noise component of the data can be very helpful. Fur-

thermore, if prior estimates of the unknown parameters of the model are included in the analysis, then the weights of the data points must be based upon estimates of the absolute values of the weights.

3.5 Uncertainty in the Model Parameters

In Section 3.4 we developed the methodology for finding the set of a_k's that minimize the objective function S. In this section we turn to the task of determining the uncertainties associated with the a_k's. The usual measures of uncertainty are standard deviation (i.e., σ) or variance (i.e., σ^2) so we seek an expression that allows us to estimate the σ_{a_k}'s. It can be shown [WO67, BA74, GA92] that the following expression gives us an unbiased estimate of σ_{a_k}:

$$\sigma_{a_k}^2 = \frac{S}{n-p} C_{kk}^{-1}$$

$$\sigma_{a_k} = \left(\frac{S}{n-p} C_{kk}^{-1}\right)^{1/2} \tag{3.5.1}$$

We see from this equation that the unbiased estimate of σ_{a_k} is related to the objective function S and the k^{th} diagonal term of the inverse matrix C^{-1}. The matrix C^{-1} is required to find the least squares values of the a_k's and once these values have been determined, the final (i.e., minimum) value of S can easily be computed. Thus the process of determining the a_k's leads painlessly to a determination of the σ_{a_k}'s.

As an example, consider the data included in Table 3.3.1. In Section 3.4 details were included for a straight-line fit to the data using unit weighting:

$$y = f(x) = a_1 + a_2 x = 0.5786 + 5.5286x \tag{3.5.2}$$

The C and C^{-1} matrices are:

$$\boldsymbol{C} = \begin{bmatrix} 7 & 28 \\ 28 & 140 \end{bmatrix} \qquad \boldsymbol{C^{-1}} = \frac{1}{196}\begin{bmatrix} 140 & -28 \\ -28 & 7 \end{bmatrix}$$

The value for $\boldsymbol{S}/(\boldsymbol{n}-\boldsymbol{p}) = \boldsymbol{S}/(7-2)$ is 1.6019. We can compute the σ_{a_k}'s from Equation 3.5.1:

$$\sigma_{a_1} = \sqrt{1.6019 * 140/196} = 1.070 \quad \text{and} \quad \sigma_{a_2} = \sqrt{1.6019 * 7/196} = 0.2392$$

The relative error in a_1 is 1.070 / 0.5786 = 1.85 and the relative error in a_2 is 0.2392 / 5.5286 = 0.043. If the purpose of the experiment was to determine, a_2, then we have done fairly well (i.e., we have determined a_2 to about 4%). However, if the purpose of the experiment was to determine, a_1, then we have done terribly (i.e., the relative error is about 185%). What does this large relative error imply? If we were to repeat the experiment many times, we would expect that the computed value of a_1 would fall within the range $0.5786 \pm 2.57 * 1.85 = 0.5786 \pm 4.75$ ninety-five percent of the time (i.e., from -4.17 to 5.33). This is a very large range of probable results. (The constant 2.57 comes from the t distribution with $(n-p) = (7-2) = 5$ degrees of freedom and a 95% confidence interval.)

If we use statistical weighting (i.e., $w_i = 1/\sigma_y^2$), can we improve upon these results? Reanalyzing the data in Table 3.3.1 using the values of σ_y included in the table, we get the following straight line:

$$y = f(x) = a_1 + a_2 x = 1.8926 + 4.9982x \tag{3.5.3}$$

The computed value of σ_{a_1} is 0.0976 and the value for σ_{a_2} is 0.0664. These values are considerably less than the values obtained using unit weighting. The reduction in the value of σ_{a_1} is more than a factor of 10 and the reduction in σ_{a_2} is almost a factor of 4. This improvement in the accuracy of the results is due to the fact that in addition to the actual data (i.e., the values of x and y) the quality of the data (i.e., the values of σ_y) were also taken into consideration.

We should also question the independence of a_1 and a_2. If for example, we repeat the experiment many times and determine many pairs of values for a_1 and a_2, how will the points be distributed in the two-dimensional space

defined by a_1 and a_2? Are they randomly scattered about the point [a_1 = 0.5786, a_2 = 5.5286] or is there some sort of correlation between these two parameters? An answer to this question is also found in the least squares formulation. The notation σ_{jk} is used for the **covariance** between the parameters j and k and is computed as follows:

$$\sigma_{jk} = \frac{S}{n-p} C^{-1}_{jk} \tag{3.5.4}$$

A more meaningful parameter is the **correlation coefficient** between the parameters j and k. Denoting this parameter as ρ_{jk}, we compute it as follows:

$$\rho_{jk} = \frac{\sigma_{jk}}{\sigma_{a_j}\sigma_{a_k}} \tag{3.5.5}$$

The correlation coefficient is a measure of the degree of correlation between the parameters. The values of ρ_{jk} are in the range from –1 to 1. If the value is zero, then the parameters are uncorrelated (i.e., independent), if the value is 1, then they fall exactly on a line with a positive slope and if the value is –1 then the fall exactly on a line with a negative slope. Examples of different values of ρ_{jk} are seen in Figure 3.5.1.

Returning to our example using unit weighting, let us estimate σ_{12} and ρ_{12}:

$$\sigma_{12} = -1.6019 * 28/196 = -0.2288$$

$$\rho_{12} = \frac{-0.2288}{1.070 * 0.2392} = -0.894$$

In other words, a_1 and a_2 are strongly negatively correlated. Larger-than-average values of a_1 are typically paired with smaller-than-average values of a_2.

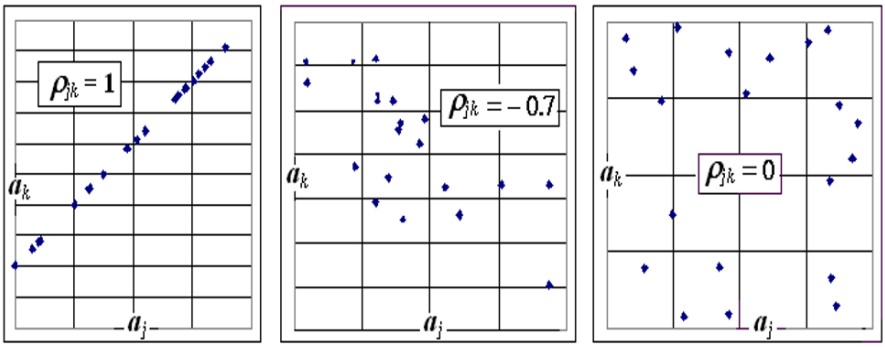

Figure 3.5.1 Correlation Coefficients for Several Different Data Distributions

As will be seen in Section 3.6, the covariance is used in evaluating the standard deviations of the least squares curves. For example, we can use Equation 3.5.2 or 3.5.3 to predict the value of y for any value of x. The covariance is needed to estimate the uncertainty σ_f associated with the predicted value of y (i.e., $f(\mathbf{X})$).

3.6 Uncertainty in the Model Predictions

In Section 3.5 the uncertainties in the model parameters were considered. If the only purpose of the experiment is to determine the parameters of the model, then only these uncertainties are of interest. However, there are many situations in which we are interested in using the model for making predictions. Once the parameters of the model are available, then the equation $f(\mathbf{X})$ can be used to predict y for any combination of the independent variables (i.e., the vector $\mathbf{X}$). In this section attention is turned towards the uncertainties σ_f of these predictions.

Typically, one assumes that the model is "correct" and thus the computed values of y are normally distributed about the true values. For a given set of values for the terms of the X vector (i.e., a combination of the independent variables $x_1, x_2, \ldots x_m$), we assume that the uncertainty in the predicted value of y is due to the uncertainties associated with the a_k's. The predicted value of y is determined by substituting $\mathbf{X}$ into $f(\mathbf{X})$:

$$y = f(\mathbf{X}; a_1, a_2, .., a_p) \tag{3.6.1}$$

Defining Δa_k as the error in a_k, we can estimate Δy (the error in y) by neglecting higher order terms in a Taylor expansion around the true value of y:

$$\Delta f \cong \frac{\partial f}{\partial a_1}\Delta a_1 + \frac{\partial f}{\partial a_2}\Delta a_2 + ... + \frac{\partial f}{\partial a_p}\Delta a_p \tag{3.6.2}$$

To simplify the analysis, let is use the following definition:

$$T_k = \frac{\partial f}{\partial a_k}\Delta a_k \tag{3.6.3}$$

Thus:

$$\Delta f \cong T_1 + T_2 + ... + T_p = \sum_{k=1}^{k=p} T_k \tag{3.6.4}$$

If we square Equation 3.6.2 we get the following:

$$\Delta f^2 \cong (T_1)^2 + (T_2)^2 .. + (T_p)^2 + 2T_1T_2 + 2T_1T_3 .. + 2T_{p-1}T_p \tag{3.6.5}$$

$$\Delta f^2 \cong \sum_{k=1}^{k=p}(T_k)^2 + \sum_{j=1}^{j=p}\sum_{k=j+1}^{k=p} 2T_jT_k \tag{3.6.6}$$

If the experiment is repeated many times and average values of the terms are taken, we obtain the following from Equation 3.6.6:

$$(\Delta f^2)_{avg} \cong \sum_{k=1}^{k=p}((T_k)^2)_{avg} + \sum_{j=1}^{j=p}\sum_{k=j+1}^{k=p} 2(T_jT_k)_{avg} \tag{3.6.7}$$

Recognizing that $((\Delta a_k)^2)_{avg}$ is just $(\sigma_{a_k})^2$ and $(\Delta a_j \Delta a_k)_{avg}$ is σ_{jk} we get the following:

$$\sigma_f^2 = \sum_{k=1}^{k=p} \left(\frac{\partial f}{\partial a_k}\sigma_{a_k}\right)^2 + \sum_{j=1}^{j=p}\sum_{k=j+1}^{k=p} 2\frac{\partial f}{\partial a_j}\frac{\partial f}{\partial a_k}\sigma_{jk} \qquad (3.6.8)$$

The number of cross-product terms (i.e., terms containing σ_{jk}) is $p(p\text{-}1)/2$. If we use the following substitution:

$$\sigma_{kk} = \sigma_{a_k}^2 \qquad (3.6.9)$$

and recognizing that $\sigma_{jk} = \sigma_{kj}$ we can simplify equation 3.6.8:

$$\sigma_f^2 = \sum_{j=1}^{j=p}\sum_{k=1}^{k=p} \frac{\partial f}{\partial a_j}\frac{\partial f}{\partial a_k}\sigma_{jk} \qquad (3.6.10)$$

Using Equations 3.5.1 and 3.5.4, we can relate σ_f to the terms of the C^{-1} matrix:

$$\sigma_f^2 = \frac{S}{n-p}\sum_{j=1}^{j=p}\sum_{k=1}^{k=p} \frac{\partial f}{\partial a_j}\frac{\partial f}{\partial a_k}C_{jk}^{-1} \qquad (3.6.11)$$

As an example of the application of this equation to a data set, let us once again use the data from Table 3.3.1 and $w_i = 1$. The data was fit using a straight line:

$$y = f(x) = a_1 + a_2 x$$

so the derivatives are:

$$\frac{\partial f}{\partial a_1} = 1 \qquad \text{and} \qquad \frac{\partial f}{\partial a_2} = x$$

We have seen that the inverse matrix is:

$$C^{-1} = \frac{1}{196}\begin{bmatrix} 140 & -28 \\ -28 & 7 \end{bmatrix}$$

and the value of $S/(n-p)$ is 1.6019. Substituting all this into Equation 3.6.11 we get the following expression:

$$\sigma_f^2 = \frac{S}{n-p}\left(\frac{\partial f}{\partial a_1}\frac{\partial f}{\partial a_1}C_{11}^{-1} + \frac{\partial f}{\partial a_2}\frac{\partial f}{\partial a_2}C_{22}^{-1} + 2\frac{\partial f}{\partial a_1}\frac{\partial f}{\partial a_2}C_{12}^{-1}\right)$$

$$\sigma_f^2 = \frac{S}{n-p}\left(C_{11}^{-1} + x^2C_{22}^{-1} + 2xC_{12}^{-1}\right)$$

$$\sigma_f^2 = \frac{1.6019}{196}\left(140 + 7x^2 - 56x\right) \qquad (3.6.12)$$

(Note that the C^{-1} matrix is always symmetric so we can use $2C_{jk}^{-1}$ instead of $C_{jk}^{-1} + C_{kj}^{-1}$.)

Equations 3.5.2 and 3.6.12 are used to predict values of y and σ_f for several values of x and the results are seen in Table 3.6.1. Note the curious fact that the values of σ_f are symmetric about $x = 4$. This phenomenon is easily explained by examining Equation 3.6.12 and noting that this equation is a parabola with a minimum value at $x = 4$.

x	$y=f(x)$	σ_f
1.5	8.871	0.766
2.5	14.400	0.598
3.5	19.929	0.493
4.5	25.457	0.493
5.5	30.986	0.598
6.5	36.514	0.766

Table 3.6.1 Predicted values of y and σ_f using w_I=1.

In the table, we see that the values of x that have been chosen are all within the range of the values of x used to obtain the model (i.e., 1 to 7). The use of a model for purposes of extrapolation should be done with extreme caution [WO06]! Note that the σ_f values tend to be least at the midpoint of the range and greatest at the extreme points. This is reasonable. Instinctively if all the data points are weighted equally, we would expect σ_f to be least in regions that are surrounded by many points. Table 3.6.1 was based

upon a least squares analysis in which all points were weighted equally. However, when the points are not weighted equally, the results can be quite different. Table 3.6.2 is also based upon the x and y values from Table 3.3.1 but using statistical weighting (i.e., $w_i=1/\sigma_y^2$).

Table 3.6.2 presents a very different picture than Table 3.6.1 (which is based upon unit weighting). When unit weighting is used differences in the quality of the data are ignored, and we see (in Table 3.3.1) that the relative errors for the first few data points are large. However, when the data is statistically weighted, the relative errors (also seen in Table 3.3.1) are all comparable. In Table 3.6.2 we observe that the values of σ_f are much less than the values in Table 3.6.1 even for points at the upper end of the range. This improvement in accuracy is a result of taking the quality of the data into consideration (i.e., using statistical weighting). Furthermore, the most accurate points (i.e., least values of σ_f) are near the points that have the smallest values of σ_y.

x	$y=f(x)$	σ_f
1.5	9.390	0.044
2.5	14.388	0.089
3.5	19.386	0.151
4.5	24.384	0.215
5.5	29.383	0.281
6.5	34.381	0.347

Table 3.6.2 Predicted values of y and σ_f using $w_i=1/\sigma_y^2$.

In Section 3.4 we noted that increasing or decreasing weight by a constant factor had no effect upon the results (i.e., the resulting A vector). Similarly, changes in the weights do not affect the computed values of σ_f and σ_{a_k}. The value of S and the terms of the C matrix will change proportionally if the w's are changed by a constant factor and the changes in the terms of the C^{-1} matrix will be inversely proportional to the changes in w. The computation of both σ_f and σ_{a_k} are based upon the products of S and terms of the C^{-1} matrix so they will be independent of proportional changes in w. What does makes a difference is the relative values of the weights and not the absolute values. This does not imply that estimates of the actual rather than the relative uncertainties of the data are unimportant.

When designing an experiment, estimates of the absolute uncertainties are crucial to predicting the accuracy of the results of the experiment.

It should be emphasized that values of σ_f computed using Equation 3.6.11 are the σ's associated with the function f and are a measure of how close the least squares curve is to the "true" curve. One would expect that as the number of points increases, the values of σ_f decreases and if the function f is truly representative of the data σ_f will approach zero as n approaches infinity. Equation 3.6.11 does in fact lead to this conclusion. The term $S / (n - p)$ approaches one and the terms of the C matrix become increasingly large for large values of n. The terms of the C^{-1} matrix therefore become increasingly small and approach zero in the limit of n approaching infinity. In fact one can draw a "95% confidence band" around the computed function f. The interpretation of this band is that for a given value of x the probability that the "true" value of f falls within these limits is 95%. Sometimes we are more interested in the "95% prediction band". Within this band we would expect that 95% of new data points will fall [MO03]. This band is definitely not the same as the 95% confidence band and the effect of increasing n has only a small effect upon the prediction band. Assuming that for a given x the deviations from the true curve and from the least squares curve are independent, the σ's associated with the prediction band are computed as follows:

$$\sigma_{pred}^2 = \sigma_f^2 + \sigma_y^2 \tag{3.6.13}$$

Knowing that as n increases, σ_f becomes increasingly small, the limiting value of σ_{pred} is σ_y. In Table 3.6.3 the values of σ_{pred} are computed for the same data as used in Table 3.6.2. The values of σ_y are interpolated from the values in Table 3.3.1. The 95% prediction band is computed using σ_{pred} and the value of t corresponding to 95% limits for $n - p$ degrees of freedom. From Table 3.3.1 we see that $n = 7$ and for the straight line fit $p = 2$. The value of t for $\alpha = 2.5\%$ and 5 degrees of freedom is 2.571. In other words, 2.5% of new points should fall above $f(x) + 2.571\sigma_{pred}$ and 2.5% should fall below $f(x) - 2.571\sigma_{pred}$. The remaining 95% should fall within this band. As n increases, the value of t approaches the value for the standard normal distribution which for a 95% confidence limit is 1.96. The 95% confidence and prediction bands for this data are seen in Figure 3.6.1.

x	$y=f(x)$	σ_y	σ_f	σ_{pred}
1.5	9.390	0.075	0.044	0.087
2.5	14.388	0.150	0.089	0.174
3.5	19.386	0.350	0.151	0.381
4.5	24.384	0.650	0.215	0.685
5.5	29.383	1.150	0.281	1.184
6.5	34.381	2.750	0.347	2.772

Table 3.6.3 Values of σ_{pred} using data from Table 3.3.1 and statistical weighting

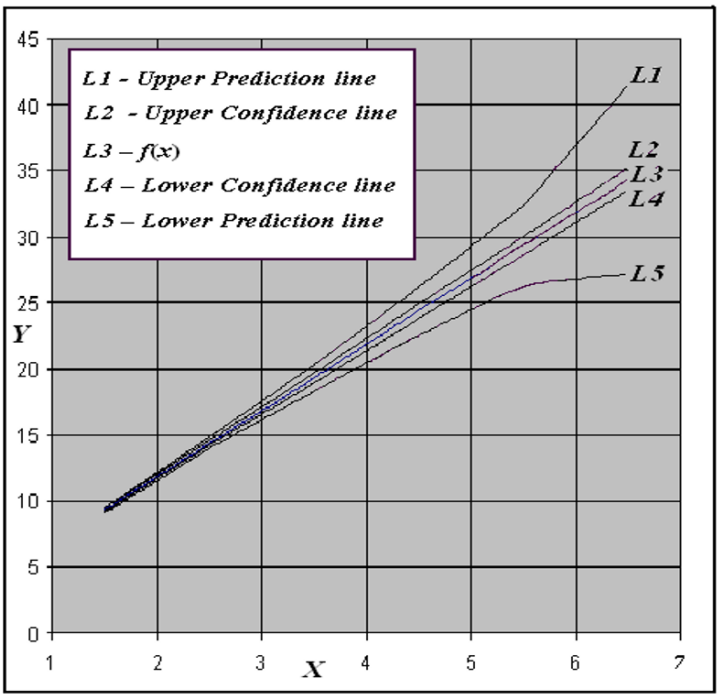

Figure 3.6.1 Confidence and Prediction Bands for Data from Table 3.6.3

3.7 Treatment of Prior Estimates

In the previous sections we noted that a basic requirement of the method of least squares is that the number of data points ***n*** must exceed ***p*** (the number of unknown parameters of the model). The difference between these two numbers ***n-p*** is called the "number of degrees of freedom". Very early in

my career I came across an experiment in which the value of $n-p$ was in fact negative! The modeling effort was related to damage caused by a certain type of event and data had been obtained based upon only two events. Yet the model included over ten unknown parameters. The independent variables included the power of the event and other variables related to position. To make up the deficit, estimates of the parameters based upon theoretical models were used to supplement the two data points. The prior estimates of the parameters are called Bayesian estimators and if the number of Bayesian estimators is n_b then the number of degrees of freedom is $n+n_b-p$. As long as this number is greater than zero, a least squares calculation can be made.

In Section 3.2 Equation 3.2.6 is the modified form that the objective function takes when prior estimates of the a_k parameters are available:

$$S = \sum_{i=1}^{i=n} w_i (Y_i - f(\mathbf{X}_i))^2 + \sum_{k=1}^{k=p} (a_k - b_k)^2 / \sigma_{b_k}^2$$

In this equation b_k is the prior estimates of a_k and σ_{b_k} is the uncertainty associated with this prior estimate. The parameter b_k is typically used as the initial guess $a0_k$ for a_k. We see from this equation that each value of b_k is treated as an additional data point. However, if σ_{b_k} is not specified, then it is assumed to be infinite and no weight is associated with this point. In other words, if σ_{b_k} is not specified then b_k is treated as just an initial guess for a_k and not as a prior estimate. The number of values of b_k that are specified (i.e., not infinity) is n_b.

In the previous sections it was stated that the weights w_i could be based upon relative and not absolute values of the uncertainties associated with the data points. When prior estimates of the a_k's are included in the analysis, we are no longer at liberty to use relative weights. Since the weights associated with the prior estimates are based upon estimates of absolute values (i.e., $1 / (\sigma_{b_k})^2$), the w_i values must also be based upon estimates of absolute values.

To find the least squares solution, we proceed as in Section 3.4 by setting the p partial derivatives of S to zero yielding p equations for the p unknown values of a_k:

$$-2\sum_{i=1}^{i=n} w_i(Y_i - f(\mathrm{X}_i))\frac{\partial f(\mathrm{X}_i)}{\partial a_k} + 2\sum_{k=1}^{k=p}(a_k - b_k)/\sigma_{b_k}^2 = 0 \qquad k = 1 \text{ to } p$$

The terms in the last summation can be expanded:

$$(a_k - b_k)/\sigma_{b_k}^2 = (a_k - a0_k + a0_k - b_k)/\sigma_{b_k}^2 = A_k/\sigma_{b_k}^2 + (a0_k - b_k)/\sigma_{b_k}^2$$

$$\sum_{i=1}^{i=n} w_i f(\mathrm{X}_i)\frac{\partial f(\mathrm{X}_i)}{\partial a_k} + \sum_{k=1}^{k=p}\frac{A_k}{\sigma_{b_k}^2} = \sum_{i=1}^{i=n} w_i Y_i \frac{\partial f(\mathrm{X}_i)}{\partial a_k} + \sum_{k=1}^{k=p}\frac{b_k - a0_k}{\sigma_{b_k}^2} \qquad (3.7.1)$$

As in Section 3.4 this equation is best treated as a matrix equation:

$$CA = V$$

The diagonal terms of the C matrix are modified but the off-diagonal terms remain the same:

$$C_{jk} = C_{kj} = \sum_{i=1}^{i=n} w_i \frac{\partial f}{\partial a_j}\frac{\partial f}{\partial a_k} \qquad (j \neq k) \qquad (3.7.2)$$

$$C_{kk} = \frac{1}{\sigma_{b_k}^2} + \sum_{i=1}^{i=n} w_i \frac{\partial f}{\partial a_j}\frac{\partial f}{\partial a_k} \qquad (3.7.3)$$

The terms of the V vector are also modified:

$$V_k = \frac{b_k - a0_k}{\sigma_{b_k}^2} + \sum_{i=1}^{i=n} w_i Y_i \frac{\partial f}{\partial a_k} \qquad (3.7.4)$$

Solution of the matrix equation $CA = V$ yields the vector A which is then used to compute the unknown a_k's (Equation 3.4.16). The computation of the σ_{a_k} terms must be modified to include the additional data points. The modified form of Equation 3.5.1 is:

$$\sigma_{a_k} = \left(\frac{S}{n + n_b - p} C_{kk}^{-1}\right)^{1/2} \tag{3.7.5}$$

In this equation n_b is the number of Bayesian estimations included in the analysis (i.e., the number of b_k's that are specified. Using the same reasoning, Equation 3.6.11 must also be modified:

$$\sigma_f^2 = \frac{S}{n + n_b - p} \sum_{j=1}^{j=p} \sum_{k=1}^{k=p} \frac{\partial f}{\partial a_j} \frac{\partial f}{\partial a_k} C_{jk}^{-1} \tag{3.7.6}$$

As an example of the application of prior estimates, let us once again use the data in Table 3.3.1 but only for the case of statistical weighting (i.e., $w=1/\sigma_y^2$). The straight line computed for this case was:

$$y = f(x) = a_1 + a_2 x = 1.8926 + 4.9982x$$

The computed value of σ_{a_1} and σ_{a_2} were 0.0976 and 0.0664. The C matrix and V vector for this case are:

$$C = \begin{bmatrix} 531.069 & 701.917 \\ 701.917 & 1147.125 \end{bmatrix} \qquad \text{and} \qquad V = \begin{bmatrix} 4513.39 \\ 7016.96 \end{bmatrix}$$

Let us say that we have a prior estimate of a_1:

$$b_1 = 1.00 \pm 0.10$$

The only term in the C matrix that changes is C_{11}. The terms of the V vector are, however, affected by the changes in the values of Y_i. Since we start from the initial guess for a_1, all the values of Y_i are reduced by a_1 (i.e., 1.00) :

$$C = \begin{bmatrix} 631.069 & 701.917 \\ 701.917 & 1147.125 \end{bmatrix} \qquad \text{and} \qquad V = \begin{bmatrix} 3982.32 \\ 6360.04 \end{bmatrix}$$

Solving for the terms of the A vector we get $A_1 = 0.4498$ and $A_2 = 5.2691$. The computed value of a_1 is therefore 1.4498 and a_2 is 5.2691. Note that the prior estimate of a_1 reduces the value previously computed from 1.8926 towards the prior estimate of 1.00. The values of σ_{a_1} and σ_{a_2} for this calculation were 0.1929 and 0.1431. These values are greater than the values obtained without the prior estimate and that indicates that the prior estimate of a_1 is not in agreement with the experimental results. Assuming that there is no discrepancy between the prior estimates and the experimental data, we would expect a reduction in uncertainty. For example, if we repeat the analysis but use as our prior estimate: $b_1 = 2.00 \pm 0.10$:

The resulting values of a_1 and a_2 are:

$$a_1 = 1.9459 \pm 0.0670 \qquad a_2 = 4.9656 \pm 0.0497$$

If we repeat the analysis and use prior estimates for both a_1 and a_2:

$$b_1 = 2.00 \pm 0.10 \qquad b_2 = 5.00 \pm 0.05$$

The resulting values of a_1 and a_2 are:

$$a_1 = 1.9259 \pm 0.0508 \qquad a_2 = 4.9835 \pm 0.0325$$

The results for all these cases are summarized in Table 3.7.1.

b_1	b_2	$n+n_b$	a_1	σ_{a_1}	a_2	σ_{a_2}
none	None	7	1.8926	0.0976	4.9982	0.0664
1.00±0.1	None	8	1.4498	0.1929	5.2691	0.1431
2.00±0.1	None	8	1.9459	0.0670	4.9656	0.0497
2.00±0.1	5.00±0.05	9	1.9259	0.0508	4.9835	0.0325

Table 3.7.1 Computed values of a_1 and a_2 for combinations of b_1 and b_2.

We see in this table that the best results (i.e., minimum values of the σ' s) are achieved when the prior estimates are in close agreement with the results obtained without the benefit of prior estimates of the unknown parameters a_1 and a_2. For example when b_1 is close to the least squares value of a_1 without prior estimates (i.e., a_1=1.89 and b_1=2.00) the accuracy is

improved (i.e., 0.0607 vs. 0.0967). However, when b_1 is not close, (i.e., b_1=1.00) resulting accuracy is worse (i.e., 0.1929 vs. 0.0967).

3.8 Applying Least Squares to Classification Problems

In the previous sections the dependent variable ***y*** was assumed to be a continuous numerical variable and the method of least squares was used to develop models that could then be used to predict the value of ***y*** for any combination of the independent ***x*** variable (or variables). There are, however, problems in which the dependent variable is a "class" rather than a continuous variable. For example the problem might require a model that differentiates between two classes: "good" or "bad" or three levels: "low", "medium" or "high". Typically we have ***nlrn*** learning points that can be used to create the model and then ***ntst*** test points that can be used to test how well the model predicts on unseen data. The method of least squares can be applied to classification problems in a very straight-forward manner.

The trick that allows a very simple least squares solution to classification problems is to assign numerical values to the classes (i.e., the ***y*** values) and then make predictions based upon the computed value of ***y*** for each test point. For example, for two class problems we can assign the values 0 and 1 to the two classes (e.g., "bad" = 0 and "good" = 1). We then fit the learning data using least squares as the modeling technique and then for any combination of the ***x*** variables, we compute the value of ***y***. If it is less than 0.5 the test point is assumed to fall within the "bad" class, otherwise it is classified as "good". For 3 class problems we might assign 0 to class 1, 0.5 to class 2 and 1 to class 3. If a predicted value of ***y*** is less than 1/3 then we would assign class 1 as our prediction, else if the value was less than 2/3 we would assign class 2 as our prediction, otherwise the assignment would be class 3. Obviously the same logic can be applied to any number of classes.

It should be emphasized that the least squares criterion is only one of many that can be used for classification problems. In their book on ***Statistical Learning***, Hastie, Tibshirani and Friedman discuss a number of alternative criteria but state that "squared error is analytically convenient and is the most popular" [HA01]. The general problem is to minimize a **loss function** $L(Y, f(\mathbf{X}))$ that penalizes prediction errors. The least squares loss function is $\Sigma(Y - f(\mathbf{X}))^2$ but other loss functions (e.g. $\Sigma \mid Y - f(\mathbf{X}) \mid$) can also be used.

A different approach to classification problems is based upon **nearest neighbors** [WO00].

In my previous book (***Data Analysis using the Method of Least Squares***) [WO06] I discuss the application of least squares to classification problems in detail with several examples to illustrate the process. However, the emphasis in this book is upon design of experiments using the method of prediction analysis and this method is not applicable to classification problem. What one would like to be able to estimate at the design stage for classification problems is the mis-classification rate and this is not possible using prediction analysis.

3.9 Goodness-of-Fit

In Section 2.4 the χ^2 (chi-squared) distribution was discussed. Under certain conditions, this distribution can be used to measure the **goodness-of-fit** of a least squares model. To apply the χ^2 distribution to the measurement of goodness-of-fit, one needs estimates of the uncertainties associated with the data points. In Sections 3.3 it was emphasized that only relative uncertainties were required to determine estimates of the uncertainties associated with the model parameters and the model predictions. However, for goodness-of-fit calculations, **estimates of absolute uncertainties are required.** When such estimates of the absolute uncertainties are unavailable, the best approach to testing whether or not the model is a good fit is to examine the residuals. This subject is considered in Section 3.9 of my book on least squares [WO06].

The goodness-of-fit test is based upon the value of $\boldsymbol{S / (n-p)}$. Assuming that $\boldsymbol{S}$ is based upon reasonable estimates of the uncertainties associated with the data points, if the value of $\boldsymbol{S / (n-p)}$ is much less than one, this usually implies some sort of misunderstanding of the experiment (for example, the estimated errors are vastly over-estimated). If the value is much larger than one, then one of the following is probably true:

1) The model does not adequately represent the data.

2) Some or all of the data points are in error.

3) The estimated uncertainties in the data are erroneous (typically vastly underestimated).

Assuming that the model, the data and the uncertainty estimates are correct, the value of S (the weighted sum of the residuals) will be distributed according to a χ^2 distribution with $n-p$ degrees of freedom. Since the expected value of a χ^2 distribution with $n-p$ degrees of freedom is $n-p$, the expected value of $S / (n-p)$ is one. If one assumes that the model, data and estimated uncertainties are correct, the computed value of $S / (n-p)$ can be compared with $\chi^2 / (n-p)$ to determine a probability of obtaining or exceeding the computed value of S. If this probability is too small, then the goodness-of-fit test fails and one must reconsider the model and/or the data.

Counting experiments (in which the data points are numbers of events recorded in a series of windows (e.g., time or space) are a class of experiments in which estimates of the absolute uncertainty associated with each data point are available. Let us use Y_i to represent the number of counts recorded in the time window centered about time t_i. According to Poisson statistics the expected value of σ_i^2 (the variance associated with Y_i) is just Y_i and the weight associated with this point is $1/Y_i$. From Equation 2.2.1 we get the following expression for S:

$$S = \sum_{i=1}^{i=n} w_i R_i^2 = \sum_{i=1}^{i=n} w_i (Y_i - y_i)^2 = \sum_{i=1}^{i=n} (Y_i - f(t_i))^2 / Y_i \qquad (3.9.1)$$

Since the expected value of $(Y_i - y_i)^2$ is $\sigma_i^2 = Y_i$ the expected value of $w_i R_i^2$ is one. The expected value of S is not n as one might expect from this equation. If the function f includes p unknown parameters, then the number of degrees of freedom must be reduced by p and therefore the expected value of S is $n–p$. This might be a confusing concept but the need for reducing the expected value can best be explained with the aid of a qualitative argument. Lets assume that n is 3 and we use a 3 parameter model to fit the data. We would expect the model to go thru all 3 points and therefore, the value of S would be zero which is equal to $n–p$.

To illustrate this process, let us use data included in Bevington and Robinson's book ***Data Reduction and Error Analysis*** [BE03]. The data is presented graphically in Figure 3.9.1 and in tabular form in Table 3.9.1. The data is from a counting experiment in which a Geiger counter was used to detect counts from an irradiated silver sample recorded in 15 second inter-

vals. The 59 data points shown in the table include two input columns (t_i and Y_i) and one output column that is the residual divided by the standard deviation of Y_i (i.e., R_i / σ_i):

$$R_i / \sigma_i = (Y_i - y_i) / \sqrt{Y_i} \tag{3.9.2}$$

The data was modeled using a 5 parameter equation that included a background term and two decaying exponential terms:

$$y = a_1 + a_2 e^{-t / a_4} + a_3 e^{-t / a_5} \tag{3.9.3}$$

i	t_i	Y_i	R_i / σ_i	i	t_i	Y_i	R_i / σ_i
1	15	775	0.9835	31	465	24	-0.0208
2	30	479	-1.8802	32	480	30	1.2530
3	45	380	0.4612	33	495	26	0.7375
4	60	302	1.6932	34	510	28	1.2466
5	75	185	-1.6186	35	525	21	0.0818
6	90	157	-0.4636	36	540	18	-0.4481
7	105	137	0.3834	37	555	20	0.1729
8	120	119	0.7044	38	570	27	1.6168
9	135	110	1.3249	39	585	17	-0.2461
10	150	89	0.4388	40	600	17	-0.1141
11	165	74	-0.2645	41	615	14	-0.7922
12	180	61	-1.0882	42	630	17	0.1231
13	195	66	0.2494	43	645	24	1.6220
14	210	68	1.0506	44	660	11	-1.4005
15	225	48	-1.0603	45	675	22	1.4360
16	240	54	0.2938	46	690	17	0.5068
17	255	51	0.3200	47	705	12	-0.7450
18	270	46	0.0160	48	720	10	-1.3515
19	285	55	1.5750	49	735	13	-0.0274
20	300	29	-2.2213	50	750	16	0.5695
21	315	28	-2.0393	51	765	9	-1.4914
22	330	37	0.0353	52	780	9	-1.4146
23	345	49	2.0104	53	795	14	0.2595
24	360	26	-1.4128	54	810	21	1.7830
25	375	35	0.5740	55	825	17	1.0567
26	390	29	-0.2074	56	840	13	0.1470
27	405	31	0.4069	57	855	12	-0.0891
28	420	24	-0.7039	58	870	18	1.3768
29	435	25	-0.2503	59	885	10	-0.6384
30	450	35	1.6670				

Table 3.9.1 Input data (t and Y) from Table 8.1, Bevington and Robinson [BE03].

The input data in the table was analyzed using the REGRESS program [see Section 3.10] and yielded the following equation:

$$y = 10.134 + 957.77e^{-t/34.244} + 128.29e^{-t/209.68}$$

For example, for the first data point (i.e., t =15), the computed value of y according to this equation is 747.62. The **relative errors** included in the table (i.e., R_i/σ_i) are computed using Equation 3.9.2. Thus the value of the relative error for the first point is (775 – 747.62) / **sqrt**(775) = 0.9835. Note that the relative errors are distributed about zero and range from -2.2213 to 2.0104. The value of S is the sum of the squares of the relative errors and is 66.08. The number of points n is 59 and the number of unknown parameters p is 5 so the value of $S/(n-p)$ is 1.224. The goodness-of-fit test considers the closeness of this number to the most probable value of one for a correct model. So the question that must be answered is: how close to one is 1.224?

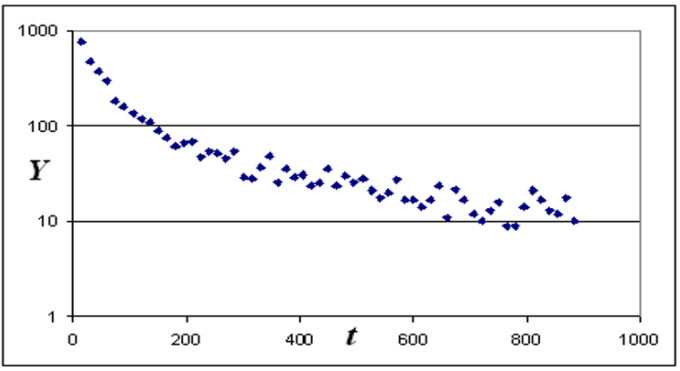

Figure 3.9.1 Bevington and Robinson data [BE03]

In Section 2.4 it was mentioned that for the χ^2 distribution with k degrees of freedom, the mean is k and the standard deviation is $\sqrt{2k}$. Furthermore as k becomes larger, the χ^2 distribution approaches a normal distribution. We can use these properties of the distribution to estimate the probability of obtaining a value of S greater or equal to 66.08 for a χ^2 distribution with 54 degrees of freedom:

$$66.08 = 54 + x_p * \sqrt{2 * 54} = 54 + x_p * 10.39 \rightarrow x_p = (66.08 - 54) / 10.39 = 1.163$$

In other words, 66.08 is approximately x_p = 1.16 standard deviations above the expected value of 54. From a table of the normal distribution we can verify that the integral from 0 to 1.163 standard deviations is 0.378, so the probability of exceeding this value is 0.5 – 0.378 = 0.122 (i.e., about 12%). Typically one sets a value of the probability at which one would reject the model. For example, if this probability is set at 1%, then the lower limit of x_p for rejecting the model would be 2.326. Since our computed value of x_p of 1.163 is much less than 2.326, we have no reason to reject the five parameter model.

If we really want to be pedantic, we can make a more accurate calculation. From the ***Handbook of Mathematical Functions*** [AB64], Equation 26.4.17 is suggested for values of $k > 30$:

$$\chi^2 = k\left[1 - \frac{2}{9k} + x_p\sqrt{\frac{2}{9k}}\right]^3 \tag{3.9.4}$$

In this equation x_p is the number of standard deviations for a standard normal distribution to achieve a particular probability level. For example, if we wish to determine the value of the χ^2 distribution with k degrees of freedom for which we could expect 1% of all values to exceed this level, we would use the standard normal distribution value of x_p = 2.326. Using Equation 3.9.4 we can compute the value of x_p corresponding to a χ^2 value of 66.08:

$$66.08 = 54[1 - 2/(9 * 54) + x_p\sqrt{2/(9 * 54)}]^3$$

Solving this equation for x_p we get a value of 1.148 which is close to the value of 1.163 obtained above using the normal distribution approximation.

Equation 3.9.3 is a 5 parameter model that includes a background term and two decaying exponential terms. If we were to simplify the equation to include only a single exponential term, would we still have a "reasonable" model? Would this equation pass the goodness-of-fit test? The proposed alternative model is:

$$y = a_1 + a_2 e^{-t/a_3} \tag{3.9.5}$$

Once again using the REGRESS program the resulting equation is:

$$y = 18.308 + 752.99e^{-t/62.989}$$

and the value of S is 226.7. The value of x_p corresponding to this value of S is estimated as follows:

$$226.7 = 56 + x_p * \sqrt{2 * 56}$$

$$x_p = (226.7 - 56) / 10.39 = 15.28$$

This value is so large that we would immediately reject the proposed model. The probability of getting a value of S that is over 15 standard deviations above the expected value for a correct model is infinitesimal. (Note that since this is only a 3 parameter fit to the data, the number of degrees of freedom is increased from 54 to 56.)

3.10 The REGRESS Program

Throughout this book much of the data analysis examples were created using the REGRESS program. The REGRESS program falls within the category of non-linear regression (NLR) programs [WO06]. I particularly like REGRESS because I wrote it! The history of the development of this program goes back to my early career when I was in charge of designing a sub-critical heavy water nuclear reactor facility. One of the experiments that we planned to run on the facility involved a nonlinear regression based upon the following nonlinear function:

$$y = \frac{\left(1 + a_1^2 x_1\right)\left(1 + a_2^2 x_1\right)}{\left(1 + a_1^2 x_2\right)\left(1 + a_2^2 x_2\right)}$$

In the 1960's commercial software was rare so we had no choice other than writing our own programs. It became quite apparent that I could generalize the software to do functions other than this particular function. All that had to be done was to supply a function to compute $f(x)$ and another function to compute the required derivatives. We would then link these functions to the software and could thus reuse the basic program with any desired function. At the time we called the program ANALYZER.

In the early 1970's I discovered a language called FORMAC that could be used for symbolic manipulation of equations. FORMAC was compatible with FORTRAN and I used FORTRAN and FORMAC to write a program similar to ANALYZER and I called the new program REGRESS. The REGRESS program accepted equations as input quantities, using FORMAC automatically generated equations for the derivatives, and then created FORTRAN subroutines that could then be used to perform the nonlinear regression. All these steps, including compilation and link-editing of the subroutines, were performed automatically without any user intervention. The REGRESS program became a commercial product on the NCSS time-sharing network and I had the opportunity to work with a number of NCSS clients and learned about many different applications of non-linear regression (NLR). In particular, working with these clients I discovered what users needed in the software.

In the mid 1970's I realized that with languages that support recursive programming, I could avoid the need to externally compile subroutines. Recursion is the ability to call a subroutine from within itself. Using recursion, it became a doable task to write a routine to symbolically differentiate functions. Using PL/1 I rewrote REGRESS and added many new features that I realized were desirable from conversations with a number of users of REGRESS. I've returned to the REGRESS program on many occasions since the original version. In the 1980's I started teaching a graduate course called Design and Analysis of Experiments and I supplied REGRESS to the students. Many of the students were doing experimental work as part of their graduate research and the feedback from their experiences with REGRESS stimulated a number of interesting developments. In the early 1990's I rewrote REGRESS in the C language and eventually it migrated into C++. Through the many version changes REGRESS has evolved over the years and is still evolving.

The REGRESS program lacks some features that are included in some of the other general NLR programs. A serious problem with REGRESS was the need to create data files in a format that the program could understand. Many users of the program gather data that ends up in an Excel Spread Sheet. The problem for such users was how to get the data into REGRESS. It turned out that the solution was quite simple: Excel allows users to create text files. A feature was added to accept Excel text files. Another important issue was the creation of graphics output. One of the features of REGRESS is that the entire interactive session is saved as a text file. The current method for obtaining graphics output is to extract the output data from the text file and then input it into programs such as Excel

and Matlab that supports graphics. Since this turns out to be a relatively painless process, the need for REGRESS to generate graphic output is not a pressing issue. It would be nice if REGRESS could seamlessly interface with a graphics program but so far I have not attempted to create such an interface.

The REGRESS program includes some features that are generally not included in other NLR programs. The most important feature in REGRESS that distinguishes it from other general purpose NLR programs is the Prediction Analysis (experimental design) feature described in Chapter 4. All of the examples in the subsequent chapters were created using the Prediction Analysis feature of REGRESS. Another important feature that I have not seen in other general purpose NLR programs is the ***int*** operator. This is an operator that allows the user to model initial value nonlinear integral equations. This operator was used extensively in the runs required to generate the results in Chapter 6. As an example of the use of the ***int*** operator, consider the following set of two equations:

$$\begin{aligned} y_1 &= a_1 \int_0^x y_2 dx + a_2 \\ y_2 &= a_3 \int_0^x y_1 dx + a_4 \end{aligned} \tag{3.10.1}$$

These highly nonlinear and recursive equations can be modeled in REGRESS as follows:

```
y1 = 'a1 * int(y2, 0, x) + a2'
y2 = 'a3 * int(y1, 0, x) + a4'
```

This model is recursive in the sense that ***y*1** is a function of ***y*2** and ***y*2** is a function of ***y*1.** Not all general purpose NLR programs support recursive models. The user supplies values of ***x***, ***y*1** and ***y*2** for ***n*** data points and the program computes the least squares values of the a_k 's.

Another desirable REGRESS feature is a simple method for testing the resulting model on data that was not used to obtain the model. In REGRESS the user invokes this feature by specifying a parameter called NEVL (number of evaluation data points). Figure 3.10.1 includes some of the REGRESS output for a problem based upon Equation 3.10.1 in which the number of data records for modeling was 8 and for evaluation was 7. Each data record included values of ***x***, ***y*1** and ***y*2** (i.e., a total of 16 modeling and

14 evaluation values of ***y***). This problem required 15 iterations to converge.

```
   Function Y1:    A1 * INT(Y2,0,X) + A2
   Function Y2:    A3 * INT(Y1,0,X) + A4

   K     A0(K)   AMIN(K)   AMAX(K)      A(K)   SIGA(K)
   1   0.50000 Not Spec Not Spec   1.00493  0.00409
   2   1.00000 Not Spec Not Spec   2.00614  0.00459
   3   0.00000 Not Spec Not Spec  -0.24902  0.00079
   4  -1.00000 Not Spec Not Spec  -3.99645  0.00663

 Evaluation of Model for Set 1:
   Number of points in evaluation data set:      14
   Variance Reduction                          100.00
 (Average)
     VR:            Y1                        100.00
     VR:            Y2                        100.00
   RMS (Y - Ycalc)                           0.01619
 (all data)
       RMS (Y-Yc) - Y1                       0.02237
   RMS (Y-Yc)/Sy) - Y1                       0.00755
       RMS (Y-Yc) - Y2                       0.00488
   RMS (Y-Yc)/Sy) - Y2                       0.00220
   Fraction Y_eval positive                :   0.214
   Fraction Y_calc positive                :   0.214

 DataSet  Var  Minimum  Maximum  Average   Std_dev
 Modeling  X1   0.0100   6.2832   1.6970    2.3504
 Modeling  Y1  -7.9282   2.0000  -1.2393    3.7499
 Modeling  Y2  -4.1189   4.0000  -2.2600    3.1043

 Evaluate  X1   0.1500   5.2360   1.6035    1.8876
 Evaluate  Y1  -8.0000   1.3900  -2.1940    3.4524
 Evaluate  Y2  -4.1169   2.9641  -2.6260    2.7180
```

Figure 3.10.1 Output for a recursive problem with the *int* operator and Evaluation data set

Chapter 4 PREDICTION ANALYSIS

4.1 Introduction

Prediction analysis is a general method for predicting the accuracy of quantitative experiments. The method is applicable to experiments in which the data is to be analyzed using the method of least squares. Use of prediction analysis allows the designer of an experiment to estimate the accuracy that should be obtained from the experiment before equipment has been acquired and the experimental setup finalized. Through use of the method, a picture is developed showing how the accuracy of the results should be dependent upon the experimental variables. This information then forms the basis for answering the following questions:

1. Can the proposed experiment yield results with sufficient accuracy to justify the effort in performing the experiment?

2. If the answer to 1 is positive, what choices of the experimental variables (e.g., the number of data points, choice of the independent variables for each data point, accuracy of the independent variables) will satisfy the experimental objectives?

3. Which combination of the experimental variables optimizes the resulting experiment?

In this chapter the method of prediction analysis is developed. The following chapters discuss a variety of experimental classes.

J. Wolberg, *Designing Quantitative Experiments*, DOI 10.1007/978-3-642-11589-9_4,

4.2 Linking Prediction Analysis and Least Squares

In Chapter 3 the method of least squares was developed. Prediction analysis is based upon the assumption that once an experiment has been performed and data collected, the data will then be analyzed using the method of least squares. There are several compelling reasons for using the method of least squares to analyze data:

1. If the errors associated with the dependent variables can be assumed to be normally distributed, then the method yields the "best estimates" of the parameters of the fitting function [WO06].

2. The method of least squares includes as by-products, estimates of the uncertainties associated with the parameters of the fitting function.

3. The method of least squares also includes estimates of the uncertainties associated with the fitting function itself (for any combination of the independent variables).

Equation 3.5.1 is the expression that is used to compute the uncertainty associated with the model parameters:

$$\sigma_{a_k}^2 = \frac{S}{n-p} C_{kk}^{-1} \tag{3.5.1}$$

Equation 3.6.11 is the expression that is used to compute the uncertainty associated with the fitting function:

$$\sigma_f^2 = \frac{S}{n-p} \sum_{j=1}^{j=p} \sum_{k=1}^{k=p} \frac{\partial f}{\partial a_j} \frac{\partial f}{\partial a_k} C_{jk}^{-1} \tag{3.6.11}$$

The terms of the $\boldsymbol{C^{-1}}$ matrix are only dependent upon the values of the independent variables, and the partial derivatives in Equation 3.6.11 are computed using the values of the independent variables and the values of the model parameters (i.e., the a_k's). To estimate the values of σ_{ak} and σ_f we only need to estimate S (i.e., the weighted sum of the squares of the residuals associated with the fitting function). The calculation of S is discussed in detail in Section 3.2.

In Section 3.9 the concept of Goodness-of-Fit was introduced. The goodness-of-fit test assumes that the value of S is χ^2 (chi-squared) distributed with $n-p$ degrees of freedom. In other words, the mean value of the S distribution is $n-p$ and the mean value of S / $n-p$ is one! This is a beautiful revelation and implies that our best estimates of the accuracies that we can expect for our proposed experiment is obtained by applying this result to Equations 3.5.1 and 3.6.11:

$$\sigma_{a_k}^2 = C_{kk}^{-1} \tag{4.2.1}$$

$$\sigma_f^2 = \sum_{j=1}^{j=p} \sum_{k=1}^{k=p} \frac{\partial f}{\partial a_j} \frac{\partial f}{\partial a_k} C_{jk}^{-1} \tag{4.2.2}$$

Equations 4.2.1 and 4.2.2 are the basic equations of the method of prediction analysis. All that is needed to estimate the expected accuracy of the parameters of a model (i.e., the a_k's) is to compute the elements of the C matrix, invert the matrix and then use the diagonal elements of the inverted matrix to estimate the values of σ_{ak}. To estimate the accuracy of the fitting function (i.e., σ_f for any point on the surface of the function), in addition to the inverse matrix, the partial derivatives must also be estimated. In the next section prediction analysis of a very simple experiment is used to illustrate the method: a straight line fit to a set of data.

4.3 Prediction Analysis of a Straight Line Experiment

In this section we analyze a proposed experiment in which the dependent variable y will be related to the independent variable x by a straight line:

$$y = f(x) = a_1 + a_2 x \tag{4.3.1}$$

Many real-world experiments use this very simple fitting function to model data. This equation is the simplest case of the multiple linear regression equation discussed in Section 4.10. This model can be analyzed analytically and thus we can obtain equations that can be used to predict accuracy. The first step is to propose an experiment. Let us make the following assumptions:

1. All values of y are measured to a given accuracy: σ_y. The actual value of σ_y is not important. This value will be one of the design parameters of the experiment.

2. The values of x are obtained with negligible uncertainty (i.e., assume $\sigma_x = 0$).

3. The number of data points is n.

4. The points are equally spaced (i.e., Δx constant).

From Equation 3.4.11 we can predict the terms of the C matrix:

$$C = \begin{bmatrix} \sum_{i=1}^{i=n} w_i & \sum_{i=1}^{i=n} w_i x_i \\ \sum_{i=1}^{i=n} w_i x_i & \sum_{i=1}^{i=n} w_i x_i^2 \end{bmatrix} \tag{3.4.11}$$

The values of w_i are all equal to $1 / \sigma_y^2$ so this constant can be removed from all the summations:

$$C = \frac{1}{\sigma_y^2} \begin{bmatrix} n & \sum_{i=1}^{i=n} x_i \\ \sum_{i=1}^{i=n} x_i & \sum_{i=1}^{i=n} x_i^2 \end{bmatrix} \tag{4.3.2}$$

The terms C_{12} and C_{21} are simply $n\, x_{avg} / \sigma_y^2$. The assumption of equally spaced values of x allows us to estimate C_{22} by replacing the summation with an integral :

$$\sum_{i=1}^{i=n} x_i^2 \rightarrow \frac{n}{b-a} \int_a^b x^2 dx = \frac{n(b^3 - a^3)}{3(b-a)} = \frac{n}{3}(b^2 + ba + a^2) \tag{4.3.3}$$

For large values of n the range a to b approaches x_1 to x_n but for smaller values of n a more accurate approximation is obtained by using the range $a = x_1 - \Delta x/2$ to $b = x_n + \Delta x/2$.

Substituting 4.3.3 into Equations 4.3.2 the $\boldsymbol{C}$ matrix can now be expressed as follows:

$$\boldsymbol{C} = \frac{\boldsymbol{n}}{\sigma_y^2}\begin{bmatrix} \boldsymbol{1} & \boldsymbol{(b+a)/2} \\ \boldsymbol{(b+a)/2} & \boldsymbol{\frac{1}{3}(b^2+ba+a^2)} \end{bmatrix} \tag{4.3.4}$$

Inverting this matrix:

$$\boldsymbol{C^{-1}} = \frac{\sigma_y^2}{\boldsymbol{nD}}\begin{bmatrix} \boldsymbol{\frac{1}{3}(b^2+ba+a^2)} & \boldsymbol{-(b+a)/2} \\ \boldsymbol{-(b+a)/2} & \boldsymbol{1} \end{bmatrix}$$

Where $\boldsymbol{D}$ is the determinant of the matrix:

$$\boldsymbol{D} = \frac{\boldsymbol{1}}{\boldsymbol{3}}\boldsymbol{(b^2+ba+a^2)} - \frac{\boldsymbol{(b+a)^2}}{\boldsymbol{4}} = \frac{\boldsymbol{1}}{\boldsymbol{12}}\boldsymbol{(b-a)^2}$$

$$\boldsymbol{C^{-1}} = \frac{\boldsymbol{12}\sigma_y^2}{\boldsymbol{n(b-a)^2}}\begin{bmatrix} \boldsymbol{\frac{1}{3}(b^2+ba+a^2)} & \boldsymbol{-(b+a)/2} \\ \boldsymbol{-(b+a)/2} & \boldsymbol{1} \end{bmatrix} \tag{4.3.5}$$

From Equations 4.3.5 and 4.2.1 we can predict the variances $\sigma_{a_1}^2$ and $\sigma_{a_2}^2$ and the covariance σ_{12}:

$$\sigma_{a_1}^2 = \boldsymbol{C_{11}^{-1}} = \frac{\boldsymbol{4}\sigma_y^2}{\boldsymbol{n(b-a)^2}}\boldsymbol{(b^2+ba+a^2)} \tag{4.3.6}$$

$$\sigma_{a_2}^2 = \boldsymbol{C_{22}^{-1}} = \frac{\boldsymbol{12}\sigma_y^2}{\boldsymbol{n(b-a)^2}} \tag{4.3.7}$$

$$\sigma_{12} = \boldsymbol{C_{12}^{-1}} = \boldsymbol{C_{21}^{-1}} = \frac{\boldsymbol{-12}\sigma_y^2\boldsymbol{(b+a)/2}}{\boldsymbol{n(b-a)^2}} \tag{4.3.8}$$

We can put these equations into a more useable form by noting that the range $b - a$ is $n\Delta x$ and defining a dimensionless parameter r_x :

$$r_x \equiv \frac{x_{avg}}{n\Delta x} = \frac{(b+a)/2}{b-a} \tag{4.3.9}$$

Using 4.3.9 we obtain the following expressions for the variances and covariance:

$$\sigma_{a_1}^2 = \frac{4\sigma_y^2}{n}(3r_x^2 + \frac{1}{4}) = \frac{\sigma_y^2}{n}(12r_x^2 + 1) \tag{4.3.10}$$

$$\sigma_{a_2}^2 = \frac{12\sigma_y^2}{n(b-a)^2} = \frac{12\sigma_y^2}{n^3(\Delta x)^2} \tag{4.3.11}$$

$$\sigma_{12} = \frac{-12\sigma_y^2 r_x}{n(b-a)} = \frac{-12\sigma_y^2 r_x}{n^2\Delta x} \tag{4.3.12}$$

From these equations some obvious conclusions can be made. If the purpose of the experiment is to accurately measure a_1 then we see from Equation 4.3.10 that the best results are obtained when r_x=0 therefore we should choose a range in which the average value of x=0. If the values of x must be positive then the minimum value of r_x is 1/2 and the predicted minimum value of $\sigma_{a_1}^2$ is:

$$\left(\sigma_{a_1}^2\right)_{min} = \frac{\sigma_y^2}{n}\left(\frac{12}{4} + 1\right) = \frac{4\sigma_y^2}{n} \quad \text{for all } x > 0.$$

It should be remembered that $\sigma_{a_1}^2$ can be less than this value in a real experiment. This value is a design minimum based upon the assumptions that the value of $S / n - p$ is one, the value of x_1 is close to zero and the values of x are equally spaced.

If the purpose of the experiment is to measure the slope of the line, then from Equation 4.3.11 it is clear that the value of r_x is irrelevant. The range of the values of x (i.e., $b - a = n\Delta x$) is the important parameter. Stated in a

different way, if n is a fixed number (i.e., the experiment must be based on a fixed number of data points), then the accuracy of the slope (i.e., a_2) is improved by increasing Δx as much as possible. The value of x_1 makes no difference when all we are interested in is the measurement of a_2.

Assume that the purpose of the straight line experiment is to determine a line that can be used to estimate the value of y for any value of x within the range of the values of x (i.e., from x_1 to x_n). The predicted value of y is computed using Equation 4.2.2. For the straight line experiment, this equation takes the following simple form:

$$\sigma_f^2 = C_{11}^{-1} + 2xC_{12}^{-1} + x^2C_{22}^{-1} = \sigma_{a_1}^2 + 2x\sigma_{12} + x^2\sigma_{a_2}^2 \qquad (4.3.13)$$

Substituting Equations 4.3.10 to 4.3.12 into 4.3.13:

$$\sigma_f^2 = \frac{12\sigma_y^2}{n}\left(r_x^2 + 1 - \frac{2r_x x}{(b-a)} + \frac{x^2}{(b-a)^2}\right) \qquad (4.3.14)$$

This equation is a parabola in the x direction and we can find the point at which the parabola has its minimum value by setting the derivative with respect to x equal to zero:

$$\frac{-2r_x}{(b-a)} + \frac{2x}{(b-a)^2} = 0 \quad \text{therefore} \quad x = (b-a)r_x = \frac{(b+a)}{2}$$

This result is not surprising: the most accurate prediction (i.e., minimum value of σ_f^2) is achieved at the midpoint of the range. And clearly, the most inaccurate predictions are made at the extreme ends of the range: (i.e., at x_1 and x_n).

Example 4.3.1:

Assume we wish to design an experiment in which the dependent variable y is related to x by Equation 4.3.1 (i.e., a straight line). The values of x are limited to the range 1 to 10 and the purpose of the experiment is to determine the value of a_1 to ±0.01 (i.e., $\sigma_{a_1} = 0.01$). This parameter is the extrapolated value of y at $x = 0$ as seen in Figure 4.3.1. Assume that the values of x are equally spaced with $x_1 = 1$ and $x_n = 10$. The range $b-a$ required in the computation of r_x in Equation 4.3.10 is slightly greater than 9 due to the addition of $\Delta x/2$ at each end of the range. Thus $b-a = 9+\Delta x = 9n / (n-1)$. The two design variables are n and σ_y. Results for several values of n are included in Table 4.3.1.

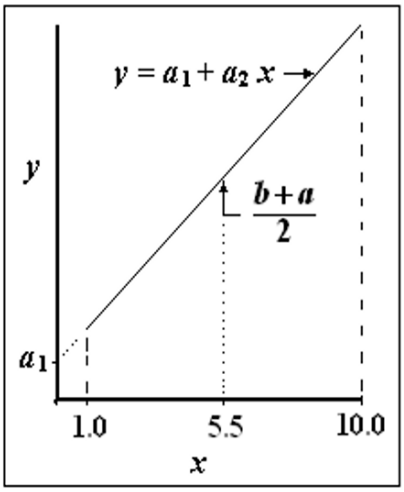

Figure 4.3.1 Straight line experiment to measure a_1

n	$b - a$	r_x	σ_{a1} / σ_y	σ_y
10	10.00	0.550	0.463	0.0147
20	9.47	0.581	0.502	0.0199
50	9.18	0.599	0.326	0.0307
100	9.09	0.605	0.232	0.0431

Table 4.3.1 Computed values of σ_y to achieve $\sigma_{a_1} = 0.01$

We can see from this table that if we can measure the values of y to an accuracy of 0.02 we will need about 20 data points to obtain an expected value of $\sigma_{a_1} = 0.01$. However, if we can only measure the values of y to

an accuracy of 0.04 we will need between and 50 and 100 data points. In other words, as the accuracy of the individual measurements of y decreases (i.e., σ_y increases), the number of data points required to obtain the desired accuracy (i.e., $\sigma_{a_1} = 0.01$) increases.

Example 4.3.2:

For the same experiment we would like to use the line for making predictions of the value of y for any point on the line. What combinations of n and σ_y are required to achieve values of $\sigma_f = 0.02$ at all points along the line? In the discussion following Equation 4.3.14 we see that the maximum values of σ_f occur at the extreme values of x (i.e., 1 and 10).

n	$b - a$	r_x	σ_f / σ_y	σ_y
10	10.00	0.550	1.201	0.0166
20	9.47	0.581	0.858	0.0233
50	9.18	0.599	0.545	0.0367
100	9.09	0.605	0.387	0.0517

Table 4.3.2 Computed values of σ_y to achieve $\sigma_f = 0.02$

The values of σ_f/σ_y in the Table 4.3.2 were computed using Equation 4.3.14 and either x =1 or x = 10. (The values are the same for both x = 1 and 10.) We see from the results that to achieve values of σ_f <= 0.02 at both ends of the line (and indeed, for all points within this range), the accuracy requirement for the individual points decreases as n increases. For example, if the values of y are measured to σ_y=0.02, then between 10 and 20 points are needed. However, if we can only measure the points to an accuracy of σ_y=0.05, then we will need close to 100 points.

4.4 Prediction Analysis of an Exponential Experiment

In this section we analyze a proposed experiment in which the dependent variable y is related to the independent variable x by an exponential function:

$$y = f(x) = a_1 e^{-a_2 x} \tag{4.4.1}$$

Typically, the x variable represents **time** and the y variable represents the number of events happening within the time-segment Δx centered at x. For example, this function can be used to analyze the decay of a radioactive isotope. In Section 4.3 the straight line experiment (i.e., Equation 4.3.1) was analyzed. Equation 4.4.1 provides another example of an experiment that can be analyzed analytically. We again start by proposing an experiment. Let us assume the following:

1. The values of x are obtained with negligible uncertainty (i.e., assume $\sigma_x = 0$).

2. Assume that the points are equally spaced (i.e., Δx constant).

3. Assume that the number of data points is n.

4. All values of y are Poisson distributed and are measured to an accuracy characterized by $\sigma_y = (Ky)^{1/2}$ therefore $w_i = 1/\sigma_y^2 = 1/(Ky)$.

The need for the constant K is required to maintain consistency of the units. If y is measured in **counts** (per time slice Δx) then σ_y and a_1 are also expressed in the same units. The units of K would thus be **counts** and the value of K would be one. However, if y are recorded in **counts** per time unit (i.e., **counts** / **second**) then $K = 1 / \Delta x$. The units of K would be the same as y. For example, if 100 counts are recorded in a time slice of 0.01 seconds, then the count rate y would be 100 / 0.01 = 10000 cps and $\sigma_y = (Ky)^{1/2} = (10000/0.01)^{1/2} = 1000$ cps.

From Equation 3.4.14 we can develop expression for the terms of the C matrix:

$$C_{11} = \sum_{i=1}^{i=n} w_i \frac{\partial f}{\partial a_1}\frac{\partial f}{\partial a_1} = \sum_{i=1}^{i=n} \frac{1}{Ky_i}(e^{-a_2 x_i})^2 = \frac{1}{Ka_1}\sum_{i=1}^{i=n} e^{-a_2 x_i} \qquad (4.4.2)$$

$$C_{22} = \sum_{i=1}^{i=n} \frac{1}{Ky_i}(-a_1 x_i e^{-a_2 x_i})^2 = \frac{a_1}{K}\sum_{i=1}^{i=n} x_i^2 e^{-a_2 x_i} \qquad (4.4.3)$$

$$C_{12} = \sum_{i=1}^{i=n} \frac{-1}{Ky_i} a_1 x_i (e^{-a_2 x_i})^2 = -\frac{1}{K}\sum_{i=1}^{i=n} x_i e^{-a_2 x_i} \qquad (4.4.4)$$

If we define a new dimensionless variable $z \equiv a_2 x$ the $\boldsymbol{C}$ matrix can be expressed as follows:

$$\boldsymbol{C} = \frac{n}{K}\begin{bmatrix} \frac{1}{a_1}(e^{-z})_{avg} & -\frac{1}{a_2}(ze^{-z})_{avg} \\ -\frac{1}{a_2}(ze^{-z})_{avg} & \frac{a_1}{a_2^2}(z^2e^{-z})_{avg} \end{bmatrix} \tag{4.4.5}$$

Inverting the matrix:

$$\boldsymbol{C}^{-1} = \frac{Ka_2^2}{nD}\begin{bmatrix} \frac{a_1}{a_2^2}(z^2e^{-z})_{avg} & \frac{1}{a_2}(ze^{-z})_{avg} \\ \frac{1}{a_2}(ze^{-z})_{avg} & \frac{1}{a_1}(e^{-z})_{avg} \end{bmatrix} \tag{4.4.6}$$

Where D is computed as follows:

$$\boldsymbol{D} = [(z^2e^{-z})_{avg}(e^{-z})_{avg} - ((ze^{-z})_{avg})^2] \tag{4.4.7}$$

The average values are only dependent upon the values of z_a and z_b where $z_a = a_2(x_1 - \Delta x/2)$ and $z_b = a_2(x_n + \Delta x/2)$. For example, if $n = 10$ and the values of x are 0.5, 1.5, 2.5, .. 9.5 (i.e., $\Delta x = 1$) and $a_2 = 0.2$ then $z_a = 0$ and $z_b = 2$. If we choose our x values so that $z_a = 0$, then z runs from 0 to a maximum value of $z = z_b = a_2 n \Delta x$. We can develop simple expressions for the averages required in Equations 4.4.6 and 4.4.7.

$$(e^{-z})_{avg} = \frac{1}{z}\int_0^z e^{-z}dz = \frac{1-e^{-z}}{z} \quad \text{from 0 to } z \tag{4.4.8}$$

$$(ze^{-z})_{avg} = \frac{1-(z+1)e^{-z}}{z} \quad \text{from 0 to } z \tag{4.4.9}$$

$$(z^2e^{-z})_{avg} = \frac{2-(z^2+2z+2)e^{-z}}{z} \quad \text{from 0 to } z \qquad (4.4.10)$$

These equations assume that the value of n is large. Values of Equations 4.4.7 through 4.4.10 are included in Table 4.4.1 for various values of z. For most experiments based upon Equation 4.4.1, the purpose of the experiment is to measure a_2. The predicted value of the variance of this measurement can be computed using Equation 4.4.6 and then Equations 4.4.7 through 4.4.10 for the average values:

$$\sigma_{a_2}^2 = C_{22}^{-1} = \frac{Ka_2^2}{na_1} \frac{(e^{-z})_{avg}}{[(z^2e^{-z})_{avg}(e^{-z})_{avg} - ((ze^{-z})_{avg})^2]} \qquad (4.4.11)$$

z	$(e^{-z})_{avg}$	$(ze^{-z})_{avg}$	$(z^2e^{-z})_{avg}$	D	$(e^{-z})_{avg}/D$
1.0	0.6321	0.2642	0.1606	0.0317	19.94
2.0	0.4323	0.2970	0.3233	0.0516	8.38
3.0	0.3167	0.2670	0.3845	0.0505	6.27
4.0	0.2454	0.2271	0.3809	0.0419	5.86
5.0	0.1987	0.1919	0.3501	0.0327	6.07
6.0	0.1663	0.1638	0.3127	0.0252	6.61
7.0	0.1427	0.1418	0.2772	0.0195	7.33
8.0	0.1250	0.1246	0.2466	0.0153	8.18
9.0	0.1111	0.1110	0.2208	0.0122	9.09
10.0	0.1000	0.1000	0.1994	0.0100	10.05

Table 4.4.1 Asymptotic Values (large n) Computed for Ranges 0 to z using Equations 4.4.7 through 4.4.10

Example 4.4.1:

Assume that we want to design an experiment based upon Equation 4.4.1 to determine a_2 to 1% accuracy. If the approximate value of a_2 is 0.2 sec^{-1}, what combinations of n, Δx and a_1 will best satisfy the accuracy requirement of the proposed experiment? From Equation 4.4.11 and the last column in Table 4.4.1 we see that the best choice for z (i.e., $a_2 n \Delta x$) is about 4. This value minimizes σ_{a2}. Thus the choice of n and Δx should satisfy

the following relationship: $a_2 n \Delta x = 4$. Since a_2 is approximately 0.2, $n\Delta x$ should be about 20 seconds. Using 4.4.11 with K=1 we can solve for $n\ a_1$:

$$\frac{\sigma_{a_2}^2}{a_2^2} = 0.01^2 = \frac{5.86}{na_1} \quad \text{therefore} \quad na_1 = \frac{5.86}{0.01^2} = 58600 .$$

The units of a_1 are counts/Δx seconds. For example, if n = 100, a_1 should be 586/ Δx and Δx should be about 20/100 = 0.2. The design count rate a_1 is therefore 586/0.2 = 2930 cps (counts per second).

Example 4.4.2:

Reconsider the experiment discussed in Example 4.4.1. Using the same values of n, Δx, a_1 and a_2 what is predicted value of σ_{a1}? From Equation 4.4.6 we can develop an equation similar to 4.4.11 for σ_{a1}:

$$\sigma_{a_1}^2 = C_{11}^{-1} = \frac{Ka_1}{n} \frac{(z^2 e^{-z})_{avg}}{[(z^2 e^{-z})_{avg}(e^{-z})_{avg} - ((ze^{-z})_{avg})^2]} \qquad (4.4.12)$$

Using the value z = 4, from Table 4.4.1 $\sigma_{a_1}^2$ = 586 * 0.3809 / (100 * 0.0419) = 53.27 and therefore σ_{a1} = 7.30 counts/0.2 seconds = 36.5 cps.

Example 4.4.3:

For the experiment discussed in Example 4.4.1 develop an equation for predicting the value of σ_f as a function of x using the same values of n, Δx, a_1 and a_2. We start with Equation 4.2.2 and substitute the expressions for the two partial derivatives of Equation 4.4.1:

$$\sigma_f^2 = e^{-2a_2 x}\sigma_{a_1}^2 - 2xa_1 e^{-2a_2 x}\sigma_{12} + x^2 a_1^2 e^{-2a_2 x}\sigma_{a_2}^2 \qquad (4.4.13)$$

The values of $\sigma_{a_1}^2$ and $\sigma_{a_2}^2$ are 53.27 and 0.000004. The value of σ_{12} is computed as follows:

$$\sigma_{12} = C_{12}^{-1} = \frac{Ka_2}{n} \frac{(ze^{-z})_{avg}}{[(z^2e^{-z})_{avg}(e^{-z})_{avg} - ((ze^{-z})_{avg})^2]} \quad (4.4.14)$$

Using the results in Table 4.4.1 for z = 4, n=100, a_2=0.2 and K=1 the predicted value of σ_{12} = 0.01084. Substituting into Equation 4.4.13 we obtain the following expression that can be used to predict σ_f :

$$\sigma_f^2 = e^{-0.4x}(53.276 - 12.705x + 1.3736x^2) \quad (4.4.15)$$

For example, at the midpoint of the range (i.e., x=10), the value of σ_f^2 is 0.0183*(53.27 - 127.05 + 137.36) = 1.164 therefore σ_f = 1.078. The value of y at this point is $586e^{-0.2*10}$ = 79.3 so the relative uncertainty of the predicted value of y at x=10 is 1.078 / 79.3 = 1.36%. Note that for a data point at x=10 the relative uncertainty would be approximately 1/*sqrt*(79.3) = 11%. **Thus the uncertainties associated with predicted values of y on the curve are much less than the uncertainties of the individual points used to fit the curve.**

In this section experiments based upon a single exponential function (i.e., Equation 4.4.1) were considered. In Section 4.6 a more complicated model that includes a constant background term (i.e., Equation 4.6.1) is discussed. In Section 5.2 the analysis is based upon separation experiments in which the mathematical model includes two exponential terms (i.e., Equation 5.2.1).

4.5 Dimensionless Groups

Analysis of a proposed experiment presents a problem: how should the results be displayed? The method of prediction analysis allows us to predict the accuracy that can be expected for a particular set of experimental variables but if the number of variables is more than a few, we need to simplify the presentation by combining the variables into manageable dimensionless groups. The experiment discussed in Section 4.4 (the exponential experiment) is used to illustrate this process.

The exponential experiment was analyzed analytically and equations for predicting accuracies (i.e., σ_{a1}, σ_{a2} and σ_f) were developed. Assuming that the x values were evenly spaced, these σ's turned out to be functions of n, Δx, a_1 and a_2. Further, it was assumed that the range of values of x started from $x = \Delta x/2$ and ended at $x = n\Delta x - \Delta x/2$. (In other words, data collection starts at 0 and ends at $n\Delta x$.) The dimensionless group z was defined as $a_2 n \Delta x$. Equations 4.4.11 and 4.4.12 can be reformulated to present the results for σ_{a1} and σ_{a2} in a dimensionless manner:

$$\Phi_1 = \sigma_{a_1}\left[\frac{n}{Ka_1}\right]^{1/2} = \left[\frac{(z^2 e^{-z})_{avg}}{D}\right]^{1/2} \tag{4.5.1}$$

$$\Phi_2 = \frac{\sigma_{a_2}(na_1)^{1/2}}{K^{1/2}a_2} = \left[\frac{(e^{-z})_{avg}}{D}\right]^{1/2} \tag{4.5.2}$$

Values of Φ_1 and Φ_2 are included in Table 4.5.1 for various values of z.

z	$(e^{-z})_{avg}$	$(z^2 e^{-z})_{avg}$	D	Φ_1	Φ_2
1.0	0.6321	0.1606	0.0317	2.25	4.47
2.0	0.4323	0.3233	0.0516	2.50	2.89
3.0	0.3167	0.3845	0.0505	2.76	2.50
4.0	0.2454	0.3809	0.0419	3.01	2.42
5.0	0.1987	0.3501	0.0327	3.27	2.46
6.0	0.1663	0.3127	0.0252	3.52	2.57
7.0	0.1427	0.2772	0.0195	3.77	2.71
8.0	0.1250	0.2466	0.0153	4.01	2.86
9.0	0.1111	0.2208	0.0122	4.25	3.01
10.0	0.1000	0.1994	0.0100	4.46	3.17

Table 4.5.1 Values of Φ_1 and Φ_2 for Ranges 0 to z (Eq 4.5.1 & 4.5.2)

A dimensionless group that characterizes σ_f can be developed from Equation 4.4.13. Substituting Equations 4.4.11, 4.4.12 and 4.4.14 into 4.4.13 we obtain the following:

$$\sigma_f^2 = \frac{Ka_1}{n} e^{-2a_2 x}\left(f_1(z) - 2a_2 x f_{12}(z) + (a_2 x)^2 f_2(z)\right) \tag{4.5.3}$$

In this equation the f's are functions of z. We thus see that the following dimensionless group Φ_f is a function of two dimensionless groups:

$$\Phi_f = \sigma_f \left[\frac{n}{K a_1} \right]^{1/2} = function(z, a_2 x) \tag{4.5.4}$$

The dimensionless group a_2x runs from 0 to z. Values of Φ_f are presented graphically in Figure 4.5.1 for $z = 4$. We see in the graph that σ_f decreases by more than a factor of 10 over the range of 0 to 4. However, the relative uncertainty (i.e., σ_f / y) actually decreases, reaches a minimum and then increases. We can show this by dividing Equation 4.5.3 by y^2, gathering terms and then taking the square root:

$$\Psi_f = \frac{\sigma_f (a_1 n)^{1/2}}{K^{1/2} y} = \left(f_1(z) - 2 a_2 x f_{12}(z) + (a_2 x)^2 f_2(z) \right)^{1/2} \tag{4.5.5}$$

We see from this equation that Ψ_f is the square root of a parabola along the a_2x axis. A plot of Ψ_f versus a_2x for $z = 4$ is shown in Figure 4.5.2.

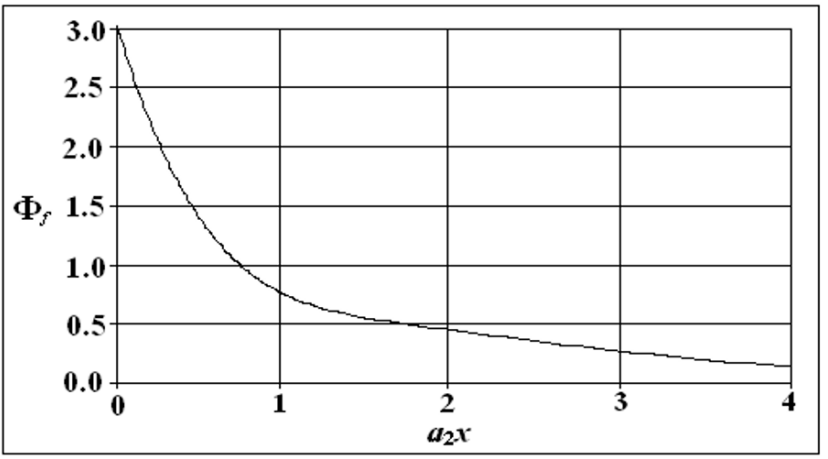

Figure 4.5.1 Φ_f versus a_2x for $z = 4$

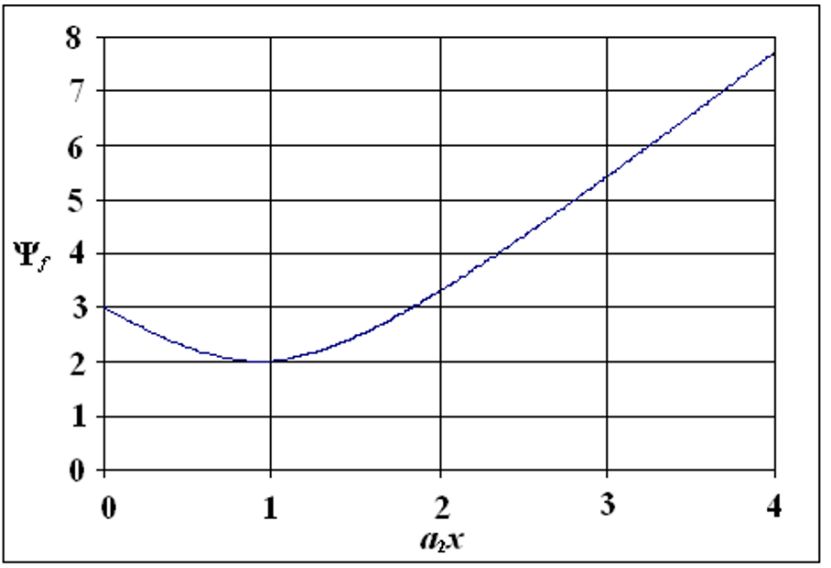

Figure 4.5.2 Ψ_f versus a_2x for $z = 4$

4.6 Simulating Experiments

Sections 4.3, 4.4 and 4.5 analyzed two very simple experiments in which the mathematical models allowed derivations of equations that could be used to predict accuracy (i.e., σ 's). We use the term "simple" in the sense that the mathematical model is simple to analyze both before and after data is collected. In reality, the actual experiment might be extremely complex and costly when one is faced with the task of running the experiment and collecting the data.

The simplicity of an experiment (from the analysis point-of-view) is measured to a large extent by the complexity of the mathematical model that will be applied to the data. For example, consider the class of experiments analyzed in Section 4.4 and 4.5 (i.e., the exponential experiment with two unknown parameter: Equation 4.4.1). If we add a background term to the equation and treat the model as containing three unknowns the task of developing an analytical solution becomes much more difficult. Consider the following equation:

$$y = f(x) = a_1 e^{-a_2 x} + a_3 \tag{4.6.1}$$

The $\boldsymbol{C}$ matrix has nine terms and inversion of the matrix leads to some very messy equations indeed. A simple approach is to avoid trying to seek analytical solutions for the terms of the $\boldsymbol{C}^{-1}$ matrix and go directly to solutions based upon simulations of the proposed experiment. What is required is simulation software. Such software can be built for a given experiment using readily available software (like MATLAB). Alternatively, a general purpose simulation tool is available in the public domain. Simulation is a feature included in the REGRESS program [WO06]. The REGRESS program is available for general purpose least squares analysis but can also be used in the prediction analysis mode.

Building upon the results in Section 4.5 three dimensionless groups were defined: $\boldsymbol{\Phi_1}$, $\boldsymbol{\Phi_2}$ and $\boldsymbol{\Phi_f}$ (Equations 4.5.1, 4.5.2 and 4.5.4). The first two are functions of $\boldsymbol{z}$ (i.e., $\boldsymbol{a_2 n \Delta x}$) and the third is a function of $\boldsymbol{z}$ and $\boldsymbol{a_2 x}$. If we add a third parameter to the model (i.e., $\boldsymbol{a_3}$) we will need an additional dimensionless parameter such as $\boldsymbol{a_3}$ / $\boldsymbol{a_1}$. The units for $\boldsymbol{a_1}$, $\boldsymbol{a_3}$, $\boldsymbol{y}$ and $\boldsymbol{K}$ are all in counts per $\boldsymbol{\Delta x}$ **so** $\boldsymbol{K = 1.}$ Examining Equation 4.5.1, if we choose values of $\boldsymbol{a_1} = \boldsymbol{n}$, then the computed value of $\boldsymbol{\sigma_{a1}}$ is equal to $\boldsymbol{\Phi_1}$. If we choose $\boldsymbol{a_2} = 1$ then $\boldsymbol{\Phi_2} = \boldsymbol{\sigma_{a2}}\, \boldsymbol{n}$. We first select values of $\boldsymbol{n}$ and $\boldsymbol{z}$ and then we can compute the initial value of $\boldsymbol{x}$ (i.e., $\boldsymbol{x_1}$) and $\boldsymbol{\Delta x}$. For example, if $\boldsymbol{a_2} = 1$ and $\boldsymbol{z} = 4$, then $\boldsymbol{n}\, \boldsymbol{\Delta x} = 4$. If we choose $\boldsymbol{n} = 1000$ then $\boldsymbol{\Delta x} = 0.004$ and $\boldsymbol{x_1} = 0.002$. (Remember that the range is from $\boldsymbol{x_1} - \boldsymbol{\Delta x}/2$ to $\boldsymbol{x_n} + \boldsymbol{\Delta x}/2$ which for this example is from 0 to 4.) A REGRESS simulation of this example using Equation 4.6.1 as the model and $\boldsymbol{a_3} = 0$ is shown in Figure 4.6.1.

```
PARAMETERS USED IN REGRESS ANALYSIS:
Thu Nov 27 14:22:20 2008

  INPUT PARMS FILE: fig461.par
  INPUT DATA  FILE: fig461.par
  REGRESS  VERSION: 4.21, Mar 31, 2008

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  :  1000
  Y VALUES - computed using A0 & X values
  SYTYPE   - Sigma type for Y             :     4
     TYPE 4: SIGMA Y = CY1 * sqrt(Y)   CY1: 1.000
  T Values - computed using interpolation table
     x01=0.002 np1=1000 dx1=0.004
  SXTYPE - Sigma type for X               :     0
     TYPE 0:  SIGMA X = 0

Analysis for Set 1
  Function Y:  A1*EXP(-A2*X)+ A3

   K       A0(K)         A(K)  PRED_SA(K)
   1     1000.00      1000.00     3.20629
   2     1.00000      1.00000     0.00462
   3     0.00000      0.00000     0.62460
```

Figure 4.6.1 - Simulation of an experiment using Eq. 4.6.1 for $z = 4$

In the results table in Figure 4.6.1 the column PRED_SA(K) is the predicted value of σ_{ak}. Plugging these values into equations 4.5.1 and 4.5.2 the computed values of $\Phi_1 = 3.21$ and $\Phi_2 = 4.62$ are obtained. We can compare the results from this simulation with similar results using the two parameter model (i.e., Equation 4.4.1). From Table 4.5.1 we see that for z=4 the values of $\Phi_1 = 3.01$ and $\Phi_2 = 2.42$ are noted. In other words, adding a third parameter to the analysis does increase the value of Φ_1 slightly but the effect upon Φ_2 is quite dramatic: an increase of almost a factor of 2 even when $a_3 = 0$. Clearly, if the purpose of the experiment is to measure a_2 as accurately as possible, one should attempt to use Equation 4.4.1 rather than 4.6.1 by measuring the background (i.e., a_3) separately.

In addition, we need a third dimensionless group that includes σ_{a3}:

$$\Phi_3 = \sigma_{a_3}\left[\frac{n}{Ka_1}\right]^{1/2} = function(z, a_3/a_1) \qquad (4.6.1)$$

The value of Φ_3 as seen in the simulation is 0.625. We have assumed that the dimensionless groups are insensitive to the value of n, a_1 and a_2. We can test this assumption by repeating the simulation using $z = 4$ and $a_2 = 1$ but changing n and a_1 to 100, $\Delta x = 0.04$ and $x_1 = 0.02$. The results are seen in Figure 4.6.2.

```
PARAMETERS USED IN REGRESS ANALYSIS:
Thu Nov 27 15:14:28 2008

  INPUT PARMS FILE: fig462.par
  INPUT DATA  FILE: fig462.par
  REGRESS  VERSION: 4.21, Mar 31, 2008

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  :    100
  Y VALUES - computed using A0 & X values
  SYTYPE  - Sigma type for Y              :      4
     TYPE 4: SIGMA Y = CY1 * sqrt(Y)   CY1: 1.000
  X Values - computed using interpolation table
  SXTYPE  - Sigma type for X              :      0
     TYPE 0:  SIGMA X = 0

Analysis for Set 1
  Function Y:  A1*EXP(-A2*X)+ A3

   K        A0(K)          A(K)  PRED_SA(K)
   1    100.00000    100.00000     3.20712
   2      1.00000      1.00000     0.04617
   3      0.00000      0.00000     0.62475
```

Figure 4.6.2 – Similar to Fig 4.6.1 but using $n = a_1 = 100$.

From Equations 4.5.1, 4.5.2 and 4.6.1 the values of $\Phi_1 = 3.21$, $\Phi_2 = 4.62$ and $\Phi_3 = 0.625$ are obtained and are in agreement to 3 decimal places of accuracy in both simulations. If we reduce n even further to $n = 10$, an additional simulation yields values of $\Phi_1 = 3.29$, $\Phi_2 = 4.72$ and $\Phi_3 = 0.640$ which are only a few percent larger than the values for much larger values of n. In other words, the value of n used in simulations of this experiment has only a small effect upon the results.

To measure the effect of the dimensionless parameter a_3 / a_1 simulations were performed for $z = 4$ and the results are presented in Figure 4.6.3.

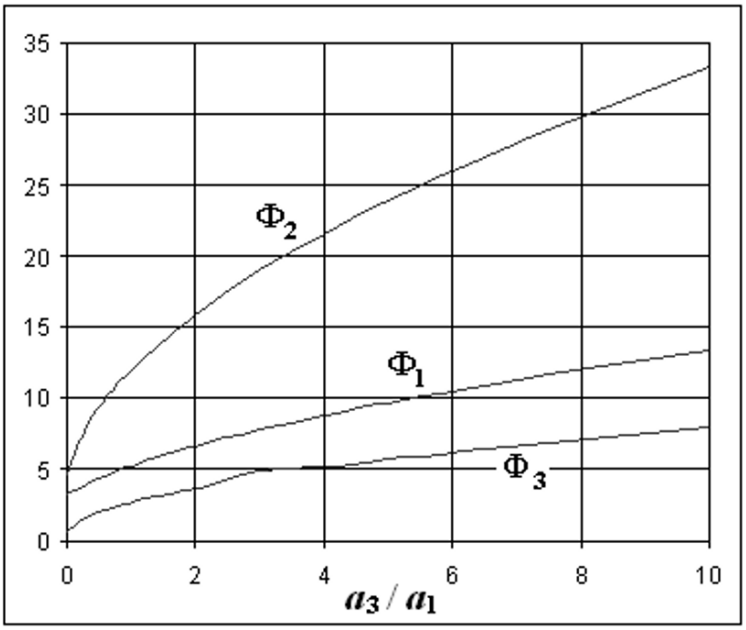

Figure 4.6.3 – Φ_1, Φ_2 and Φ_3 versus a_3 / a_1 for $z = 4$.

Example 4.6.1:

We wish to measure a_2 to 1% accuracy in a very noisy experiment in which the signal to noise ratio (i.e., a_1 / a_3) is approximately 1/8. What combinations of initial amplitude (i.e., $a_1 + a_3$ per Δx) and n are required so that the predicted values of $\sigma_{a2} / a_2 = 0.01$? Assume that the choice of values for x_1, Δx and n satisfy the relationship $z = 4$.

From Figure 4.6.3 for $a_3 / a_1 = 8$ we see that the value of Φ_2 is approximately 30. With $K = 1$, from the definition of Φ_2 (Equation 4.5.2) we can develop the following relationship:

$$\frac{a_2}{\sigma_{a_2}}\Phi_2 = 100 * 30 = (na_1/K)^{1/2} = (n(a_1 + a_3)/9)^{1/2}$$

$$n(a_1 + a_3) = 81{,}000{,}000$$

Example 4.6.2:

Example 4.6.1 is based upon the choice of $z = 4$. This value was chosen because in Table 4.5.1 it was seen that this choice of z minimizes Φ_2. However, Table 4.5.2 was based upon the two parameter model (i.e., Equation 4.4.1) and Example 4.6.1 is based upon the three parameter model (i.e., Equation 4.6.1). For the noisy experiment discussed in Example 4.6.1 is the choice of $z = 4$ still an optimum or close to an optimum choice for z?

To answer this question a series of simulations were performed in which the parameters a_1 and n were set equal to 100, a_3 was set to 800 and the values of x_1 and Δx were varied to cover the range $z = 3, 4 .. 8$. The value of Φ_2 for $z = 3$ was 35.1, and this value decreased to a minimum of 27.5 for $z = 6$. Thus for the proposed experiment the choice of $z = 4$ is not optimal but is not very far from optimal. Increasing z to 6 should reduce the predicted value σ_{a2} / a_2 by about 2.5/30 (i.e., 8%) compared to the value that would be obtained if the experiment were to be performed using $z = 4$.

4.7 Predicting Calculational Complexity

One can predict whether or not the least squares analysis phase of an experiment will be plagued by numerical problems. The ***Condition*** number of the ***C*** matrix is a number that is indicative of several issues:

1) The sensitivity of the results due to variations in the data.

2) The difficulty that one might encounter in converging to a solution for nonlinear models.

3) The numerical problems associated with ill-conditioned systems.

The ***Condition*** number is defined as the ratio of the maximum to minimum absolute values of the Eigenvalues of the ***C*** matrix. Since the ***C*** matrix is symmetric, all the Eigenvalues are real numbers but not necessarily positive. Eigenvalues are not easy to calculate but functions included in widely available software (like MATLAB and MAPLE) can be used to deter-

mine the ***Condition*** number of a matrix. The REGRESS program includes a parameter that directs the program to include the condition number of the ***C*** matrix in the program output.

To illustrate the effect of ***Condition*** number on the sensitivity of the results, consider the following matrix equation:

$$\boldsymbol{CA} = \boldsymbol{V} = \begin{bmatrix} \boldsymbol{n}+2 & \boldsymbol{n}+1 \\ \boldsymbol{n}+1 & \boldsymbol{n} \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} 2\boldsymbol{n}+3 \\ 2\boldsymbol{n}+1 \end{bmatrix} \tag{4.7.1}$$

The solution to this equation is easily seen to be $\boldsymbol{a}_1 = \boldsymbol{a}_2 = 1$ for any value of $\boldsymbol{n}$. The condition of the ***C*** matrix is shown in Table 4.7.1 for values of $\boldsymbol{n}$ from 1 to 1000. The value of $\boldsymbol{v}_2$ is then perturbed from $2\boldsymbol{n}+1$ to $2\boldsymbol{n}+1.001$ and Equation 4.7.1 is solved for the terms of the perturbed ***A*** vector.

N	***Condition***	$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	*Perturbed* $\boldsymbol{a}_1$	*Perturbed* $\boldsymbol{a}_2$
1	17.9	1	1	1.002	0.997
10	486.0	1	1	1.011	0.988
20	1766.0	1	1	1.021	0.978
50	10496.0	1	1	1.051	0.948
100	40806.0	1	1	1.101	0.898
200	161606.0	1	1	1.201	0.798
500	1004006.0	1	1	1.501	0.490
1000	4008006.0	1	1	2.001	-0.002

Table 4.7.1 – Effect of perturbing $\boldsymbol{v}_2$ from $2\boldsymbol{n}+1$ to $2\boldsymbol{n}+1.001$

We see in this table that as the condition of the matrix becomes more and more ill-conditioned (i.e., larger), the effect of the small perturbation in the data causes an ever increasing perturbation in the computed value of the terms of the ***A*** vector. The sensitivity of the results for the case of $\boldsymbol{n} = 1000$ is particularly dramatic. The condition of the matrix is about 4 million and we see that a change in $\boldsymbol{v}_2$ from 2001 to 2001.001 (i.e., a change of 0.00005%) causes an increase of about 100% in $\boldsymbol{a}_1$ and a decrease of close to 100% in $\boldsymbol{a}_2$. Note that $\boldsymbol{c}_{21}\boldsymbol{a}_1 + \boldsymbol{c}_{22}\boldsymbol{a}_2 = 1001*2.001 - 1000*0.002$ is exactly 2001.001 so the reason for the large changes in $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are not attributable to numeric problems such as round-off errors. It is a fact of life that solving equations based upon ill-conditioned matrices will yield results that are highly sensitive to small changes in the data.

Converging to a solution can be extremely difficult for nonlinear problems in which the ***C*** matrix is highly ill-conditioned. The U.S. National Institute of Standards and Technology (NIST) initiated a project to develop a standard group of statistical reference datasets (StRD's) [WO06]. In their words the object of the project was "to improve the accuracy of statistical software by providing reference datasets with certified computational results that enable the objective evaluation of statistical software." One of the specific areas covered was datasets for nonlinear regression. The problems in the nonlinear regression group were graded as *Lower*, *Average* and *Higher* levels of difficulty. The Bennett5 problem includes a 3 parameter fit to 154 data points and is graded as highly difficult. This problem required over 536,000 iterations using the default settings of REGRESS and starting from the "far" (according to the NIST system) initial values. The fitting function for this problem is:

$$y = b_1 + (b_2 + x)^{-1/b_3} \tag{4.7.2}$$

The ***Condition*** number of the matrix using the NIST initial values is greater than 10^{17}. As seen in Figure 4.7.1 the value decreases to $>10^{14}$ after 3 iterations and after more than 536,000 settles at a value $>10^{16}$. Note that after an initial rapid decrease in the value of ***S*/(*n*-*p*)** (from 437 down to $7.36*10^{-6}$) it takes more than 500,000 additional iterations to achieve a further reduction to $3.47*10^{-6}$. This problem illustrates the extreme difficulty in achieving convergence for least squares problems that require solutions of highly ill-condition systems of equations. (Note – for those readers interested in trying to solve the Bennett5 problem using software other than REGRESS, the data and solutions can be obtained directly from the NIST website: http://www.itl.nist.gov/div898/strd/index.html. To examine the datasets, go into *Dataset Archives* and then *Nonlinear Regression*.)

The lesson to be learned from this example is that one should avoid least squares problems in which the ***C*** matrix will be highly ill-conditioned. The condition of the matrix can be estimated at the design stage, and if ill-conditioning is noticed the problem should be reformulated. For example, a different fitting function might be considered. Another possible approach is to choose different or additional values of the independent variable (or variables).

```
PARAMETERS USED IN REGRESS ANALYSIS: Wed Dec 24 15:26:05 2008

  INPUT PARMS FILE: bennett5.par
  INPUT DATA  FILE: bennett5.par
  REGRESS  VERSION: 4.22, Dec 24, 2008

  N - Number of recs used to build model  :   154
  YCOL1 - Column for dep var Y               :  1
  SYTYPE1 - Sigma type for Y              :     1
     TYPE 1:  SIGMA Y = 1
Analysis for Set 1
  Function Y:  B1 * (B2+X)^(-1/B3)

  EPS - Convergence criterion           : 0.00100
  CAF - Convergence acceleration factor :   1.000

 ITERATION         B1          B2          B3  S/(N.D.F.)  CONDITION
         0   -2000.00    50.00000    0.80000   437.23475    >10^17
         1 -696.26923    39.53721    1.14423    99.76370    >10^17
         2 -772.85887    33.14625    1.18089     1.62694    >10^15
         3 -790.25485    33.86791    1.19260  0.00028004    >10^14
         4 -791.34742    34.12556    1.19369  0.00000845    >10^14
         5 -791.60305    34.17606    1.19391  0.00000736    >10^14
         .
    536133   -2523.51    46.73656    0.93218  0.00000347    >10^16
```

Figure 4.7.1 – REGRESS output for the Bennett5 problem

To demonstrate the numerical problems associated with ill-conditioned systems, the NIST data sets include a data series called ***Filip.dat*** (in the Linear Regression directory). This data set was developed by A. Filippelli from NIST and includes 82 data points and a solution for an 11 parameter polynomial. For p parameters the function is:

$$y = \sum_{i=0}^{i=p-1} a_i x^i \tag{4.7.3}$$

The C matrix for this function is generated using Equation 3.4.14 and w_i=1:

$$C_{jk} = \sum_{i=1}^{i=n} w_i \frac{\partial f}{\partial a_j}\frac{\partial f}{\partial a_k} = \sum_{i=1}^{i=n} x^{j+k} \quad \text{for } j=0 \text{ to } p\text{-}1,\ k=0 \text{ to } p\text{-}1 \tag{4.7.4}$$

As p increases, the ***Condition*** of the C matrix explodes. It is well known that fitting higher degree polynomials of the form of Equation 4.7.3 to data is a poor idea. If one really would like to find a polynomial fit to the data, then the alternative is to use orthogonal polynomials [WO06, FO57]. How-

ever, trying to fit Equation 4.7.3 with increasing values of ***p*** illustrates the numerical problems inherent with ill-conditioned systems.

p	*Condition*	*VarReduction*	*Num Iter*	*Solution*
2	$7.52*10^{02}$	87.54	1	Yes
3	$4.89*10^{05}$	90.64	1	Yes
4	$5.05*10^{08}$	93.45	1	Yes
5	$4.86*10^{11}$	97.30	1	Yes
6	$6.28*10^{14}$	97.42	1	Yes
7	$7.67*10^{17}$	98.99	1	Yes
8	$1.57*10^{19}$	99.00	1	Yes
9	$8.18*10^{20}$	99.48	6	Yes
10	$1.04*10^{23}$	/	/	No
11	$1.46*10^{23}$	99.43	1142	Yes

Table 4.7.2 – Results for the NIST *filip* Data Set using Eq 4.7.3

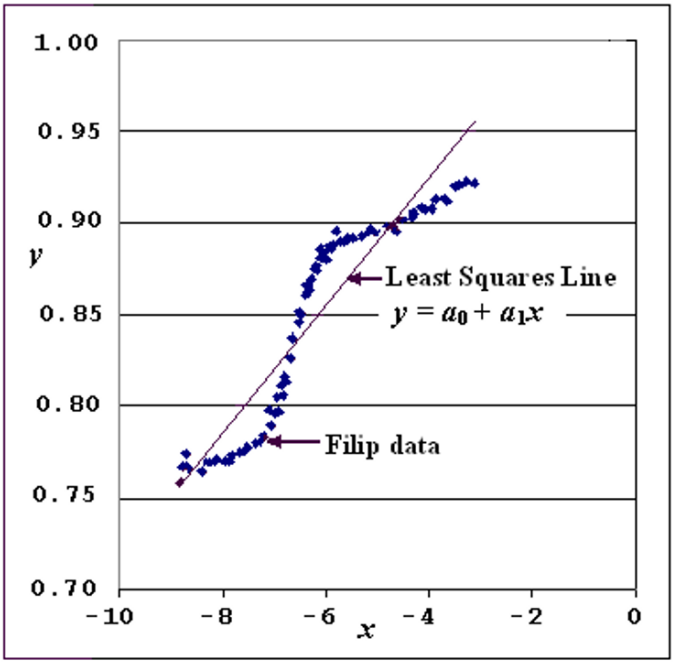

Figure 4.7.2 – Least Squares Line and *filip* Data Set

Table 4.7.2 includes results obtained using REGRESS for ***p*** = 2 to 11. For ***p*** = 9 we see the first signs of trouble brewing. Theoretically for linear

problems the least squares method should converge to a solution with a single iteration. Within each iteration, REGRESS checks the error vector (i.e., $\boldsymbol{CA - V}$) and iteratively corrects $\boldsymbol{A}$ until the error vector is smaller than a specified tolerance level or until 10 attempts have been made to reduce the error vector. For $\boldsymbol{p} = 9$ six iterations were required until the $\boldsymbol{A}$ vector satisfied the convergence criterion. Beyond $\boldsymbol{p} = 9$ nothing helped. We see that for $\boldsymbol{p} = 10$ REGRESS failed to obtain a solution. The search for a solution was halted after 10,000 iterations. For $\boldsymbol{p} = 11$, although convergence was achieved after many iterations, the values obtained were completely different than the reference values included with the data set. The huge values of ***Condition*** ($> 10^{23}$) were just too much for the program to handle. Note that the variance reduction increased from a value of 87.5% for $\boldsymbol{p} = 2$ to almost 99.5% for $\boldsymbol{p} = 9$. The large value of variance reduction using only a straight line (i.e., $\boldsymbol{p} = 2$) might seem surprising, however in Figure 4.7.2 we see that the data is not very far from the straight line obtained using the 2 parameter fit.

4.8 Predicting the Effects of Systematic Errors

In Section 1.6 the subject of systematic errors was discussed. Obvious sources of systematic errors are measurement errors. For some experiments the effect of a systematic error is quite obvious. For example, in Section 4.3 prediction analysis of a straight line experiment using Equation 4.3.1 was discussed. If there is a calibration error $\boldsymbol{\varepsilon_y}$ in the measurements of the values of $\boldsymbol{y}$ the effect upon the computed value of $\boldsymbol{a_1}$ will be just $\boldsymbol{\varepsilon_y}$ and the effect upon the computed value of $\boldsymbol{a_2}$ will be 0. The actual value of $\boldsymbol{\varepsilon_y}$ is unknown (otherwise we would correct for it) however, we can usually estimate a maximum absolute value for the systematic error. Similarly, if there is a calibration error in the measurement of $\boldsymbol{x}$ equal to $\boldsymbol{\varepsilon_x}$ the effect upon the computed value of $\boldsymbol{a_1}$ will be just $\boldsymbol{-\varepsilon_x a_2}$ and the effect upon the computed value of $\boldsymbol{a_2}$ will also be 0.

Even for simple experiments (from the analysis point-of-view), the nature of the systematic errors can lead to a complicated process of estimating the effects upon the computed parameters. For example, consider the exponential experiment discussed in Section 4.4. Equation 4.4.1 is an exponentially decaying function in which the background term is either negligible or the values of $\boldsymbol{y}$ have been corrected to remove the background. One way to approach such problems is through the use of simulations. In Example 4.4.1 an experiment was discussed in which the requirement was to

measure the decay constant a_2 to 1% accuracy. Using 100 points over a range of $z = a_2 n \Delta x = 4$ and a_2 equal to approximately 0.2, the design value for a_1 was computed to be equal to 586 counts per Δx seconds. Let us estimate that the error in the background count rate is in the range -10 to 10 cps. Since $\Delta x = 4/(0.2*100) = 0.2$, the maximum background error $\varepsilon_b = 10 * 0.2 = 2$ counts per Δx. To determine the effect of such an error, a simulation can be performed using the following steps.

1) Create a data series based upon 100 data points starting from $x = \Delta x/2 = 0.1$ and increasing in increments of 0.2 with corresponding y values computed using $y = 586*exp(-0.2*x)$.

2) Create a 2nd data series in which the values of y are increased by 2.

3) Run least squares analyses using both series and Equation 4.4.1.

4) The effect of ε_b upon both a_1 and a_2 is just the differences between the two results.

The results of the least squares analysis for the first data series should return values of $a_1 = 586$ and $a_2 = 0.2$ and confirm that the series has been created correctly. The least squares analysis for the 2nd data series is shown in Figure 4.8.1 and we note values of $a_1 = 582.7$ and $a_2 = 0.196$. The experiment had been designed for an accuracy of 1% in a_2 which is 0.002. We note, however, that if there is a 10 cps error in the determination of the background count rate, this causes an error of -0.004 in a_2. In other words, if our background count rate is truly in error by 10 cps then the resulting error in a_2 is twice as large as our design objective!

```
PARAMETERS USED IN REGRESS ANALYSIS: Wed Feb 11, 2009

  N - Number of recs used to build model  :   100
  YCOL1 - Column for dep var Y                :  3
  SYTYPE1 - Sigma type for Y                  :     4
     TYPE 4: SIGMA Y = CY1 * sqrt(Y)   CY1: 1.000

Analysis for Set 1
  Function Y:  A1*EXP(-A2*X)

   K      A0(K)       A(K)    SIGA(K)
   1    586.000    582.733      0.911
   2      0.200      0.196   0.000248
```

Figure 4.8.1 – Least Squares Analysis of Data with Background Error

4.9 P.A. with Uncertainty in the Independent Variables

In the examples shown to this point it was assumed that the values of $\boldsymbol{x}$ are known (i.e., $\boldsymbol{\sigma_x} = 0$). Equation 3.3.7 shows how one must weight the data if the assumption that $\boldsymbol{\sigma_x} = 0$ is not valid. For some problems this additional term requires only a minor modification. For example, for the straight line experiment with a constant value of $\boldsymbol{\sigma_y}$ for all data points, the resulting values of $\boldsymbol{\sigma_{a1}}$, $\boldsymbol{\sigma_{a2}}$ and $\boldsymbol{\sigma_{12}}$ are computed using Equations 4.3.10, 4.3.11 and 4.3.12. Note that a constant value of $\boldsymbol{\sigma_y}$ implies that there is a random error component for all the values of $\boldsymbol{y}$ with a mean of zero and a constant standard deviation of $\boldsymbol{\sigma_y}$. The weights for all the points were equal to $1 / \boldsymbol{\sigma_y^2}$ but if $\boldsymbol{\sigma_x}$ is a constant other than zero (i.e., a random error with mean zero and standard deviation of $\boldsymbol{\sigma_x}$), the weights would all equal the following:

$$w = \frac{1}{\sigma_y^2 + \left(\dfrac{\partial f}{\partial x}\sigma_x\right)^2} = \frac{1}{\sigma_y^2 + \left(a_2\sigma_x\right)^2} \tag{4.9.1}$$

The resulting modifications in 4.3.10 through 4.3.12 are:

$$\sigma_{a_1}^2 = \frac{12(\sigma_y^2 + a_2^2\sigma_x^2)}{n}(r_x^2 + 1) \tag{4.9.2}$$

$$\sigma_{a_2}^2 = \frac{12(\sigma_y^2 + a_2^2\sigma_x^2)}{n^3(\Delta x)^2} \tag{4.9.3}$$

$$\sigma_{12} = \frac{-12(\sigma_y^2 + a_2^2\sigma_x^2)r_x}{n^2\Delta x} \tag{4.9.4}$$

Alternatively the effect of non-zero values of σ_x can be determined using simulations. As an example of the use of a simulation to predict the effect of non-zero values of σ_x consider the exponential experiment discussed in Section 4.4. Example 4.4.1 considers an experiment in which the requirement was to measure the decay constant a_2 to 1% accuracy. Using 100 points over a range of $z = a_2 n \Delta x = 4$ and a_2 equal to approximately 0.2, the design value for a_1 was computed to be equal to 586 counts per Δx seconds. The value of Δx for this experiment was 4/(0.2*100) = 0.2. In Figure 4.9.1 a prediction analysis for this experiment in which $\sigma_x = 0.2$ (i.e., the same as Δx) is used. The results in Figure 4.9.1 should be compared to the results in Examples 4.4.1 and 4.4.2 (i.e., $\sigma_{a1} = 7.30$ and $\sigma_{a2} = 0.00200$). The effect of σ_x increasing from 0 to 0.2 causes an increase in σ_{a1} to 9.17 and σ_{a2} to 0.00224.

It should be mentioned that the examples used in this section are based upon a very simple model for σ_x (i.e., a constant value for all data points). In reality any desired model can be used in a simulation.

```
PARAMETERS USED IN REGRESS ANALYSIS: Wed Feb 11, 2009

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  :    100
  Y VALUES - computed using A0 & X values
  SYTYPE1 - Sigma type for Y              :      4
     TYPE 4: SIGMA Y = CY1 * sqrt(Y)   CY1: 1.000
  T Values - computed using interpolation table
  STTYPE1 - Sigma type for X1             :      2
     TYPE 2:  SIGMA X1 = CX1           CX1: 0.200

Analysis for Set 1
  Function Y:  A1*EXP(-A2*X)

   K        A0(K)         A(K)  PRED_SA(K)
   1   586.00000   586.00000     9.16842
   2     0.20000     0.20000     0.00224
```

Figure 4.9.1 – Prediction Analysis with non-zero σ_x

4.10 Multiple Linear Regression

In the previous sections the mathematical models discussed were based upon a single independent variable. Many experiments utilize models that require several independent variables. For example, thermodynamic experiments to measure the properties of a gas as a function of both temperature and pressure are typical. Health studies of children might consider height, weight and age as independent variables. To illustrate a prediction analysis of an experiment with more than one independent variable, the following simple model is used:

$$y = a_1 + a_2 x_1 + ... + a_{d+1} x_d \tag{4.10.1}$$

In the statistical literature the subject "multiple linear regression" is based upon this model [FR92,ME92]. In this equation a d dimensional plane is used to model y as a function of the values of a vector $\mathbf{X}$. In Section 4.3 this model was analyzed for the simplest case: $d = 1$. The analysis was based upon the following assumptions:

1. All values of y are measured to a given accuracy: σ_y.

2. The values of x are obtained with negligible uncertainty (i.e., assume $\sigma_x = 0$).

3. The number of data points is n.

4. The points are equally spaced (i.e., Δx is constant).

The only modification required to analyze 4.10.1 is that the data points are generated by taking n_1 values in the x_1 directions, n_2 values in the x_2 directions, etc. so that the total number of points is n which is the product of all the n_j values. Furthermore, the n_j values are generated as follows:

$$x_{ji} = x_{j1} + (i-1)\Delta x_j \qquad i = 1 \text{ to } n_j \tag{4.10.2}$$

The C matrix for this model is:

$$C = \begin{bmatrix} \sum_{i=1}^{i=n} w_i & \sum_{i=1}^{i=n} w_i x_{1i} & \cdots & \sum_{i=1}^{i=n} w_i x_{di} \\ \sum_{i=1}^{i=n} w_i x_{1i} & \sum_{i=1}^{i=n} w_i x_{1i}^2 & \cdots & \sum_{i=1}^{i=n} w_i x_{1i} x_{di} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^{i=n} w_i x_{di} & \sum_{i=1}^{i=n} w_i x_{1i} x_{di} & \cdots & \sum_{i=1}^{i=n} w_i x_{di}^2 \end{bmatrix} \tag{4.10.3}$$

Since w_i is constant for all points (i.e., $w_i = 1/\sigma_y^2$) it can be removed from the summations. In Section 4.3 we proceeded to develop an analytical solution for the predicted values of σ_{a1}^2 and σ_{a2}^2, the covariance σ_{12} and σ_f^2. At first glance the C matrix looks messy but by simply subtracting out the mean values of x_j from all the values of x_{ji}, all the off-diagonal terms of the C matrix become zero and the matrix can easily be inverted. We can use Equations 4.3.10, 4.3.11, 4.3.12 and 4.3.14 to anticipate the results that one should expect for d greater than one. In Section 4.3 the dimensionless midpoint of the range x was defined by Equation 4.3.9. We can define a more general r_j as follows:

$$r_j \equiv \frac{x_{javg}}{n_j \Delta x_j} = \frac{(b_j + a_j)/2}{b_j - a_j} \tag{4.10.4}$$

where $a_j = x_{j1} - \Delta x_j/2$ and $b_j = x_{jn} + \Delta x_j/2$. If the x_{ji}'s have been adjusted so that the average value in each direction is zero, then all the r_j's are zero. However, if the original values of the x_{ji}'s are used, then the r_j's have values other than zero. The results can be presented using the following dimensionless groups:

$$\Phi_1 \equiv \frac{\sigma_{a_1} n^{1/2}}{\sigma_y} \tag{4.10.5}$$

$$\Phi_{j+1} \equiv \frac{\sigma_{a_{j+1}} n^{1/2} (n_j \Delta x_j)}{\sigma_y} \qquad j = 1 \text{ to } d \tag{4.10.6}$$

The simplest case to analyze is when all the values of r_j are zero (i.e., the data is centered about zero in each direction). For this case all the off-diagonal terms of the C matrix are zero and the results reduce to the following for large n:

$$\Phi_1 = 1 \quad \& \quad \Phi_j = 12^{1/2} \tag{4.10.7}$$

These rather surprisingly simple results are independent of the dimensionality of the model (i.e., d) and are in agreement with the results from Section 4.3 for the case of d=1. Furthermore, varying the values of the r_j parameters only changes the value of Φ_1:

$$\Phi_1 \equiv \frac{\sigma_{a_1} n^{1/2}}{\sigma_y} = (12 \sum_{j=1}^{j=d} r_j^2 + 1)^{1/2} \tag{4.10.8}$$

Confirmation of these results is shown in Figure 4.10.1. This is a simulation of a case in which d =3 and all three values of r_j are equal to 1/2. Since σ_y = 1 and n = 10000, $\Phi_1 = 100\,\sigma_{a_1}$ = 3.172. From Equation 4.10.8

we get a value of $(12*(1/4+1/4+1/4)+1)^{1/2} = 10^{1/2} = 3.162$ which is close to 3.172. The values of $\boldsymbol{n_1\Delta x_1}$, $\boldsymbol{n_2\Delta x_2}$ and $\boldsymbol{n_3\Delta x_3}$ are 1, 2 and 1 and therefore the values of the remaining Φ_j's are :

$$\Phi_2 \equiv \frac{n^{1/2}(n_1\Delta x_1)\sigma_{a_2}}{\sigma_y} = \frac{100*1*0.03464}{1} = 3.464$$

$$\Phi_3 \equiv \frac{n^{1/2}(n_2\Delta x_2)\sigma_{a_3}}{\sigma_y} = \frac{100*2*0.01741}{1} = 3.482$$

$$\Phi_4 \equiv \frac{n^{1/2}(n_3\Delta x_3)\sigma_{a_4}}{\sigma_y} = \frac{100*1*0.03482}{1} = 3.482$$

From Equation 4.10.7 the limiting values (for large $\boldsymbol{n}$) of these Φ_j's are $12^{1/2} = 3.464$ which is exact to 4 decimal place for Φ_2 and close for Φ_3 and Φ_4. As expected, the simulation yields a more accurate result for Φ_2 because $\boldsymbol{n_1} = 100$ while $\boldsymbol{n_2}$ and $\boldsymbol{n_3}$ are only 10. The values of the Φ_j's are independent of the values of $\boldsymbol{a_1}$ to $\boldsymbol{a_4}$.

```
PARAMETERS USED IN REGRESS ANALYSIS: Tue Mar 17, 2009

  INPUT PARMS FILE: fig410_1.par
  INPUT DATA  FILE: fig410_1.par
  REGRESS  VERSION: 4.25, Mar 15, 2009

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  : 10000
  Y VALUES - computed using A0 & X values
  SYTYPE1 - Sigma type for Y              :     1
     TYPE 1:  SIGMA Y = 1
  M - Number of independent variables     :     3
  X1 Values - computed using interpolation table
    NP1=100  X01=0.005  DELX1=0.01
  X2 Values - computed using interpolation table
    NP2=10   X02=0.1    DELX2=0.2
  X3 Values - computed using interpolation table
    NP3=10   X03=0.05   DELX3=0.1

Analysis for Set 1
  Function Y:  A1+A2*X1+A3*X2+A4*X3

   K       A0(K)        A(K)  PRED_SA(K)
   1     1.00000     1.00000     0.03172
   2     2.00000     2.00000     0.03464
   3     3.00000     3.00000     0.01741
   4     0.00000     0.00000     0.03482
```

Figure 4.10.1 – Prediction Analysis of 3D plane model

For some experiments based upon Equation 4.10.1 the purpose is to determine a plane that can be used to estimate the value of y for any set of values of x_j within the range of the values of the x_j's. For the straight line experiment (i.e., d = 1) in Section 4.3 we obtained a parabola (Equation 4.3.14). The accuracy of the predicted value of y was seen to be best at the midpoint of the range and worst at the edge points. For d greater than one we would expect similar results: however, the parabola would be replaced by a d dimensional multinomial of degree 2. For example, for d=2 we would obtain an equation of the form:

$$\sigma_f^2 = K_0 + K_1 x_1 + K_2 x_2 + K_3 x_1^2 + K_4 x_2^2 + K_5 x_1 x_2 \qquad (4.10.9)$$

If all the x_j's have been adjusted so that their mean values are zero, Equation 4.10.9 is reduced to a simpler form as seen in Equation 4.10.10. The minimum value of σ_f will be obtained at the midpoint of the ranges of the x_j 's and the maximum values will be obtained at the corners of the d dimensional cube. For a given value of d it can be shown that the value of

σ_f^2 can be computed for any point within the d dimensional cube from the following equation:

$$\frac{\sigma_f^2 n}{\sigma_y^2} = 1 + \sum_{j=1}^{j=d} 12 \frac{(x_j - x_{javg})^2}{(n_j \Delta x_j)^2} \tag{4.10.10}$$

The derivation of this equation is simplified by considering the case where all the average values of x_j are zero. We can, of course, transform any set of experimental data to just this case by subtracting x_{javg} from all the values of x_j. For cases satisfying this condition, it can be shown that the inverse C matrix is:

$$C^{-1} = \frac{\sigma_y^2}{n} \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 12/(n_1 \Delta x_1) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 12/(n_d \Delta x_d) \end{bmatrix} \tag{4.10.11}$$

Recall that n is the product of all the n_j's. Equation 4.10.10 follows from substitution of Equation 4.10.11 into Equation 4.2.2. The best accuracy (minimum value of σ_f) is obtained at the midpoint of the d dimensional cube (i.e., where $x_j = x_{javg}$ for all j). The worst accuracy (maximum value of σ_f) is obtained at the corners of the d dimensional cube. Note that these minimum and maximum values of σ_f are not functions of the values of the r_j's and $n_j \Delta x_j$'s. The minimum value is not a function of d and the maximum value of σ_f^2 increases linearly with d.

$$\frac{\sigma_{f\,min}^2 n}{\sigma_y^2} = 1 \tag{4.10.12}$$

$$\frac{\sigma_{f\,max}^2 n}{\sigma_y^2} = 1 + \sum_{j=1}^{j=d} 12(1/2)^2 = 1 + 3d \tag{4.10.13}$$

For example, if d = 4, the predicted values of σ_f range from $\sigma_y/n^{1/2}$ to $13^{1/2}\sigma_y/n^{1/2}$.

Example 4.10.1:

Assume that we wish to perform a multiple linear regression within a 3 dimensional space. Assume that Equation 4.10.1 is a realistic representation of the dependence of the values of y upon the values of x_1, x_2 and x_3. Assume that we can measure each value of y to an accuracy of σ_y and that we can assume that the values of the x's are error free. The purpose of the experiment is to predict values of y for any point within the space. We want to design the experiment so that within the experimental volume all values of y will be computed to an accuracy of 0.01 (i.e., $\sigma_f <= 0.01$).

From Equation 4.10.13, within the design region, the maximum values of σ_f are :

$$\frac{\sigma_{f\,max}^2 n}{\sigma_y^2} = 1 + 3d = 10$$

Setting σ_{fmax} to 0.01 we get the following design requirement:

$$\frac{\sigma_{f\,max}^2 n}{\sigma_y^2} = 10 = \frac{0.01^2 n}{\sigma_y^2} \Rightarrow \frac{n}{\sigma_y^2} = 100{,}000$$

From this equation we see that if we can only measure the values of y to an accuracy of σ_y = 1, then we would require 100,000 randomly distributed points to satisfy the design requirement. However, if we could measure the values of y to an accuracy of σ_y = 0.1, then the number of points required is reduced to 1000. If we relax our requirement regarding σ_f from 0.01 to 0.02 then the value of n is reduced by an additional factor of 4.

Example 4.10.2:

In Example 4.10.1 with the requirement σ_{fmax} = 0.01 it was seen that 1000 points were required to achieve the desired result if the values of y could be measured to an accuracy of σ_y = 0.1. Assuming that the values of the x's are adjusted so that their means are zero (i.e., all the r_j's are zero), what

is the predicted accuracy for the coefficients in Equation 4.10.1? From Equations 4.10.5, 4.10.6 and 4.10.7 we get:

$$\Phi_1 \equiv \frac{\sigma_{a_1} n^{1/2}}{\sigma_y} = 1 = \frac{\sigma_{a_1} 1000^{1/2}}{0.1} \Rightarrow \sigma_{a_1} = \frac{0.1}{1000^{1/2}} = 0.0032$$

$$\Phi_{j+1} \equiv \frac{\sigma_{a_{j+1}} n^{1/2} (n_j \Delta x_j)}{\sigma_y} = 12^{1/2} = \frac{\sigma_{a_{j+1}} 1000^{1/2} (n_j \Delta x_j)}{0.1}$$

$$\sigma_{a_{j+1}} = \frac{0.1(12/1000)^{1/2}}{n_j \Delta x_j} = \frac{0.011}{n_j \Delta x_j} \qquad j = 1 \text{ to } d$$

It should be remembered that these results are based upon the assumption that Equation 4.10.1 is a realistic model for y as a function of x_1, x_2 and x_3. When the experiment is performed, this assumption can be tested using the Goodness-of-Fit tools discussed in Section 3.9.

Chapter 5 SEPARATION EXPERIMENTS

5.1 Introduction

Many experiments fall within the category of "separation experiments". The purpose of these experiments is to separate two or more competing phenomena. As an example consider experiments in which counts are recorded from several decaying radioactive species over a given time period. The purpose of the experiment might be to determine the relative amounts of the species or perhaps to identify the species by computing the various decay constants. Another example relates to spectroscopy in which spectral lines tend to overlap. In astronomy measurements of light coming from overlapping light sources present difficult problems of analysis.

Probably the most famous separation experiment is carbon dating which is used to determine the age of organic material. In living material the relative amounts of three carbon isotopes C_{12}, C_{13} and C_{14} is very close to being constant. Once living material dies, the fraction of carbon in the form of C_{14} starts to decrease because this isotope is not replenished through interaction with atmospheric carbon dioxide. The C_{14} content decreases through radioactive decay with a half life of approximately 5730 years. The age of the material can thus be estimated by measuring the relative amount of C_{14} to C_{12} and C_{13} in the specimen. Willard Libby developed the method of radioactive carbon dating in 1949 and received the Nobel prize in 1960 for this work.

In this chapter we will discuss several broad classes of separation experiments in which the mathematical models include exponential functions, Gaussian functions and sin functions. Prediction analyses of these classes are developed and can be used as starting points for planning experiments. The methodology for obtaining the results included in this chapter is described in Section 4.6. The common thread thru all these types of experi-

J. Wolberg, *Designing Quantitative Experiments*, DOI 10.1007/978-3-642-11589-9_5,

ments is that the difficulty in achieving accurate results increases dramatically as the separation between the competing phenomena decreases.

5.2 Exponential Separation Experiments

In this class of experiments we consider cases in which the purpose of the experiment is to analyze data based upon a model that includes several decaying exponential functions. Such experiments are well known in many areas of science and technology (e.g., chemical kinetics and radioisotope analysis). Considering experiments analyzing data from decaying radioactive isotopes, the purpose of the experiments might be to identify the isotopes or perhaps measure the relative amounts of the species included in the specimen. In chemical kinetics the purpose might be to measure the reaction rates of the competing processes. Typically experiments in this class utilize data that is obtained by counting the number of events occurring within a time "window". The size of the time window is usually an experimental variable and is one of the parameters that can be varied as part of the process of designing the experiment for optimum results. Other experimental variables include the number of data points and often the initial amplitude of the count rate. Clearly the greater the initial amplitude, the more accurate the measured parameters will be. However, when the cost involved in increasing the initial amplitude is considered, this parameter becomes part of the optimization process.

Assuming that the background number of counts per time window is negligible, the model might be based upon a two-exponential version of Equation 4.4.1:

$$y = f(t) = a_1 e^{-a_2 t} + a_3 e^{-a_4 t} \tag{5.2.1}$$

Once again assuming Poisson statistics (i.e., $\sigma_y = \sqrt{y}$), the dimensionless groups used to characterize the results of the prediction analysis of Equation 4.4.1 were expressed as Equations 4.5.1 and 4.5.2. For Equation 5.2.1 since the $\boldsymbol{C}$ matrix (Equation 3.4.14) is four by four, analytical expressions for the dimensionless groups will be so large as to make them useless. However we can express the results of a prediction analysis of Equation 5.2.1 in a similar manner using the following dimensionless groups:

$$\Phi_1 = \sigma_{a_1}\left[\frac{n}{K(a_1+a_3)}\right]^{1/2} \qquad \Phi_2 = \frac{\sigma_{a_2}(n(a_1+a_3))^{1/2}}{K^{1/2}a_2}$$

$$\Phi_3 = \sigma_{a_3}\left[\frac{n}{K(a_1+a_3)}\right]^{1/2} \qquad \Phi_4 = \frac{\sigma_{a_4}(n(a_1+a_3))^{1/2}}{K^{1/2}a_4}$$

In addition, three additional groups are required to completely specify the experiment: a_3/a_1 , a_4/a_2 , and $z \equiv a_2 n\Delta t$. In these equations if a_1 and a_3 are measured in counts within the Δt window, then K is one and $(a_1 + a_3)$ is the initial counts per Δt at t=0. (Alternatively, if a_1 and a_3 are measured in counts per unit time then $K = 1/\Delta t$.) Results for $a_3/a_1 = 1$ and $z = 8$ are seen in Tables 5.2.1 for various values of a_4/a_2.

a_4/a_2	Φ_1	Φ_2	Φ_3	Φ_4
1.10	10489.8	927.7	10486.5	1085.5
1.20	1497.8	237.9	1494.4	321.8
1.50	150.4	46.7	147.1	89.1
2.00	38.1	18.2	35.3	49.8
2.50	20.2	12.1	18.4	41.7
3.00	13.9	9.6	13.4	39.1

Table 5.2.1 – Values of Φ_k vs a_4/a_2 for a_3/a_1=1 and z=8

a_4 / a_2	z	Φ_1	Φ_2	Φ_3	Φ_4
1.10	4	28555.6	2634.5	28552.6	2821.9
1.10	8	10489.8	927.7	10486.5	1085.5
1.10	10	9237.8	809.7	9244.3	965.4
1.10	12	9060.0	791.1	9056.2	951.0
1.10	14	9363.5	816.5	9359.6	984.4
1.20	4	3779.2	647.1	3776.2	744.7
1.20	8	1497.8	237.9	1494.4	321.8
1.20	10	1368.3	214.4	1364.7	298.8
1.20	12	1378.6	215.0	1374.9	303.0
1.20	14	1446.7	225.2	1442.7	318.6

Table 5.2.2 – Values of Φ_k vs z for a_3/a_1=1 and a_4/a_2=1.1 and 1.2

We see that as a_4/a_2 approaches one, all the Φ's become very large. The effect of z is seen in Table 5.2.2 for $a_3/a_1 = 1$ $a_4/a_2 = 1.1$ and 1.2. We see

that for a_4/a_2 = 1.1 the optimum value of z is about 12 and for a_4/a_2 = 1.2 the optimum is about 10 to 11. The implication of these results upon the design of an experiment is shown in the following examples.

Example 5.2.1:

We wish to design an experiment in which our objective is to measure both a_2 and a_4 to about 1% accuracy. We know that a_3/a_1 is approximately 1 and a_4/a_2 is about 1.2. After choosing a value of z = 10 what values of n, a_1 and a_3 satisfy the objective of the proposed experiment? From Table 5.2.2 we see that Φ_4 is greater than Φ_2 (299 vs 214) so we need only design to satisfy the requirement that $\sigma_{a4} / a_4 = 0.01$. From our expression for Φ_4 we get the following:

$$\Phi_4 = \frac{\sigma_{a_4}(n(a_1 + a_3))^{1/2}}{K^{1/2} a_4} = 298.8 = 0.01(n(a_1 + a_3))^{1/2}$$

Solving, we obtain the following expression:

$$n(a_1 + a_3) = \frac{298.8^2}{0.01^2} = 8.93 * 10^8$$

If n=1000 then a_1 and a_3 will each have to be about $4.46*10^5$ counts per Δt. If a_2 = 1 sec^{-1}, then $\Delta t = z / (a_2\, n)$ = 10/1000 = 0.01 sec and the initial count rate would 89.3 million cps (counts per second) which is a huge number! If a_2 = 1 $hour^{-1}$ = 1/3600 sec^{-1} then Δt = 36 sec and the initial count rate would be $8.93*10^5$ / 36 which is about 24,800 cps.

The experiment becomes more complicated when the value of a_3/a_1 is not equal to one. Results for several values of a_3/a_1 are included in Table 5.2.3 for z = 10 and a_4/a_2 = 1.2. As one might expect, as the ratio a_3/a_1 increases the fractional uncertainty in a_2 increases and the fractional uncertainty in a_4 decreases. From the definitions of Φ_1 and Φ_3 we see that the values of σ_{a1} and σ_{a3} decrease as a_3/a_1 increases. The explanation for this effect is explained by the separation phenomenon. As time increases the fraction of the counts coming from the longer lived species with decay constant a_2 increases. However, increasing a_3/a_1 results in an increase in the fraction of counts for the shorter lived species. This increase results in decreases for both σ_{a1} and σ_{a3}.

a_3 / a_1	Φ_1	Φ_2	Φ_3	Φ_4
0.25	1524.9	150.3	1521.3	826.0
0.50	1458.9	172.1	1455.4	475.7
1.00	1368.3	214.4	1364.7	298.8
2.00	1264.8	295.8	1261.3	208.5
4.00	1167.8	452.7	1164.4	161.5

Table 5.2.3 – Values of Φ_k vs a_3/a_1 for a_4/a_2=1.2 and z=10

Example 5.2.2:

We wish to design an experiment to determine both a_1 and a_3 to at least 10% accuracy. We know that a_4/a_2 is approximately 1.2 and a_3/a_1 is approximately 2. Assuming that we design the experiment so that z = 10, what conditions must be satisfied to achieve predicted values of σ_{a1}/a_1 and σ_{a3}/a_3 = 0.1? From the definitions of Φ_1 and Φ_3 we obtain the following expressions:

$$\Phi_1 = \sigma_{a_1}\left[\frac{n}{K(a_1+a_3)}\right]^{1/2} = \sigma_{a_1}\left[\frac{n}{K(3a_1)}\right]^{1/2} = \frac{\sigma_{a_1}}{a_1}\left[\frac{na_1}{3K}\right]^{1/2}$$

$$\Phi_3 = \sigma_{a_3}\left[\frac{n}{K(a_1+a_3)}\right]^{1/2} = \sigma_{a_3}\left[\frac{n}{K(1.5a_3)}\right]^{1/2} = \frac{\sigma_{a_3}}{a_3}\left[\frac{na_3}{1.5K}\right]^{1/2}$$

Using K=1 and values from Table 5.2.3 we obtain the following:

$$\Phi_1 = 1265 = 0.1\left[\frac{na_1}{3}\right]^{1/2} \Rightarrow na_1 = \frac{3*1265^2}{0.1^2} = 4.80*10^8$$

$$\Phi_3 = 1261 = 0.1\left[\frac{na_3}{1.5}\right]^{1/2} \Rightarrow na_3 = 2na_1 = \frac{1.5*1261^2}{0.1^2} = 2.39*10^8$$

From these two expressions, we see that the expression for Φ_1 is the more difficult to satisfy and should therefore be the design criterion.

An alternative approach to this experiment is to treat the values of a_2 and a_4 as known parameters. Thus instead of Equation 5.2.1 we could use:

$$y = f(t) = a_1 e^{-d_1 t} + a_2 e^{-d_2 t} \tag{5.2.2}$$

In this equation d_1 and d_2 are the known decay constants. This equation is linear with respect to the a_k's. The two dimensionless groups that characterize the results are:

$$\Phi_1 = \sigma_{a_1}\left[\frac{n}{K(a_1 + a_2)}\right]^{1/2} \quad \text{and} \quad \Phi_2 = \sigma_{a_2}\left[\frac{n}{K(a_1 + a_2)}\right]^{1/2}$$

For the case of z =10, a_2/a_1 =1 and d_2/d_1 = 1.2, the values of Φ_1 and Φ_2 are 17.0 and 20.4. Comparing these to the values of Φ_1 and Φ_3 from Table 5.2.3 (i.e., 1368.3 and 1364.7) we see that the predicted values of σ_{a_1} and σ_{a_2} using Equation 5.2.2 are several orders of magnitude less than what could be expected if Equation 5.2.1 with 4 unknown parameters is used.

The use of Equation 5.2.2 is perhaps misleading. It assumes that both a_2 and a_4 are known without any uncertainty. A more realistic approach to this problem is to include the estimated values of σ_{a_2} and σ_{a_4} as Bayesian estimators. For example, assuming that z =10, a_3/a_1 =1 and a_4/a_2 = 1.2, Table 5.2.4 includes the values of the Φ_k's as functions of the

Bayes Fractional Uncertainties (i.e., σ_{a2}/a_2 and σ_{a4}/a_4). The results in this table show that decreasing the Bayes Fractional Uncertainties (BFU) from 20% down to 0.1% reduces Φ_1 and Φ_3 down to the equivalent values for Φ_1 and Φ_2 that are obtained using Equation 5.2.2 (i.e., 17.0 and 20.4). As the BFU increases the Φ_k's approach the values obtained when no Bayesian estimators are specified.

Bayes Frac	Φ_1	Φ_2	Φ_3	Φ_4
Not Spec	1368.3	214.4	1364.7	298.8
0.200	656.0	104.4	654.6	143.6
0.100	362.8	60.0	362.1	79.9
0.050	189.3	34.7	189.8	42.6
0.020	80.6	17.5	81.8	18.7
0.010	43.6	9.6	45.4	9.8
0.001	17.5	1.0	20.8	1.2

Table 5.2.4 – Φ_k's vs Bayes Frac for a_3/a_1, a_4/a_2=1.2 and z=10

The values of the Φ_k's in Table 5.2.4 were obtained using the Prediction Analysis mode of REGRESS with specification of Bayesian estimators. The run shown in Figure 5.2.1 are for BFU = 0.02. Note that the value of $(n / (a_1 + a_3)) = 1$ and $K = 1$ therefore $\Phi_1 = \sigma_{a1}$ and $\Phi_3 = \sigma_{a3}$. Also, the value of $(n(a_1 + a_3))^{1/2} = 1000$ so $\Phi_2 = 1000\sigma_{a2}/a_2 = 1000\sigma_{a2}$ and $\Phi_4 = 1000\sigma_{a4}/a_4 = 1000\sigma_{a4}/1.2$. Note that PRED_SA(K) is the predicted value of σ_{ak} . Note that T1 is the mid-point of the first time window of width DelT and therefore $z = a_2 n \Delta t = 1*1000 * 0.01 = 10$.

```
PARAMETERS USED IN REGRESS ANALYSIS: Wed Jan 21, 2009

  INPUT PARMS FILE: tab524.par
  REGRESS  VERSION: 4.22, Jan 11, 2009

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  :  1000
  Y VALUES - computed using A0 & T values

  SYTYPE1 - Sigma type for Y              :     4
     TYPE 4: SIGMA Y = CY1 * sqrt(Y)   CY1: 1.000
  T Values - computed using interpolation table
             T1 = 0.005   DelT = 0.01


  Function Y:  A1*EXP(-A2*T)+ A3*EXP(-A4*T)

   K       A0(K)     SIGA0(K)        A(K)  PRED_SA(K)
   1   500.00000    Not Spec   500.00000    80.58974
   2     1.00000     0.02000     1.00000     0.01749
   3   500.00000    Not Spec   500.00000    81.83280
   4     1.20000     0.02400     1.20000     0.02249
```

Figure 5.2.1 – Prediction Analysis run for BFU = 0.02

Table 5.2.4 is misleading! Since all the results in this table were generated using the same value of n, a_1 and a_3 we don't see the effect of changes of these parameters when BFU is specified. An examination of Equation 3.7.7 shows that the specification of Bayesian estimators is accomplished by adding a term to the diagonal element of the C matrix for each Bayesian estimator. For example, for the run shown in Figure 5.2.1, the added term to the element C_{22} is $1/\sigma_{a2}^2 = 1/0.02^2 = 2500$ and the added term to the element C_{44} is $1/0.024^2 = 1736$. Note that the elements of the C matrix prior to these additions are function of n, a_1 and a_3. For example, the C_{22} element before addition of the Bayesian term is 8016.6 and with the addition of 2500 is 10516.6 (i.e., an increase of 31%). Reducing n, a_1 and a_3 each by a factor of 10 reduces C_{22} before addition of the Bayesian term to 80.17 but after addition of the Bayesian term it is 2580.17 (i.e., an increase of 3200%)!! In other words increases in n, a_1 or a_3 reduce the impact of the Bayesian estimators upon the results.

5.3 Gaussian Peak Separation Experiments

In this class of experiments the mathematical model includes two or more Gaussian peaks that are typically overlapping. The x variable might be a measure of energy or wavelength or perhaps distance. The y variable is some measure of the strength of a signal emanating from x and is often a count rate. Typical examples of these types of experiments are found in spectroscopy, astronomy, image processing and many other areas of science and engineering.

For our initial analysis we assume that the data is gathered along the x axis in windows of equal width Δx and that the y variable is a measure of counts in the Δx window. Since y is measured in counts it is reasonable to assume that Poisson statistics (i.e., $\sigma_y = \sqrt{y}$) is applicable. If we assume that the background number of counts in each window is negligible, the model might be based upon a two-peak Gaussian function:

$$y = f(x) = a_1 e^{-\left(\frac{x-a_3}{a_2}\right)^2} + a_4 e^{-\left(\frac{x-a_6}{a_5}\right)^2} \tag{5.3.1}$$

A typical plot of y versus x is shown in Figure 5.3.1. A set of dimensionless groups that can be used to display the results are:

$$\Phi_1 = \sigma_{a_1}\left[\frac{n}{Ka_1}\right]^{1/2} \qquad \Phi_2 = \frac{\sigma_{a_2}(na_1)^{1/2}}{a_2 K^{1/2}} \qquad \Phi_3 = \frac{\sigma_{a_3}(na_1)^{1/2}}{a_2 K^{1/2}}$$

$$\Phi_4 = \sigma_{a_4}\left[\frac{n}{Ka_4}\right]^{1/2} \qquad \Phi_5 = \frac{\sigma_{a_5}(na_4)^{1/2}}{a_5 K^{1/2}} \qquad \Phi_6 = \frac{\sigma_{a_6}(na_4)^{1/2}}{a_5 K^{1/2}}$$

A reasonable question to ask is why it is preferable to use a_2 in the denominator of ϕ_3 instead of a_3 (and a_5 in the denominator of ϕ_6). The locations of the peaks are not really important. Note that if the starting and ending points of x and the values of a_3 and a_6 are all increased an equal amount the values of the σ_{ak}'s will not be effected. The values of σ_{a3} and σ_{a6} should be compared with the widths of the peaks and not the locations. Clearly the experiment increases in difficulty as a_3 approaches a_6. We define the separation between the peaks as follows:

$$SEP = \frac{2(a_6 - a_3)}{a_2 + a_5} \tag{5.3.2}$$

Results for various values of ***SEP*** are shown in Table 5.3.1 for the case $a_2 = a_5$ and the range of x values starting from $a_3 - 2a_2$ and ending at $a_6 + 2a_5$. We see that as ***SEP*** decreases the values of the ϕ's increase rapidly.

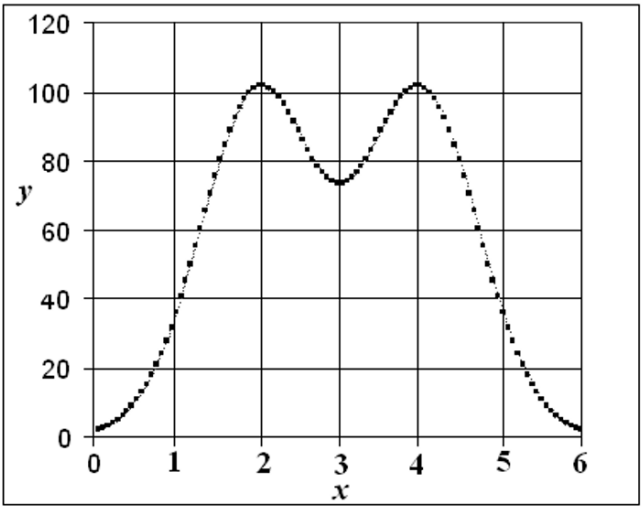

Figure 5.3.1 – Equation 5.3.1: a_3=2, a_6=4, a_1= a_4=100, a_2= a_5=1

SEP	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6
2.0	2.5	3.0	3.3	2.5	3.0	3.3
1.5	6.6	4.9	8.0	6.6	4.9	8.0
1.0	54.3	12.1	33.0	54.3	12.1	33.0
0.8	168.6	21.6	75.8	168.6	21.6	75.8
0.6	724.2	48.1	23.2	724.2	48.1	23.2
0.4	5721.2	159.5	1176.3	5721.2	159.5	1176.3

Table 5.3.1 – Φ_k's vs *SEP* for a_2=a_5, range from a_3-2a_2 to a_6+2a_5 Assuming Poisson Statistics (i.e., $\sigma_y = y^{1/2}$)

For the sake of comparison consider an experiment based upon data that will be modeled as a single Gaussian peak:

$$y = f(x) = a_1 e^{-\left(\frac{x-a_3}{a_2}\right)^2} \tag{5.3.3}$$

If the data range is similar to the range used in Table 5.3.1 (i.e., from a_3-$2a_2$ to a_3+$2a_2$), the values of Φ_1, Φ_2 and Φ_3 are 1.88, 1.18 and 1.09 respectively. Notice that for large values of ***SEP*** the comparable values are not very much greater that the values for a single peak (for ***SEP***=2, Φ_1 is 46% greater, Φ_2 is 18% greater and Φ_3 is a bit more than a factor of 2 greater). However, for small values of ***SEP*** the Φ_k's are orders of magnitude greater.

Example 5.3.1:

We wish to design an experiment to determine both peak locations a_3 and a_6 to an accuracy of ±0.1. The peak amplitudes (i.e., a_1 and a_4) are approximately the same and the widths of both peaks (i.e., a_2 and a_5) are about 10. Furthermore, we know that the dimensionless peak separation ***SEP*** is approximately 1. The data is to be gathered using a 256 channel analyzer (i.e., n = 256). All the data will be gathered at the same time so that Poisson statistics are applicable. By adjusting the time to gather the data we can design the experiment so that any value of a_1 and a_4 can be obtained. What should be the design value of the peak amplitude (i.e., both a_1 and a_4) that satisfies the requirement of the experiment?

For this experiment the relevant dimensionless groups are Φ_3 and Φ_6. From Table 5.3.1 both of these groups should be about 33. Since the amplitude and width of both peaks are the same, we need only design for one of the peaks and the other one will also be satisfied. Using the equation for Φ_3:

$$\Phi_3 = 33 = \frac{\sigma_{a_3}(na_1)^{1/2}}{a_2 K^{1/2}} = \frac{0.1(256a_1)^{1/2}}{10} = 0.16a_1^{1/2}$$

Defining y_i as counts in channel i, the value of K is one. From this equation we see that the design values of a_1 and a_4 should be $(33/0.16)^2$ = 42590. The duration of the experiment would be the amount of time required to obtain approximately this amount of counts in the peak channels.

In this section up to this point the analyses have all assumed Poisson statistics (i.e., the values of σ_y are equal to or proportional to $y^{1/2}$). For many experiments in which the y values are measured in terms of counts within a "window" centered at a value of x the assumption of Poisson statistics is reasonable. However, there are experiments when the x variable is varied over the range of interest and the values of y are each measured independently. For such experiments the time spent measuring y at each value of x can be varied so that each value of y can be measured to a constant value of σ_y / y. For example, if we would like to design an experiment such that all the values of σ_y / y are 0.01 (i.e., 1% accuracy), we would have to wait a sufficient time at each value of x so that 10000 counts are recorded. The values of y would thus be computed as ***counts*** / **measurement_*time*** for each value of x.

Defining F as the constant fractional uncertainty (i.e., σ_y / y) of each measured value of y, a set of dimensionless groups that can be used to display the results are:

$$\Theta_1 = \frac{\sigma_{a_1}}{a_1}\frac{n^{1/2}}{F} \qquad \Theta_2 = \frac{\sigma_{a_2}}{a_2}\frac{n^{1/2}}{F} \qquad \Theta_3 = \frac{\sigma_{a_3}}{a_2}\frac{n^{1/2}}{F}$$

$$\Theta_4 = \frac{\sigma_{a_4}}{a_4}\frac{n^{1/2}}{F} \qquad \Theta_5 = \frac{\sigma_{a_5}}{a_5}\frac{n^{1/2}}{F} \qquad \Theta_6 = \frac{\sigma_{a_6}}{a_5}\frac{n^{1/2}}{F}$$

Once again the ratios σ_{a3} / a_2 and σ_{a6} / a_5 are used in the definitions of the dimensionless groups. Results for various values of ***SEP*** are shown in Table 5.3.2 for the case $a_2 = a_5$ and the range of x values starting from $a_3 - 2a_2$ and ending at $a_6 + 2a_5$. We again see that as ***SEP*** decreases the values of the Θ's increase rapidly.

SEP	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	Θ_6
2.0	2.2	1.4	2.4	2.2	1.4	2.4
1.5	4.7	2.3	4.8	4.7	2.3	4.8
1.0	27.7	5.3	16.3	27.7	5.3	16.3
0.8	79.8	9.4	35.5	79.8	9.4	35.5
0.6	329.0	20.9	104.8	329.0	20.9	104.8
0.4	2593.7	71.2	532.9	2593.7	71.2	532.9

Table 5.3.2 – Θ_k's vs *SEP* for a_2=a_5, range from a_3-2a_2 to a_6+2a_5 AssumingConstant Fractional Uncertainty (i.e., σ_y / y= F)

For the sake of comparison consider an experiment based upon data that will be modeled as a single Gaussian peak (i.e., Equation 5.3.3). If the data range is similar to the range used in Table 5.3.2 (i.e., from a_3-2a_2 to a_3+2a_2), the values of Θ_1, Θ_2 and Θ_3 are 1.50, 0.42 and 0.43 respectively.

Example 5.3.2:

We wish to design an experiment similar to the experiment discussed in Example 5.3.1. The difference is that each value of y will be measured independently to a constant fractional uncertainty F. The objective is still to determine both peak locations a_3 and a_6 to an accuracy of ±0.1. The peak amplitudes (i.e., a_1 and a_4) are approximately the same and the widths of both peaks (i.e., a_2 and a_5) are about 10. Furthermore, we know that the dimensionless peak separation *SEP* is approximately 1. The data is to be gathered using a 256 channel analyzer (i.e., n = 256). What should be the design value of F that satisfies the requirements of the experiment?

For this experiment the relevant dimensionless groups are Θ_3 and Θ_6. From Table 5.3.2 both of these groups should be about 16.3. Since all the peak amplitudes and widths are the same for both groups we need only design for one of the peaks and the other one will also be satisfied. Using the equation for Θ_3:

$$\Theta_3 = 16.7 = \frac{\sigma_{a_3}}{a_2}\frac{n^{1/2}}{F} = \frac{0.1*16}{10F} \Rightarrow F = 0.0096$$

From this equation we see that the design values of F is approximately 0.01. To obtain this level of accuracy we would need to record about 10000 counts in each of the 256 channels.

To this point in this section all the analysis has been directed towards one specific case: the amplitudes of both peaks are the same, the widths of both peaks are the same and the range of x values is from a_3-$2a_2$ to a_6+$2a_5$. Only SEP was varied. We next consider 2 cases of varying amplitudes: the first case using Poisson statistics (Table 5.3.3) and the second case using constant fractional error for the values of y (Table 5.3.4). The results are symmetric with respect to x. It doesn't matter whether the peak at a_3 or at a_6 is larger. In Tables 5.3.3 and 5.3.4 we make the peak at a_3 the larger peak (i.e., a_1 / a_4 is greater than one), however, if a_1 / a_4 is less than one, we could just renumber the dimensionless groups. It makes no difference if the higher peak is located at the lower or higher value of x. For both cases we maintain the same range (from a_3-$2a_2$ to a_6+$2a_5$) and equal peak widths (i.e., $a_2 = a_5$). We also limit the cases to $SEP = 1$.

a_4 / a_1	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6
1.0	54.3	12.1	33.0	54.3	12.1	33.0
0.75	50.4	11.6	30.9	58.7	12.8	35.2
0.50	45.7	10.8	28.4	65.8	13.9	38.9
0.25	39.1	9.8	24.9	81.1	16.2	46.8
0.15	35.2	9.2	22.8	95.7	18.3	54.3
0.10	32.7	8.8	21.5	109.8	20.4	61.6

Table 5.3.3 – Φ_k's vs a_4 / a_1 for SEP=1, a_2=a_5, range a_3-$2a_2$ to a_6+$2a_5$, Assuming Poisson Statistics (i.e., $\sigma_y = y^{1/2}$)

The results for the Φ_k's representing the first peak decrease steadily as the 2nd peak becomes smaller. In other words, the larger peak parameters are measured to a greater accuracy as interference from the smaller peak decreases. The results for the smaller peak rise gradually, but the effect upon the σ's of this peak is more dramatic due to the definition of the Φ's for this peak. This effect is seen in Example 5.3.3.

Example 5.3.3:

We wish to design an experiment to determine both peak locations a_3 and a_6 to an accuracy of ±0.1. The ratio of the peak amplitudes (i.e., a_4 / a_1) is

approximately 0.15 and the widths of both peaks (i.e., a_2 and a_5) are about 10. Furthermore, we know that the dimensionless peak separation ***SEP*** is approximately 1. The data is to be gathered using a 256 channel analyzer (i.e., $n = 256$). All the data will be gathered at the same time so that Poisson statistics are applicable. By adjusting the time to gather the data we can design the experiment so that any value of a_1 can be obtained. What should be the design value of the peak amplitude (i.e., a_1) that satisfies the requirement of the experiment?

Clearly, the most difficult task of this experiment is to determine the peak location of the smaller peak to the specified accuracy. For this reason we use the definition of Φ_6 as our design criterion. From Table 5.3.3 the value of Φ_6 should be 54.3.

$$\Phi_6 = 54.3 = \frac{\sigma_{a_6}(na_4)^{1/2}}{a_5 K^{1/2}} = \frac{0.1(na_4)^{1/2}}{10} = \frac{1.6(a_4)^{1/2}}{10}$$

$$a_4 = 339^2 \approx 115{,}000$$

The design value for a_1 is a_4 /0.15 = 768,000. In example 5.3.1 the comparable value was 42590 so we see that satisfying the accuracy requirement for the 2nd peak requires a massive increase in the amplitude of the larger peak to satisfy the design requirements of the experiment. We can also compute the predicted resulting accuracy for the peak location of the larger peak:

$$\Phi_3 = 35.2 = \frac{\sigma_{a_3}(na_1)^{1/2}}{a_2 K^{1/2}} = \frac{\sigma_{a_3}(256 * 768000)^{1/2}}{10} \Rightarrow \sigma_{a_3} = 0.025$$

We see that the resulting measurement of a_3 will be about 40 times better than required in the design statement (i.e., ±0.025 compared to ±0.1).

Results for the 2nd case (i.e., constant fractional uncertainty for the values of y) are shown is Table 5.3.4. They are comparable to the results seen in Table 5.3.3 (which are based upon Poisson statistics). The values of the Θ_k's representing the first peak decrease steadily as the 2nd peak becomes smaller. The results for the smaller peak rise gradually, but the effect upon the σ's of this peak is more dramatic due to the definition of the Θ's for this peak. This effect is seen in Example 5.3.4.

a_4 / a_1	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	Θ_6
1.0	27.7	5.3	16.3	27.7	5.3	16.3
0.75	24.0	4.9	14.4	32.3	5.9	18.5
0.50	19.7	4.3	12.3	40.2	7.0	22.4
0.25	14.3	3.5	9.5	60.1	9.4	31.8
0.15	11.6	3.1	8.0	82.3	12.0	42.1
0.10	9.9	2.8	0.7	107.1	14.8	53.4

Table 5.3.4 – Θ_k's vs a_4/a_1 for *SEP*=1, $a_2=a_5$, range a_3-2a_2 to a_6+2a_5 AssumingConstant Fractional Uncertainty (i.e., $\sigma_y / y = F$)

Example 5.3.4:

We wish to design an experiment to determine both peak locations a_3 and a_6 to an accuracy of ±0.1. The ratio of the peak amplitudes (i.e., a_4 / a_1) is approximately 0.15 and the widths of both peaks (i.e., a_2 and a_5) are about 10. Furthermore, we know that the dimensionless peak separation ***SEP*** is approximately 1. The data is to be gathered using a 256 channel analyzer (i.e., $n = 256$). All the data will be gathered individually in a manner that permits each measurement of y to attain the same fractional uncertainty (i.e., σ_y / y are equal to a constant F). What should be the design value of F that satisfies the requirements of the experiment?

For this experiment the two relevant dimensionless groups are Θ_3 and Θ_6. Since Θ_6 is larger we use this as our design criterion.

$$\Theta_6 = 42.1 = \frac{\sigma_{a_6}}{a_5}\frac{n^{1/2}}{F} = \frac{0.1*16}{10F} \Rightarrow F = 0.0038$$

The requirement that each value of y should be measured to 0.38% accuracy means that the number of points counts to be recorded in each of the 256 channels of the analyzer are $(1/0.0038)^2 \approx 69000$. Compare this to the results obtained in Example 5.3.2 in which only about 10000 counts per channel were required to measure both locations to ±0.1. As noted in Example 5.3.3 the measurement of the location of the peak at a_3 will be much more accurate. Using the definition of Θ_3 and the value of 8.0 (Table 5.3.4) we obtain the following:

$$\Theta_3 = 8.0 = \frac{\sigma_{a_3}}{a_2}\frac{n^{1/2}}{F} = \frac{16\sigma_{a_3}}{10 * 0.0038} \Rightarrow \sigma_{a_3} = 0.019$$

We could next vary ***SEP***, the ratios a_4 / a_1 and a_2 / a_5 and the range of x values and generate many additional tables and figures but this would be "over-kill". The main point to note is the application of the method of prediction analysis to simulation of a proposed experiment. By limiting the range of cases studied to those relevant to the proposed experiment one can quickly get an understanding regarding the expected accuracy of the proposed experiment.

In this section the analysis is based upon a single independent variable x. For many problems the peaks are two dimensional. For example, in image processing one might be interested in examining images in which two "hot spots" are observed and one is interested in modeling these hot spots. The bivariate normal distribution is a generalization of the normal distribution into two independent variables x_1 and x_2 and can be used to model problems of this nature. A separation problem based upon use of two bivariate normal distribution is discussed in Section 7.3.

5.4 Sine Wave Separation Experiments

Sine waves are used to model phenomena in many areas of science and technology (e.g., music, power generation, radio transmission, radar, etc.). There are many different possibilities in this broad class of experiments. A general purpose model is a sine series with p terms, each term with an unknown amplitude, frequency and phase angle:

$$y = f(x) = \sum_{k=1}^{k=p} a_k \, sin(f_k 2\pi x + \varphi_k) \tag{5.4.1}$$

This equation has $3p$ unknown parameters: (a_k, f_k, φ_k) k = 1, p. The simplest case is a sum of harmonic terms in which all the phase angles are zero:

$$y = f(x) = \sum_{k=1}^{k=p} a_k \, sin(k 2\pi x) \tag{5.4.2}$$

This equation has only p unknowns and x is a dimensionless variable related to z, a variable with dimensions of length:

$$x = z / H \tag{5.4.3}$$

where H is the value of z for which $x = 1$. All the terms in the harmonic series are zero at $z = cH$ for all integer values of c. When Equation 5.4.2 is applicable, the data can be combined into the range $x = 0$ to 1 to improve the accuracy of the measured values of y. For example, assume that each value of y is measured to given accuracy (i.e., $\sigma_y = 0.1$). If we measure y for a range $x = 0$ to $x = 100$, we actually have 100 data points that are equivalent to each data point in the range 0 to 2π. Taking the average value of each set of 100 values of y, the value of σ_y is reduced by a factor of $n^{1/2} = 10$ and is thus 0.01.

The difference between sine series and series composed of exponential and Gaussian terms is that the dependent variable y can be negative. If all points are weighted equally, negative values of y are not a problem. However, if CFU (constant fractional uncertainty) is used to weigh the data, if the value of y is zero at a particular value of x then the weight of that point will be infinite and the C matrix will be singular. This effect is also possible if statistical weighting (i.e., $\sigma_y = y^{1/2}$) is used. One approach to this problem is to disregard points for which the weight is infinite. To avoid this problem the REGRESS program sets the weight of a point to zero if the denominator of Equation 3.3.7 is zero.

One of the nice features of the harmonic series 5.4.2 is that over the range from 0 to $2\pi c$ (where c is an integer) the various terms are orthogonal:

$$\int_0^{2\pi c} \sin(k_1 x)\sin(k_2 x)\,dx = 0 \quad \text{for } k_1 \neq k_2 \tag{5.4.4}$$

and

$$\int_0^{2\pi c} \sin(k_1 x)\sin(k_1 x)\,dx = \frac{1}{2} \tag{5.4.5}$$

As a result, for the case where we can assume $\sigma_y = K_y$ and $\sigma_x = 0$ (i.e., equal weight for all data points), the separation of harmonics is quite simple. In Table 5.4.1 values of Ψ_k are shown for values of p up to 5 for harmonics with equal amplitudes. We define the following dimensionless groups:

$$\Psi_k = \sigma_{a_k} \frac{n^{1/2}}{K_y} \tag{5.4.6}$$

When the range of x values is from 0 to c (an integer) we see that all values of Ψ_k are equal to $\sqrt{2}$.

p	Ψ_1	Ψ_2	Ψ_3	Ψ_4	Ψ_5
1	1.414				
2	1.414	1.414			
3	1.414	1.414	1.414		
4	1.414	1.414	1.414	1.414	
5	1.414	1.414	1.414	1.414	1.414

Table 5.4.1 – Ψ_k's vs p assuming $\sigma_y = K_y$

Using Equations 5.4.4 and 5.4.5 the results in Table 5.4.1 can easily be obtained analytically. The C matrix is p by p but all the off-diagonal terms are 0 due to Equation 5.4.4 and all the diagonal terms are $n/(2K_y^2)$. The results in Table 5.4.1 are valid for any p and for any values of the a_k's. For harmonic series such as Equation 5.4.2, there is no problem in separating the various harmonics because the Ψ_k's are not dependent upon the value of p.

Example 5.4.1:

The purpose of the experiment is to measure each of the first three amplitudes in a sine series to an accuracy of at least 0.1%. Assume that the amplitudes are proportional to $1/k^2$ so the 3rd amplitude will be the most difficult to measure. Assume that all the amplitudes can be increased by increasing the time T devoted to the experiment: $a_k = CT / k^2$ where C is a constant. Assume that our measurements are limited to the range 0 to 2π (i.e., x from 0 to 1). What is the design criterion that should be satisfied for this experiment?

$$\Psi_3 = \sigma_{a_3} \frac{n^{1/2}}{K_y} = 1.414 = 0.001 a_3 \frac{n^{1/2}}{K_y} = 0.001 \frac{CT}{3^2} \frac{n^{1/2}}{K_y}$$

$$1.414 \frac{9}{0.001} = CT \frac{n^{1/2}}{K_y} \Rightarrow CT \frac{n^{1/2}}{K_y} = 12726$$

For the case of CFU ($\sigma_y/y = F$), we define Θ_k as follows:

$$\Theta_k = \frac{\sigma_{a_k} n^{1/2}}{a_k F} \tag{5.4.7}$$

Note that Equations 5.4.4 and 5.4.5 are not applicable for this case as the terms of the C matrix include the weights of the individual points and these are not constant. In Table 5.4.2 values of Θ_k are shown for p = 1 to 5. Θ_k's are one (for any value of p) regardless of the values of the a_k's.

p	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5
1	1.000				
2	1.000	1.000			
3	1.000	1.000	1.000		
4	1.000	1.000	1.000	1.000	
5	1.000	1.000	1.000	1.000	1.000

Table 5.4.2 – Θ_k's vs p assuming $\sigma_y/y = F$

We see from the results in Table 5.4.2 that for experiments based upon Equation 5.4.2 and with the y values measured to a constant fractional uncertainty, there is no problem in separating the various harmonics. The Θ_k's are not dependent upon the value of p.

Example 5.4.2:

The purpose of the experiment is the same as in Example 5.4.1: to measure each of the first three amplitudes in a sine series to an accuracy of at least 0.1%. For this example we assume that each value of y is measured to a constant fractional uncertainty of F. Since all the values of Θ_k are one, from Equation 5.4.7 we see that each of the amplitudes will be determined to the same fractional accuracy and the results are not dependent upon the individual amplitudes. Assume that our measurements are limited to the

range 0 to 2π (i.e., x from 0 to 1). What is the design criterion that should be satisfied for this experiment?

$$\Theta_k = \frac{\sigma_{a_k} n^{1/2}}{a_k F} = 1 = 0.001 \frac{n^{1/2}}{F} \Rightarrow \frac{n^{1/2}}{F} = 1000$$

The next level of complexity is to include phase angles in the model:

$$y = f(x) = \sum_{k=1}^{k=p} a_k \, sin(k 2\pi x + \varphi_k) \tag{5.4.8}$$

This model has $2p$ unknowns: (a_k, φ_k) k = 1, p. For the case where the values of y are determined to a constant level of uncertainty (i.e., $\sigma_y = K_y$), The values of the Ψ_k's (Equation 5.4.6) are still $\sqrt{2}$ for this model. The appropriate dimensionless groups for the uncertainties that we can expect for the computed values of the φ_k's are:

$$\Psi_{\varphi k} = \frac{\sigma_{\varphi k}}{a_k} \frac{n^{1/2}}{K_y} = \sqrt{2} \tag{5.4.9}$$

Example 5.4.3:

The purpose of the experiment is similar to Example 5.4.1: to measure all the parameters of the first 3 harmonics to an accuracy of at least 0.1%. Assume that the amplitudes are proportional to $1/k^2$ so the 3rd amplitude will be the most difficult to measure. In example 5.4.1 we developed a design criterion for measuring a_k's to the required accuracy. We now need to develop similar criteria for the measurement of the φ_k's. From Equation 5.4.9 we see that the predicted values of the fractional uncertainties of the phase angles are not dependent upon the values of the a_k's. Assume that our measurements are limited to the range 0 to 2π (i.e., x from 0 to 1), the design criterion for the phase angles is:

$$\Psi_{\varphi k} = \sqrt{2} = \frac{\sigma_{\varphi k}}{a_k} \frac{n^{1/2}}{K_y} = 0.001 \frac{n^{1/2}}{K_y} \Rightarrow \frac{n^{1/2}}{K_y} = 1414$$

In both Tables 5.4.1 and 5.4.2 it was assumed that the values of $\boldsymbol{x}$ are error free. For experiments of this type random errors in the values of $\boldsymbol{x}$ are usually very small. However, the experiment might be affected by a systematic error in the $\boldsymbol{x}$'s. For example, the measured value of $\boldsymbol{x} = 0$ might not correspond to the true zero point of the sine series. For example, using the model Equation 5.4.2, for the case of $\boldsymbol{p}$ =3 and all the $\boldsymbol{a_k}$'s equal, a systematic error of 2 degrees ($\boldsymbol{\varepsilon_x}$ = 2/360) causes errors in $\boldsymbol{a_1}$ = -0.07%, in $\boldsymbol{a_2}$ = -0.24%, and in $\boldsymbol{a_3}$ = -0.64%.

For problems in which the general purpose Equation 5.4.1 is valid, then the complexity of the experiment is vastly increased even if we can assume that the phase angles are zero. Let us consider one specific case of this equation: $\boldsymbol{p} = 2$, $\boldsymbol{f_1} = 1$ and the range of $\boldsymbol{f_1 x}$ is from 0 to $\boldsymbol{2\pi}$. Let us further assume that both $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ are positive and $\boldsymbol{f_2} < \boldsymbol{f_1}$. For the weighting schemes discussed above, as $\boldsymbol{f_2}$ approaches $\boldsymbol{f_1}$ the predicted values of the $\boldsymbol{\sigma_{ak}}$'s and $\boldsymbol{\sigma_{fk}}$'s approach infinity. To illustrate this point using CFU ($\boldsymbol{\sigma_y/y} = \boldsymbol{F}$), the relevant dimensionless groups are:

$$\Theta_{ak} = \frac{\sigma_{a_k} n^{1/2}}{a_k F} \quad \text{and} \quad \Theta_{fk} = \frac{\sigma_{f_k} n^{1/2}}{f_k F} \tag{5.4.10}$$

Results are included in Table 5.4.3 for various values of $\boldsymbol{f_2/f_1}$. The results in this table are limited to the specific case $\boldsymbol{a_1} = \boldsymbol{a_2}$. The results in the table are easily explainable. For small values of $\boldsymbol{f_2/f_1}$ (e.g., 0.1), the range of $\boldsymbol{x}$ includes only a small portion of the first cycle of $\boldsymbol{f_2}$ so the values of $\boldsymbol{\Theta_{a2}}$ and $\boldsymbol{\Theta_{f2}}$ are very large. For $\boldsymbol{f_2/f_1} = 0.5$ the best results are obtained for the 2nd harmonic because the range $\boldsymbol{f_2 x}$ is from 0 to $\boldsymbol{\pi}$. Also, as $\boldsymbol{f_2/f_1}$ approaches one, all the $\boldsymbol{\Theta}$'s become large confirming that separation become increasingly difficult under this condition.

$\boldsymbol{f_2/f_1}$	$\boldsymbol{\Theta_{a1}}$	$\boldsymbol{\Theta_{a2}}$	$\boldsymbol{\Theta_{f1}}$	$\boldsymbol{\Theta_{f2}}$
0.10	1.1	121.5	0.5	131.3
0.25	1.1	6.1	0.5	11.1
0.50	1.1	2.1	0.5	1.0
0.75	35.8	36.7	5.2	6.2
0.90	336.6	337.7	18.4	17.6

Table 5.4.3 – $\boldsymbol{\Theta}$'s vs $\boldsymbol{f_2/f_1}$ for case $\boldsymbol{a_1} = \boldsymbol{a_2}$, assuming $\boldsymbol{\sigma_y/y} = \boldsymbol{F}$

Example 5.4.4:

We wish to design an experiment to measure both f_1 and f_2 to 1% accuracy. Assuming that the range of the experiment will be $0 \leq f_1 x \leq 2\pi$ and that f_2/f_1 is approximately 0.75, if $n = 100$, what fractional accuracy F is required to satisfy the experimental objectives?

From Table 5.4.3 we see that the more difficult objective is to measure f_2 so we should satisfy the requirement that $\Theta_{f2} = 6.2$. From Equation 5.4.10:

$$\Theta_{f2} = 6.2 = \frac{\sigma_{f_2} n^{1/2}}{f_2 F} = \frac{0.01 * 100^{1/2}}{F} \Rightarrow F = \frac{0.1}{6.2} = 0.016$$

Thus to satisfy the requirements of the experiment, the values of y should be measured to about 1.6% accuracy.

5.5 Bivariate Separation

In the previous sections the separation problems were all based upon a single independent variable. In this section we consider a generalization of the Gaussian separation problems considered in Section 5.3: separation of bivariate normally distributed peaks. The bivariate normal distribution (Equation 7.3.1) is discussed in Section 7.3 in detail.

There are applications in which the models require two-dimensional peaks. For example, astronomical data related to overlapping heavenly bodies (e.g., stars, galaxies, black holes) can be modeled using this distribution. Experiments of these types typically are based upon numbers of counts within a specified area on a grid and therefore statistical weighting ($\sigma_y = C_y y^{1/2}$) most often appropriate. For cases in which the values of the independent variables are uncorrelated (i.e., $\rho = 0$), Equation 7.3.1 (the two-dimensional distribution for a single peak) reduces to:

$$f(x_1, x_2) = \frac{1}{2\pi} exp\left(\frac{-x_1^2 - x_2^2}{2}\right) \tag{5.5.1}$$

Equation 5.5.1 assumes that both x_1 and x_2 have been normalized. However, if the mean values and standard deviations are unknown parameters then the equation must be expanded accordingly:

$$f(x_1, x_2) = \frac{1}{2\pi} exp\left(\frac{-((x_1 - \mu_1)/\sigma)^2 - ((x_2 - \mu_2)/\sigma)^2}{2}\right) \quad (5.5.2)$$

This equation assumes that the standard deviation σ is the same in both directions. If we have two peaks, we can choose as our x_1 axis the line between the mid-points of the two peaks and thus $\mu_2 = 0$ for both peaks. Our distribution for each peak becomes:

$$f(x_1, x_2, \mu, \sigma) = \frac{1}{2\pi} exp\left(\frac{-((x_1 - \mu)/\sigma)^2 - (x_2/\sigma)^2}{2}\right) \quad (5.5.3)$$

Furthermore is we set $x_1 = 0$ at the mid-point between the two peaks and define μ as the distance between the peaks, our two peak model is:

$$y = a_1 f(x_1, x_2, -\mu/2, \sigma_1) + a_2 f(x_1, x_2, \mu/2, \sigma_2) \quad (5.5.4)$$

This model has 5 unknown: a_1, a_2, μ, σ_1 and σ_2. If we assume that $\sigma_1 = \sigma_2 = \sigma$ then the model is reduced to 4 unknowns:

$$y = a_1 f(x_1, x_2, -\mu/2, \sigma) + a_2 f(x_1, x_2, \mu/2, \sigma) \quad (5.5.5)$$

In addition to the model parameters, there are experimental parameters. Assuming that data is collected by observing the number of "events" taking place within a square box on a grid that is placed symmetrically about the two peaks, an example of the experimental layout is shown in Figure 5.5.1. This layout is for an experiment in which a very coarse grid of only 5 values of x_1 and 5 values of x_2 are used for a total of 25 data points. In Figure 5.5.2 a prediction analysis for the same peak separation and range is show but a much finer grid with 10000 (i.e., 100 by 100) data points is specified.

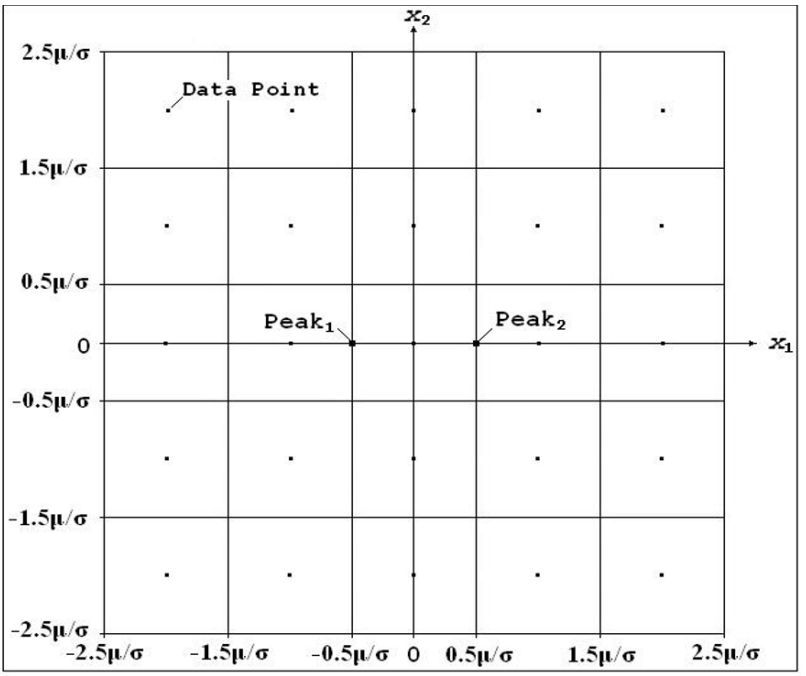

Figure 5.5.1 – Experimental Layout: 25 data points positioned equally in a square grid of size 5μ/σ. The peak separation is μ/σ.

```
PARAMETERS USED IN REGRESS ANALYSIS: Mon Jun 08, 2009
  INPUT PARMS FILE: fig552.par
  INPUT DATA  FILE: fig552.par
  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  : 10000
  Y VALUES - computed using A1,A2.X1,X2,MU & SIG values
  SYTYPE1 - Sigma type for Y              :     4
     TYPE 4: SIGMA Y = CY1 * sqrt(Y)   CY1: 1.000
  M - Number of independent variables     :     2
  X1 Values - computed using interpolation table
       NP1=100  X01=-2.4750  DELX1=0.0500
  X2 Values - computed using interpolation table
       NP2=100  X02=-2.4750  DELX2=0.0500
Analysis for Set 1
  Function Y:  A1*EXP(-((X1+MU/2)^2+X2^2)/(2*SIG^2))/(2*PI)
              +A2*EXP(-((X1-MU/2)^2+X2^2)/(2*SIG^2))/(2*PI)

   PARAMETER  INIT_VALUE       VALUE  PRED_SIGMA
          A1     0.50000     0.50000     0.07574
          A2     0.50000     0.50000     0.07574
         SIG     1.00000     1.00000     0.04359
          MU     1.00000     1.00000     0.25188
```

Figure 5.5.2 – Prediction Analysis: 10000 data points positioned equally in a square grid of size 5μ/σ. The peak separation is μ/σ.

Note that in Figure 5.5.2 the predicted values of $\boldsymbol{\sigma_a}$ were the same for both $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ (i.e., 0.07574). This is due to the fact that $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ were assumed to be equal (i.e., 0.5). The results can be expressed using the following dimensionless groups:

$$\Phi_{a_k} \equiv \frac{\sigma_{a_k} n^{1/2}}{C_y (a_1 + a_2)^{1/2}} \qquad (k = 1 \ \& \ 2) \tag{5.5.6}$$

$$\Phi_{\mu} \equiv \frac{\sigma_{\mu}}{\sigma} \frac{n^{1/2} (a_1 + a_2)^{1/2}}{C_y} \tag{5.5.7}$$

$$\Phi_{\sigma} \equiv \frac{\sigma_{\sigma}}{\sigma} \frac{n^{1/2} (a_1 + a_2)^{1/2}}{C_y} \tag{5.5.8}$$

Values of the $\boldsymbol{\phi}$'s are included in Table 5.5.1 for values of $\boldsymbol{N}$ = 2 to 6 where $\boldsymbol{N}$ is the number of multiples of $\boldsymbol{\mu/\sigma}$ (i.e., ***range*** = $\boldsymbol{N}$ $\boldsymbol{\mu/\sigma}$). The results in the table are limited to the specific case where $\boldsymbol{\mu = \sigma}$ and $\boldsymbol{a_1 = a_2}$.

N	Φ_{a_1} & Φ_{a_2}	Φ_{σ}	Φ_{μ}
2	9.425	10.845	51.92
3	7.707	6.092	31.16
4	6.929	4.690	25.64
5	7.574	4.359	25.19
6	8.613	4.530	27.22

Table 5.5.1 – Effect of Range $N\mu/\sigma$ for the Case $\mu = \sigma$ and $a_1 = a_2$.

We see from Table 5.5.1 that the optimum range is of the order 4 to 5 times $\boldsymbol{\mu/\sigma}$ for this specific case (i.e., $\boldsymbol{\mu = \sigma}$ and $\boldsymbol{a_1 = a_2}$). For $\boldsymbol{N}$ = 5, the range is thus from $\boldsymbol{2\sigma}$ to the left of the first peak to $\boldsymbol{2\sigma}$ to the right of the 2nd peak.

In Table 5.5.2 we see the effect of $\boldsymbol{\mu/\sigma}$ upon the values of the $\boldsymbol{\phi}$'s. All calculations in this table are for the range $\boldsymbol{peak_1 - 2\sigma}$ to $\boldsymbol{peak_2 + 2\sigma}$ (i.e., $\boldsymbol{-\mu/2-2\sigma}$ to $\boldsymbol{\mu/2+2\sigma}$) in the $\boldsymbol{x_1}$ direction and an equal range in the $\boldsymbol{x_2}$ direction. The results in the table yield an interesting insight into experiments of this type. Note that the values of $\boldsymbol{\phi}$'s increase when $\boldsymbol{\mu/\sigma}$ increases from 2 to 4. In the previous sections of this chapter we discussed separation experiments and in all the models considered, as separation increased predicted

results were improved. The explanation for the predicted decrease in accuracy as the separation increases from 2 to 4 is due to the range in the x_2 direction. For example, when $\mu/\sigma = 4$, all points at the extreme values of x_2 (i.e., $\pm 4\sigma$) are at least 4σ from the nearest peak and thus there is very little information in these distant points. By reducing the range in the x_2 direction, values of the ϕ's can be reduced. For example, reducing the range in the x_2 direction to $\pm 2\sigma$ the values for μ/σ equal to 4 are reduced to ϕ_a=5.52 from 7.29, Φ_σ=4.30 from 4.66 and Φ_μ=12.75 from 17.53.

μ/σ	Φ_{a_1} & Φ_{a_2}	Φ_σ	Φ_μ
0.50	11.270	4.425	49.89
0.75	8.670	4.381	33.13
1.00	7.574	4.359	25.19
1.50	6.764	4.342	18.26
2.00	6.566	4.337	15.93
4.00	7.289	4.658	17.53

Table 5.5.2 – Effect of Peak Separation for the Case $a_1 = a_2$ and range $-\mu/2-2\sigma$ to $\mu/2+2\sigma$.

In Table 5.5.3 the effect of differing amplitudes in the two peaks is considered. The assumptions used to generate this table are that $\mu/\sigma = 1$, $\sigma_1 = \sigma_2$ and the range is $-\mu/2-2\sigma$ to $\mu/2+2\sigma$ (i.e., -2.5σ to 2.5σ) in both directions.

a_1/a_2	Φ_{a_1}	Φ_{a_2}	Φ_σ	Φ_μ
10	8.670	10.079	4.200	23.59
8	8.550	9.990	4.219	23.79
6	8.357	9.818	4.245	24.05
4	8.026	9.445	4.282	24.43
2	7.486	8.476	4.337	24.97
1	7.574	7.574	4.359	25.19

Table 5.5.3 – Effect of Amplitude Variation for the Case $\mu/\sigma = 1$ and range $-\mu/2-2\sigma$ to $\mu/2+2\sigma$.

.

Example 5.5.1:

We want to design an experiment to analyze data based upon two bivariate distributions with negligible correlation coefficients. The purpose of the experiment is to determine the separation **μ** between the peaks of the two distributions to approximately 1% accuracy. We can assume that **σ** of both peaks are equal and approximately equal to the peak separation. The ratio of the amplitudes of the two peaks is approximately 5. The data is to be obtained in a square symmetric grid of sides approximately equal to **5σ.**

From interpolation in Table 5.5.3 an estimated value of $\Phi_\mu = 24.2$ is obtained. Assuming that the values of a_1 and a_2 are specified in the same units as the measured values of y, the value of C_y is one. From Equation 5.5.7 we obtain the following design criterion:

$$\Phi_\mu \equiv \frac{\sigma_\mu}{\sigma}\frac{n^{1/2}(a_1+a_2)^{1/2}}{C_y} = 24.2 = 0.01 * n^{1/2}(a_1+a_2)^{1/2}$$

$$2420 = n^{1/2}(a_1+a_2)^{1/2} \Rightarrow n(a_1+a_2) = 5.86 * 10^6$$

If, for example, the grid is 100 by 100 (i.e., $n = 10000$), then the sum of the two "hot spots" (i.e., $a_1 + a_2$) should be 586 (i.e., approximately 98 in the smaller peak and 488 in the larger peak). In other words, the number of counts that should be observed in the area of size 5**σ**/100 by 5**σ**/100 near the larger peak should be about 488.

Example 5.5.2:

For the experiment analyzed in Example 5.5.1 what fractional accuracy can we expect for the measured values of a_1, a_2 and **σ**? Assume that the square grid is divided into 100 by 100 equally sized sections.

From Table 5.5.3 we estimate the values of Φ_{a1}, Φ_{a2} and Φ_σ to be approximately 8.19, 9.66 and 4.26. From Equation 5.5.6:

$$\Phi_{a_1} \equiv \frac{\sigma_{a_1} n^{1/2}}{C_y (a_1+a_2)^{1/2}} = 8.19 = \frac{\sigma_{a_1} n^{1/2}}{(1.2a_1)^{1/2}}\frac{a_1^{1/2}}{a_1^{1/2}}$$

$$8.19 * 1.2^{1/2} = \frac{\sigma_{a_1}}{a_1} n^{1/2} a_1^{1/2} \Rightarrow \frac{\sigma_{a_1}}{a_1} = \frac{8.97}{(na_1)^{1/2}} = \frac{8.97}{(10^4 * 586)^{1/2}} = 0.0037$$

$$\Phi_{a_2} \equiv \frac{\sigma_{a_2} n^{1/2}}{C_y (a_1 + a_2)^{1/2}} = 9.66 = \frac{\sigma_{a_2} n^{1/2}}{(6a_2)^{1/2}} \frac{a_2^{1/2}}{a_2^{1/2}}$$

$$9.66 * 6^{1/2} = \frac{\sigma_{a_2}}{a_2} n^{1/2} a_2^{1/2} \Rightarrow \frac{\sigma_{a_2}}{a_2} = \frac{23.66}{(na_2)^{1/2}} = \frac{23.66}{(10^4 * 98)^{1/2}} = 0.024$$

From Equation 5.5.8:

$$\Phi_\sigma \equiv \frac{\sigma_\sigma}{\sigma} \frac{n^{1/2} (a_1 + a_2)^{1/2}}{C_y} = 4.26 = \frac{\sigma_\sigma}{\sigma} 10000^{1/2} * 586^{1/2}$$

$$\frac{\sigma_\sigma}{\sigma} = 4.26 / (586 * 10000)^{1/2} = 0.00176$$

We see from these results that the experiment designed to measure the peak separation μ to 1% accuracy also yields the three other parameters but to accuracies different than 1%. The larger peak amplitude a_1 will be measured to a predicted accuracy of 0.37%, the smaller peak amplitude a_2 will be measured to a predicted accuracy of 2.4% and σ to a predicted accuracy of 0.18%.

Chapter 6 INITIAL VALUE EXPERIMENTS

6.1 Introduction

Many mathematical models are expressed as initial value problems starting from a differential equation (or equations). Often the differential equation can be solved analytically and thus the model is expressed as an equation that includes some unknown parameters. Experiments to determine the unknown parameters might then be performed. However, some initial value problems do not have a simple analytical solution. Even if there is no analytical solution one can still use the method of least squares to determine the unknown parameters of the model. The solution is simply expressed as an integral (or integrals) and the integrals are computed numerically. In the REGRESS program the integral operator INT is used to perform numerical integration. An example of the use of such an operator is shown in Section 6.2. In Section 6.3 an initial value problem with an analytical solution is discussed. This example is based upon the well-known logistics population model. In Section 6.4 we consider an initial value problem in which the model is based upon several simultaneous first order differential equations. This problem is an example of the type of chain reactions that are encountered in many branches of science and engineering (e.g., nuclear decay, chemical kinetics). The design of an experiment for estimating the coefficients used to model a Chemostat is considered in Section 6.5. In Section 6.6 astronomical observations to determine orbits of rotating bodies are considered. The determination of orbits is an initial value problem best analyzed in polar coordinates. The solution utilizes two coupled integral equations.

6.2 A Nonlinear First Order Differential Equation

As an example of a nonlinear first order differential equation, the general form of the Ricatti equation is:

J. Wolberg, *Designing Quantitative Experiments*, DOI 10.1007/978-3-642-11589-9_6,

$$\frac{dy}{dx} = p(x)y^2 + q(x)y + r(x) \tag{6.2.1}$$

Let us consider one case within this family of equations: $p(x)=a_1$, $q(x)=0$ and $r(x)=x$:

$$\frac{dy}{dx} = a_1 y^2 + x \tag{6.2.2}$$

There is no solution to this equation based upon elementary mathematical functions but we can express the solution as an integral. Using the boundary condition $y=a_2$ at $x=0$ we can solve for the values of a_1 and a_2 by fitting the following equation to a set of experimentally determined values of x and y in the range 0 to $xmax$:

$$y = \int_{x=0}^{x} (a_1 y^2 + x)dx + a_2 \tag{6.2.3}$$

This is an interesting equation and for any combination of a_1 and a_2 there are an infinite number of values of x for which the value of y becomes infinite. Analytical equations for the x singularities cannot be expressed with elementary mathematical functions but can be determined using numerical methods. As long as the range of values of x doesn't include one of these singularities then the experiment can proceed to a solution for a_1 and a_2. As an example of an analysis of one case, Figure 6.2.1 includes a REGRESS run using a_1=0.9, a_2=1 with 11 values of x ranging from 0 to 1. Unit weighting (UNIT) was used for the y values. Note that the value of y explodes as x reaches 1. The singularity is thus slightly greater than x=1 for this combination of a_1 and a_2.

```
PARAMETERS USED IN REGRESS ANALYSIS: Tue Feb 03 16:56:47 2009

  INPUT PARMS FILE: riccati.par
  INPUT DATA  FILE: riccati.par
  REGRESS  VERSION: 4.23, Jan 27, 2009

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  :    11
  Y VALUES - computed using A0 & X values
  SYTYPE1 - Sigma type for Y              :     1
     TYPE 1:  SIGMA Y = 1
   X Values - computed using interpolation table

Analysis for Set 1
  Function Y:  A2+INT(X+A1*Y^2,0,X)

   K        A0(K)         A(K)  PRED_SA(K)
   1     0.90000      0.90000      0.06520
   2     1.00000      1.00000      0.09873

 POINT          X1         YCALC    PRED_SIGY
     1     0.00000      1.00000      0.09873
     2     0.10000      1.10424      0.11022
     3     0.20000      1.24262      0.12454
     4     0.30000      1.42717      0.14287
     5     0.40000      1.67769      0.16710
     6     0.50000      2.02929      0.20049
     7     0.60000      2.55065      0.24899
     8     0.70000      3.39656      0.32469
     9     0.80000      5.00430      0.45534
    10     0.90000      9.21683      0.71033
    11     1.00000     55.34458      0.99737
```

Figure 6.2.1 – Prediction Analysis with unit weighting

The results in Figure 6.2.1 show that for this particular case the predicted value of σ_{a1} is 0.065 and the predicted value of σ_{a2} is 0.099. These results are based upon an assumption that all data points are weighted equally. We could assume that the values of σ_y are proportional to y and thus the weights would be proportional to $1/y^2$. This method of weighting is called CFU (Constant-Fractional-Uncertainty). In Figure 6.2.2 we see results similar to those in Figure 6.2.1 but using CFU in place of UNIT weighting of the data points.

```
 PARAMETERS USED IN REGRESS ANALYSIS: Tue Feb 03 16:49:46 2009

   INPUT PARMS FILE: riccati.par
   INPUT DATA  FILE: riccati.par
   REGRESS  VERSION: 4.23, Jan 27, 2009

   Prediction Analysis Option  (MODE='P')
   N - Number of recs used to build model  :    11
   Y VALUES - computed using A0 & X values
   SYTYPE1 - Sigma type for Y              :     3
      TYPE 3:  SIGMA Y = CY1 * Y        CY1: 1.000
   X Values - computed using interpolation table

 Analysis for Set 1
   Function Y:  A2+INT(X+A1*Y^2,0,X)

    K        A0(K)         A(K)  PRED_SA(K)
    1      0.90000      0.90000      0.00899
    2      1.00000      1.00000      0.01383

 POINT          X1         YCALC    PRED_SIGY
     1      0.00000      1.00000      0.01383
     2      0.10000      1.10424      0.01546
     3      0.20000      1.24262      0.01748
     4      0.30000      1.42717      0.02008
     5      0.40000      1.67769      0.02351
     6      0.50000      2.02929      0.02825
     7      0.60000      2.55065      0.03514
     8      0.70000      3.39656      0.04591
     9      0.80000      5.00430      0.06450
    10      0.90000      9.21683      0.10043
    11      1.00000     55.34458      0.01807
```

Figure 6.2.2 – Prediction Analysis with Constant Fractional Uncertainty

At first glance a comparison of Figures 6.2.1 and 6.2.2 leads one to the conclusion that fractional weighting leads to a much more accurate experiment than if all data points are weighted equally. The average values of $\boldsymbol{\sigma_y}$ are very different for the two cases (UNIT and CFU). In Figure 6.2.1 it was assumed that $\boldsymbol{\sigma_y} = 1$. If the estimated values of $\boldsymbol{\sigma_y} = \boldsymbol{C}$ where $\boldsymbol{C}$ is a constant, then the values of the **PRED_SA(K)** and **PRED_SIGY** shown in the figure will be proportional to $\boldsymbol{C}$. In Figure 6.2.2 it was assumed that $\boldsymbol{\sigma_y} = \boldsymbol{Cy}$ and $\boldsymbol{C} = 1$. The values of the **PRED_SA(K)** and **PRED_SIGY** shown in the figure will also be proportional to $\boldsymbol{C}$. For this case the weights decrease rapidly as $\boldsymbol{x}$ increases so the average weight for the CFU case is much less than one (i.e., the average weight for the unit weight case). The values of **PRED_SA(K)** and **PRED_SIGY** increase with increasing weights so this explains why the CFU results are so much less.

For both cases the results will be inversely proportional to $n^{1/2}$. In other words, if the range remains the same and the value of ***n*** is increased by a factor of 4, then the values of **PRED_SA(K)** and **PRED_SIGY** will be decreased by a factor of 2.

Equation 6.2.2 is an example of an equation that exhibits explosive growth (i.e., growth that reaches infinity for a finite value of ***x***). To compute the value of ***x*** at which the singularity occurs we would have to obtain an analytical solution for the equation but this nonlinear equation cannot be solved based upon elementary mathematical functions. We can, however, obtain an approximate solution by substituting the ***x*** term in the equation with a constant [GR09]:

$$\frac{dy}{dx} = a_1 y^2 + k^2 \tag{6.2.4}$$

At first glance we might expect that the value of ***k*** would be somewhere within the range of ***x*** values. Using the initial value y_0 as the value of **y** at ***x***=0, the solution for this equation is:

$$y(x) = k\, tan(kx + tan^{-1}\frac{y_0}{k}) \tag{6.2.5}$$

The smallest positive singular value of ***x*** occurs when the argument of the ***tan*** function is equal to $\pi/2$. So this singularity for Equation 6.2.4 is at:

$$x = \frac{1}{k}(\frac{\pi}{2} - tan^{-1}\frac{y_0}{k}) \tag{6.2.6}$$

In the examples shown in the figures above, the value of y_0 is one. Using Equation 6.2.3 and increasing x in increments of 0.001, the singularity was noted to be between 1.020 and 1.021 for values of a_1=0.9 and a_2=1. No value of k will yield a value of x greater than one so we see that Equation 6.2.6 gives only a hint as to the values of the singularities of Equation 6.2.2.

Another interesting result from simulations of experiments based upon Equation 6.2.2 is the effect of the range of x values upon the predicted values of σ_{a1}, σ_{a2} and $\sigma_f(x)$. Setting the range from x_1=**0** to x_n=*xmax* with Δx= *xmax*/(*n*-**1**), results are tabulated in Table 6.2.1 for n = 11 and CFU. Note that as the value of *xmax* approaches the singularity (between 1.020 and 1.021), the values of all the predicted σ's approach zero and the values of y_n and *COND* (the condition of the matrix) approach infinity. Clearly these results indicate that for the best results the experimenter should try to obtain a value of *xmax* as close to the singularity as possible. However, the increasing value of *COND* suggests that convergence problems might arise as *xmax* approaches the singularity.

xmax	y_n	$\sigma_f(xmax)/F$	σ_{a1}/F	σ_{a2}/F	*COND*
0.96	18.5	0.0541	0.0272	0.0340	$4.8*10^3$
0.98	27.7	0.0361	0.0175	0.0235	$1.0*10^4$
1.00	55.3	0.0181	0.0090	0.0138	$3.6*10^4$
1.02	3811.6	0.0003	0.0003	0.0033	$3.5*10^9$

Table 6.2.1 – Effect of Range upon the Prediction Analysis. Values of y measured to CFU ($\sigma_y / y = F$)

6.3 First Order ODE with an Analytical Solution

The simplest population model that includes decreasing per capita growth as the population increases is called the logistics population model. The following ODE was first proposed by Verhulst in 1838 [VE38]:

$$\frac{dy}{dt} = ry\left(1 - \frac{y}{K}\right) \tag{6.3.1}$$

The solution of this equation exhibits exponential growth while the population y is small but as the population reaches the limiting value of K the growth approaches zero. The solution for this equation is known as the logistic population model [BR01]:

$$y = \frac{K}{1 + (\frac{K}{y_0} - 1)e^{-rt}} \tag{6.3.2}$$

In this equation y_0 is the value of y at time $t = 0$. Equation 6.3.2 is valid for y_0 in the range $0 < y_0 < K$. To analyze this model, Brauer and Castillo-Chavez [BR01] used the U.S. census data from 1790 to 1990 (21 data points) shown in Table 6.3.1.

To test the model, a least squares analysis of the data in Table 6.3.1 can be performed. It is reasonable to assume that the values of σ_y are proportional to y (i.e., $\sigma_y = Cy$). The only question that must be answered is what is a reasonable value for the constant C? As explained in the discussion of Goodness-of-Fit in Section 3.9, the expected value of $S / (n-p)$ is 1 if the estimated values of σ_y are reasonable and the model is an accurate description of the phenomenon being modeled. Using a value of $C = 0.01$ the analysis yielded a value of $S / (n-p)$ of 35.5. One possible reason for this huge discrepancy is that the value of C was grossly underestimated. In fact, to reduce $S / (n-p)$ to one we would have to increase C by a factor of **sqrt**(35.5) or approximately 6. In other words, if we assumed that the accuracy of the data was approximately 6%, we would compute a value of $S/(n-p)$ of approximately one. Rerunning the least squares analysis using $\sigma_y = 0.06y$, yielded the results shown in Figure 6.3.1. The values of REL_ERROR in this figure are computed as (Y – CALC_VALUE) / SIGY.

We assumed that the initially large value of $S/(n-p)$ was due to an underestimate of C (the fractional error in the values of y). However, another possibility that should be considered is the validity of the logistics population model for the range of data included in Table 6.3.1. We can immediately see that the results are not realistic as the limiting value of population is 258.6 ± 13 million and this value has already been exceeded in more recent U.S. censuses and is continuing to grow. To accurately model the U.S. population one would need a more sophisticated model, however, let us consider this model as valid for other populations and for other time ranges.

One alternative to running the analysis as a model with 3 unknown parameters (i.e., K, r and y_0) is to treat y_0 as a known constant with a value of 3.9. Rerunning the analysis as a 2 parameter model yielded values K = 250 ± 11 and r = 0.0302 ± 0.00036. Comparing these results with the results in Figure 6.3.1 it is seen that the values of σ_K and σ_r are smaller but is this a real improvement in accuracy? If there is uncertainty in the value of y_0 then this uncertainty should be included in the analysis. For this particular problem do we really know that the population of the U.S. was exactly 3.9 million in 1790?

Year	*t*	*y*		*Year*	*t*	*Y*
1790	0	3.9		1900	110	76.0
1800	10	5.3		1910	120	92.0
1810	20	7.2		1920	130	105.7
1820	30	9.6		1930	140	122.8
1830	40	12.9		1940	150	131.7
1840	50	17.1		1950	160	150.7
1850	60	23.1		1960	170	179.0
1860	70	31.4		1970	180	205.0
1870	80	38.6		1980	190	226.5
1880	90	50.2		1990	200	248.7
1890	100	62.9				

Table 6.3.1 – U.S. Census Data from 1790 to 1990 (in millions)

```
PARAMETERS USED IN REGRESS ANALYSIS: Thu Feb 05, 2009

  N - Number of recs used to build model  :     21
  SYTYPE1 - Sigma type for Y              :      3
      TYPE 3:   SIGMA Y = CY1 * Y       CY1: 0.060
  Function Y:  K/(1+((K-Y0)/Y0)*EXP(-R*T))

ITERATION          K           R          Y0  S/(N.D.F.)
        0   250.0000     0.03000     3.90000     1.11862
        1   257.5599     0.02946     4.08560     0.98749
        2   258.4723     0.02945     4.08665     0.98694

 POINT      T        Y        SIGY  CALC_VALUE   REL_ERROR
     1      0     3.90     0.23400     4.08713    -0.79969
     2     10     5.30     0.31800     5.45699    -0.49366
     3     20     7.20     0.43200     7.27285    -0.16863
     4     30     9.60     0.57600     9.66987    -0.12131
     5     40    12.90     0.77400    12.81663     0.10771
     6     50    17.10     1.02600    16.91779     0.17759
     7     60    23.10     1.38600    22.21266     0.64022
     8     70    31.40     1.88400    28.96604     1.29191
     9     80    38.60     2.31600    37.44736     0.49769
    10     90    50.20     3.01200    47.89399     0.76561
    11    100    62.90     3.77400    60.45807     0.64704
    12    110    76.00     4.56000    75.14247     0.18806
    13    120    92.00     5.52000    91.74178     0.04678
    14    130   105.70     6.34200   109.81261    -0.64847
    15    140   122.80     7.36800   128.69687    -0.80034
    16    150   131.70     7.90200   147.60607    -2.01292
    17    160   150.70     9.04200   165.74759    -1.66419
    18    170   179.00    10.74000   182.45233    -0.32145
    19    180   205.00    12.30000   197.26244     0.62907
    20    190   226.50    13.59000   209.95833     1.21719
    21    200   248.70    14.92200   220.53163     1.88771

   PARAMETER INIT_VAL   MINIMUM   MAXIMUM      VALUE      SIGMA
           K   250.00   Not Spec  Not Spec  258.5610     13.021
           R     0.03   Not Spec  Not Spec    0.0294   0.000583
          Y0     3.90   Not Spec  Not Spec    4.0871     0.1316
 Variance Reduction:         98.63
 S/(N - P)          :       0.98694
 RMS (Y - Ycalc)    :       8.97145
 RMS ((Y-Ycalc)/Sy):       0.91975
```

Figure 6.3.1 – REGRESS Analysis of U.S. Census Data

To perform a prediction analysis of the logistic population model we first assume that CFU (constant fractional uncertainty) is valid (i.e., $\boldsymbol{\sigma_y = Cy}$). From Equation 6.3.2 the results should be functions of 2 dimensionless groups: the dimensionless range of values of $\boldsymbol{t}$ (i.e., $\boldsymbol{r\, t_{max}}$) and the dimensionless limiting value (i.e., $\boldsymbol{K / y_0}$). The results can be expressed as 3 dimensionless groups:

$$\Theta_K = \frac{\sigma_K}{K}\frac{n^{1/2}}{C} \qquad \Theta_r = \frac{\sigma_r}{r}\frac{n^{1/2}}{C} \qquad \Theta_{y0} = \frac{\sigma_{y0}}{y_0}\frac{n^{1/2}}{C}$$

The most interesting of these groups is Θ_K. This dimensionless number allows us to predict the accuracy to which the limiting value of the population will be determined. In Figure 6.3.2 values of Θ_K are shown as a function of the dimensionless range rt for several values of the ratio K / y_0.

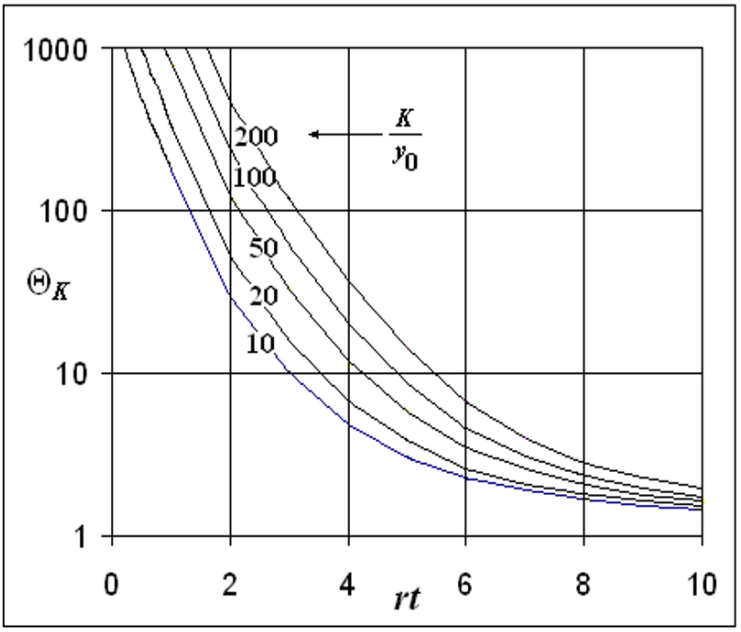

Figure 6.3.2 – Θ_K versus rt for several values of K / y_0

Example 6.3.1:

We wish to design an experiment in which we are trying to determine the limiting population of a type of bacteria (measured in units of millions of cells per ml). Our assumption is that Equation 6.3.2 is a valid model. Our preliminary estimates of K / y_0 and rt are 20 and 2 based upon a least squares analysis with n = 25 data points. The value of Θ_K from Figure 6.3.2 is between 10 and 100 and the more accurate value using the original data upon which the figure is based is 51.9. Using the resulting value of $S/(n-p)$ and then setting it to one, an estimate of the fractional uncertainty

of the data points of $C = 0.1$ is obtained. From the definition of Θ_K the predicted value of $\sigma_K / K = \Theta_K C/n^{1/2} = 51.9 * 0.1 / 5 = 1.04$. In other words, the predicted value of σ_K is approximately equal to K and the results from the least squares analysis verify this conclusion. If we double the time devoted to the measurement of K and take an additional 25 measurements, what is the predicted fractional uncertainty in the limiting population (i.e., σ_K / K)?

From Figure 6.3.2 if the value of rt is increased from 2 to 4, the predicted value of Θ_K for $K / y_0 = 20$ should be less than 10 (and using the original data it is 6.64). The predicted value of σ_K / K (using all 50 data points) is $6.64 * 0.1 / 50^{1/2} = 0.094$. In other words, doubling the time of the experiment and the number of data points should result in a decrease in σ_K / K by about a factor of 10!

Example 6.3.2:

In Example, 6.3.1 the predicted value of σ_K / K for $K / y_0 = 20$, $rt = 4$ and $n = 50$ was computed to be 0.094. What are the predicted values of σ_r / r and σ_{y0} / y_0 for this experiment? A prediction analysis for this experiment is shown in Figure 6.3.3. Note that the values of C, r and y_0 are all 1 and n=100 in the prediction analysis so the values of $\Theta_r = n^{1/2} \sigma_r/Cr = 100^{1/2}*0.300 = 3.00$ and $\Theta_{y0} = n^{1/2} \sigma_y/Cy_0 = 100^{1/2}*0.281 = 2.81$. Thus for the experiment discussed in Example 6.3.1 the predicted value of $\sigma_r /r = \Theta_r C/ n^{1/2} = 3.00*0.1/50^{1/2} = 0.042$ and $\sigma_{y0} / y_0 = 2.81 * 0.1/50^{1/2} = 0.040$. Checking the value of Θ_K: $\Theta_K = n^{1/2} \sigma_K/CK = 100^{1/2}*13.27/20 = 6.64$ which is the value used in Example 6.3.1.

```
PARAMETERS USED IN REGRESS ANALYSIS: Sun Feb 08, 2009

  INPUT PARMS FILE: fig633.par
  INPUT DATA  FILE: fig633.par
  REGRESS  VERSION: 4.23, Jan 27, 2009

  Prediction Analysis Option  (MODE='P')
  STARTREC - First record used            :     1
  N - Number of recs used to build model  :   100
  Y VALUES - computed using A0 & X values
  SYTYPE1 - Sigma type for Y              :     3
     TYPE 3:  SIGMA Y = CY1 * Y        CY1: 1.000
  T Values - computed using interpolation table
     T1=0.02  DELT=0.04

Analysis for Set 1
  Function Y:  K/(1+((K-Y0)/Y0)*EXP(-R*T))

   PARAMETER  INIT_VALUE       VALUE  PRED_SIGMA
           K    20.00000    20.00000    13.27388
           R     1.00000     1.00000     0.29984
          Y0     1.00000     1.00000     0.28137
```

Figure 6.3.3 – Prediction Analysis for Example 6.3.2

6.4 Simultaneous First Order Differential Equations

In the previous examples in this chapter the dependent variable y was a scalar. In this section experiments based upon simultaneous dependent variables are considered. A class of experiments that is modeled using simultaneous first order differential equations is the following chain reaction:

$$S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow\rightarrow\rightarrow S_d$$

Examples of such schemes are found in the decay of radioactive isotopes, in chemical kinetics and in many other areas of science and technology. The reaction is started with species S_1 decaying into species S_2 which decays into S_3 until the final product S_d is obtained. The value of d is the length of the chain. The dependent variable is thus a d dimensional vector with time dependent terms S_1, S_2, … S_d. The initial values of the various terms are S_{01}, S_{02}, … S_{0d}. The differential equation for the first term is:

$$\frac{dS_1}{dt} = -k_1 S_1 \tag{6.4.1}$$

The differential equations for the intermediate terms are:

$$\frac{dS_i}{dt} = k_{i-1}S_{i-1} - k_iS_i \tag{6.4.2}$$

The differential equation for the final term is:

$$\frac{dS_d}{dt} = k_{d-1}S_{d-1} \tag{6.4.3}$$

These equations are linear and can be solved analytically. However, as the chain gets larger the analytical solutions can be cumbersome. The solution for Equation 6.4.1 is always:

$$S_1 = S_{01}e^{-k_1t} \tag{6.4.4}$$

The solution for the intermediate terms can be expressed using the integral operator:

$$S_i = S_{0i} + \int_0^t (k_{i-1}S_{i-1} - k_iS_i)dt \tag{6.4.5}$$

The solution for the final term utilizes the fact that at $t = \infty$ S_d is the sum of all the initial values as all species eventually decay to the final product:

$$S_d = \sum_{i=1}^{i=d} S_{0i} - \sum_{i=1}^{i=d-1} S_i \tag{6.4.6}$$

Assuming that the initial values of all the species except S_1 are zero, the model has d unknown parameters: the d-1 reaction rates k_1 thru k_{d-1} and the initial value S_{01}. Assuming that the values of S_i are measured in numbers of counts, we can assume that $\sigma_{Si}^2 = S_i$ and the results can be expressed using the following dimensionless variables:

$$\Phi_{S_{01}} = \sigma_{S_{01}}\left[\frac{n}{S_{01}}\right]^{1/2} \qquad \Phi_{k_i} = \frac{\sigma_{k_i}}{k_i}(nS_{01})^{1/2} \quad i = 1 \text{ to } d\text{-}1$$

A typical plot if the values of S_i are shown in Figure 6.4.1 for the case where all three values of k_i are equal (i.e., k_i=k), d=4 and the dimensionless variable kt runs from 0 to 5. The value of S_{01} is 100 and all the other initial values are zero. The results of the prediction analysis for a similar run are shown in Figure 6.4.2. In this run the values of S_{01} and n are both 100 so that the ϕ_{S01} is equal to σ_{S01} and the 3 ϕ_{Ki}'s are equal to $100\sigma_{Ki}$. The data generated for this run is not included in Figure 6.4.2 but is shown graphically in 6.4.1. We see from Figure 6.4.2 that the predicted values of the σ_{ki} 's are 1.52, 2.00 and 2.43 and the value of $\sigma_{S_{01}}$ = 1.12. In Table 6.4.1 values of the ϕ's are listed for several values of kt_{max}. The results in this table obtained using S_{01} = n = 1000 and explain the slight differences with the results in Figure 6.4.2. Note that the optimum value of kt_{max} is different for the different k's.

kt_{max}	$\Phi_{S_{01}}$	Φ_{k_1}	Φ_{k_2}	Φ_{k_3}
1	1.13	1.95	5.36	21.54
3	1.15	1.43	2.12	3.28
5	1.11	1.53	2.00	2.44
7	1.08	1.72	2.17	2.45
9	1.06	1.92	2.41	2.65

Table 6.4.1 – ϕ's for various values of kt_{max} (all k's are equal)

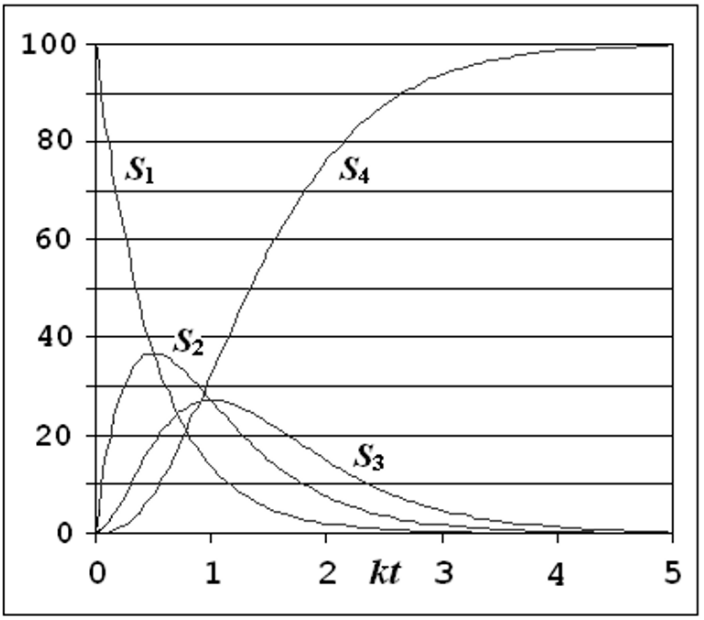

Figure 6.4.1 – S_i vs kt for case of $k_1 = k_2 = k_3 = k$

```
PARAMETERS USED IN REGRESS ANALYSIS: Thu Feb 19, 2009
  Prediction Analysis Option  (MODE='P')

  SYTYPE1 - Sigma type Si (i=1 to 4):
     SIGMA Si = CYi * sqrt(Si)          CYi: 1.000

  T Values - computed using interpolation table
     T0=0.025  DELT=0.05  N=100

  Function S1:  SO1 * EXP(-K1*T)
  Function S2:  SO2 + INT(S1*K1 - S2*K2,0,T)
  Function S3:  SO3 + INT(S2*K2 - S3*K3,0,T)
  Function S4:  SO1 + SO2 + SO3 +SO4 - S1 - S2 - S3

    K         CONSTANT        VALUE
    1              SO2      0.00000
    2              SO3      0.00000
    3              SO4      0.00000

   PARAMETER  INIT_VALUE       VALUE  PRED_SIGMA
         SO1   100.00000   100.00000     1.11515
          K1     1.00000     1.00000     0.01519
          K2     1.00000     1.00000     0.01999
          K3     1.00000     1.00000     0.02430
```

Figure 6.4.2 – Prediction Analysis for $kt_{max} = 5$

6.5 The Chemostat

The chemostat is a device found in laboratories studying chemical and biological phenomena. A detailed discussion of chemostats and their usage is included in books by Leah Edelstein-Keshet and Michael Greenberg [ED88, GR09]. A schematic diagram of a chemostat used in a biological experiment is shown in Figure 6.5.1. The volume V of laboratory chemostats can range from about 0.5 liters to over 10 liters. Industrial chemostats can exceed 1000 m^3. The C variable is the concentration of a nutrient and the N variable is the number of the microorganism strain of interest within the chamber. The chemostat includes a stirring device to reduce the spatial effects within the chamber. The purpose of the experiment might be to study the dynamics of this biological process or the equilibrium state (i.e., when the time variable t is large).

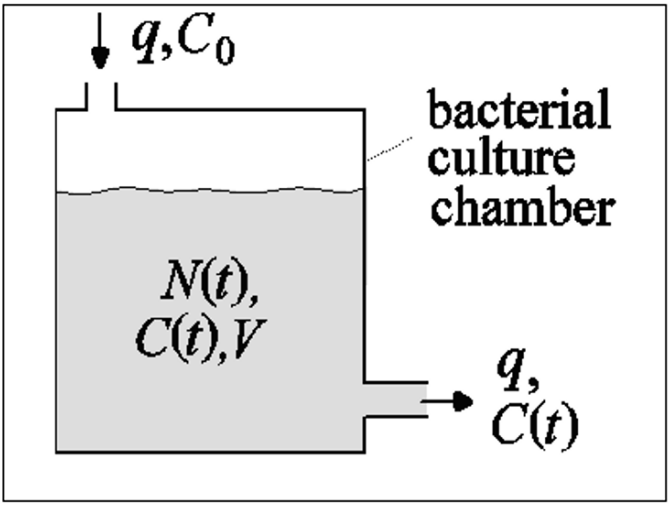

Figure 6.5.1 – Schematic Diagram of a Chemostat

Two first order initial value differential equations are used to model this system:

$$\frac{dN}{dt} = k(C)N - \frac{q}{V}N \tag{6.5.1}$$

$$\frac{dC}{dt} = -\alpha k(C)N - \frac{q}{V}C + \frac{q}{V}C_0 \tag{6.5.2}$$

The $k(C)$ term is growth rate for the N variable and is dependent upon the nutrient concentration. Edelstein-Keshet proposes Michaelis-Menten kinetics as the model for $k(C)$:

$$k(C) = \frac{\beta C}{\gamma + C} \tag{6.5.3}$$

This model is shown in Figure 6.5.2 and is based upon the assumption that there is a saturation value of k equal to β which is the growth rate when C is very large. The parameter γ is the value of C that causes k to equal $\beta/2$.

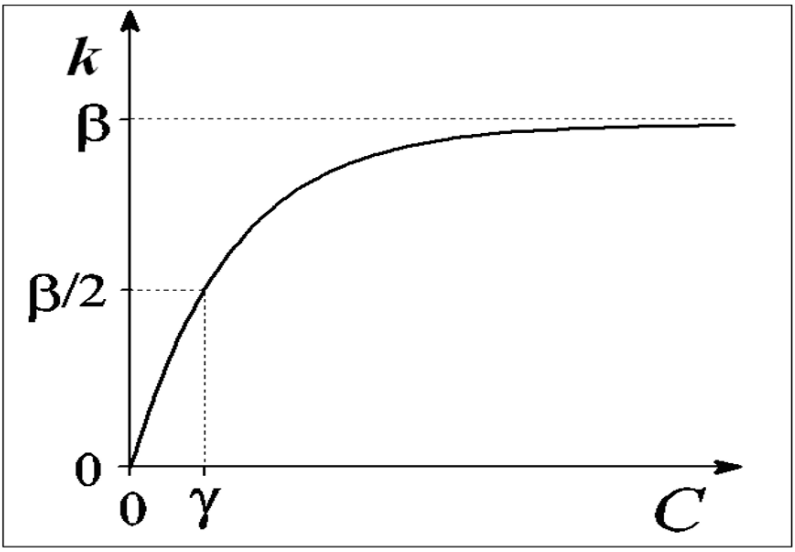

Figure 6.5.2 – The Michaelis-Menten kinetics model

In this section an experiment to determine the parameters α, β, and γ is considered. An analysis of this set of equations shows that the asymptotic value of N can be either 0 (i.e., the microorganisms all die) or a non-zero value that is a function of the parameters of the system: q (the flow rate), V (the chamber volume), C_0 (the input nutrient concentration), and the three coefficients: α, β, and γ. Greenberg [GR09] defines two parameters that are combinations of the system parameters:

$$\delta = \frac{\beta V}{q}, \quad \varepsilon = \frac{C_0}{\gamma} \tag{6.5.4}$$

The purpose of the chemostat (both in the laboratory and in industrial applications) is to maintain a stable constant population level of the culture. The conditions that must be satisfied so that N reaches a non-zero equilibrium state are:

$$\delta > 1, \quad \varepsilon > \frac{1}{\delta - 1} \tag{6.5.5}$$

This region of stability is shown in Figure 6.5.3 (the shaded zone). When these conditions are satisfied, the equilibrium values of N and C are:

$$\frac{\alpha\beta V}{\gamma q} N = \frac{\delta}{\delta - 1}(\varepsilon\delta - \varepsilon - 1) \tag{6.5.6}$$

$$\frac{C}{\gamma} = \frac{1}{\delta - 1} \tag{6.5.7}$$

Substituting 6.5.4 into 6.5.6 and 6.5.7 and gathering terms:

$$N = \frac{\gamma Q / \alpha\beta}{1 - Q/\beta}\left(\frac{C_0}{\gamma}\left(\frac{\beta}{Q} - 1\right) - 1\right) \tag{6.5.8}$$

$$C = \frac{\gamma}{(\beta / Q) - 1} \tag{6.5.9}$$

In these equation since V is constant, q/V is replaced by Q and has the units of $time^{-1}$.

We are now in a position to propose an experiment to determine the values of the parameters α, β, and γ. We can treat C_0 and Q as our independent variables and N and C as the dependent variables. Varying the independent variables over a range that satisfies the conditions specified by Equation 6.5.5 the values of N and C are measured and the coefficients can then be determined using a 3 parameter least squares analysis.

An example of a prediction analysis using 9 pairs of values of C_0 and Q is shown in Figure 6.5.4. The pairs chosen satisfy the two conditions specified in Equation 6.5.5. For this prediction analysis the values of α, β, and γ were all set to one and a constant fractional error of 100% was used (i.e., $\sigma_N = N$ and $\sigma_C = C$). The predicted results are proportional to the expected fractional error of the measurements. For example, if both N and C are measured to 10% accuracy, then the predicted values of σ_α, σ_β and σ_γ would be one-tenth the PRED_SIGMA values shown in Figure 6.5.4.

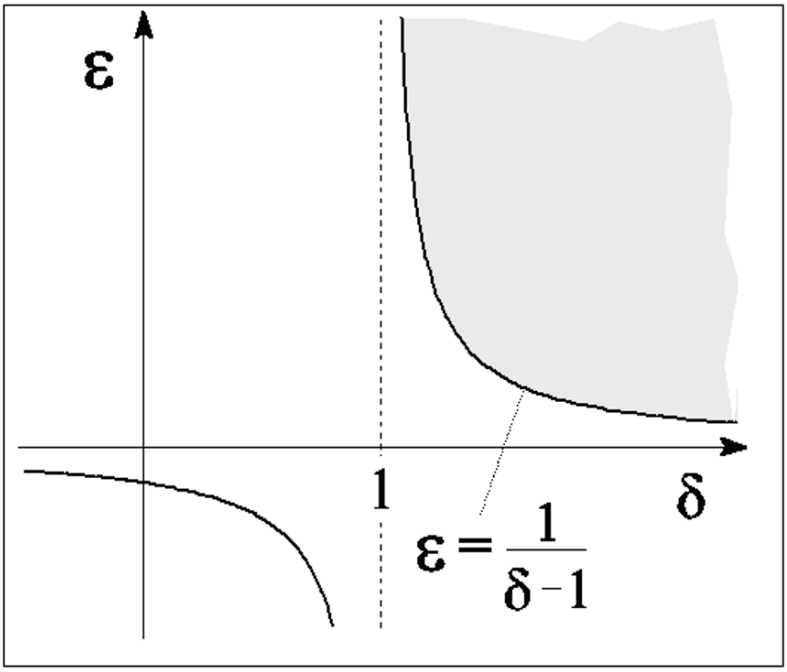

Figure 6.5.3 – The Region of Stability (the shaded zone)

The prediction analysis shown in Figure 6.5.4 is based upon 9 data points that include three values of $\boldsymbol{Q}$ and for each $\boldsymbol{Q}$, 3 values of $\boldsymbol{C_0}$. These points were chosen arbitrarily to demonstrate the process. One does not necessarily have to vary both $\boldsymbol{C_0}$ and $\boldsymbol{Q}$ to determine the coefficients. For example, the value of $\boldsymbol{C_0}$ can be held constant while $\boldsymbol{Q}$ is varied. One will obtain a least squares solution. However, if $\boldsymbol{Q}$ is held constant and only $\boldsymbol{C_0}$ is varied, the prediction analysis yields indeterminate values for $\boldsymbol{\sigma_\alpha}$, $\boldsymbol{\sigma_\beta}$ and $\boldsymbol{\sigma_\gamma}$. An explanation for this phenomenon is seen in Equation 6.5.9. Since the equilibrium value of $\boldsymbol{C}$ is only a function of $\boldsymbol{Q}$ and not of $\boldsymbol{C_0}$ all values of $\boldsymbol{C}$ will be the same. From Figure 6.5.2 it is obvious that estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ cannot be made from a single value of $\boldsymbol{C}$. This is an interesting point. If real data were obtained based upon a constant value of $\boldsymbol{Q}$ and a series of measurements taken for various values of $\boldsymbol{C_0}$, then the measured values of $\boldsymbol{C}$ would not be exactly the same (due to random variations) and a least squares analysis could be performed. However, one could expect that the computed values of the parameters would be subject to very large uncertainties (i.e., $\boldsymbol{\sigma_\alpha}$, $\boldsymbol{\sigma_\beta}$ and $\boldsymbol{\sigma_\gamma}$). Once again, looking at Figure 6.5.2 one can see that if all the values of $\boldsymbol{C}$ are concentrated in one portion of the curve, the computed values of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ will not be very accurate.

```
PARAMETERS USED IN REGRESS ANALYSIS: Tue Feb 24, 2009
  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  :      9
  NCOL - Number of data columns           :      2
  NB & C VALUES - computed using C0 & Q values
  SYTYPE1 - Sigma type for NB             :      3
     TYPE 3:  SIGMA NB = CY1 * NB      CY1:  1.00
  SYTYPE2 - Sigma type for C              :      3
     TYPE 3:  SIGMA C = CY2 * C        CY2:  1.00

Analysis for Set 1
  Function NB:(GAMMA*Q/ALPHA*BETA)*((C0/GAMMA)*(BETA/Q-1)-1)/(1-Q/BETA)
  Function C: GAMMA/((BETA/Q)-1)

REC  Y-INDEX        Q        C0         NB      SIGNB  CALC_VALUE
   1       1    0.900    11.000      2.000      2.000      2.000
   2       1    0.900    10.000      1.000      1.000      1.000
   3       1    0.900     9.500      0.500      0.500      0.500
   4       1    0.800     6.000      2.000      2.000      2.000
   5       1    0.800     5.000      1.000      1.000      1.000
   6       1    0.800     4.500      0.500      0.500      0.500
   7       1    0.700     5.000      2.667      2.667      2.666
   8       1    0.700     4.000      1.667      1.667      1.667
   9       1    0.700     3.500      1.167      1.167      1.167

 REC Y-INDEX        Q        C0          C       SIGC  CALC_VALUE
   1       2    0.900    11.000      9.000      9.000      9.000
   2       2    0.900    10.000      9.000      9.000      9.000
   3       2    0.900     9.500      9.000      9.000      9.000
   4       2    0.800     6.000      4.000      4.000      4.000
   5       2    0.800     5.000      4.000      4.000      4.000
   6       2    0.800     4.500      4.000      4.000      4.000
   7       2    0.700     5.000      2.333      2.333      2.333
   8       2    0.700     4.000      2.333      2.333      2.333
   9       2    0.700     3.500      2.333      2.333      2.333

   PARAMETER     VALUE  PRED_SIGMA
       ALPHA     1.000     0.54640
        BETA     1.000     0.02564
       GAMMA     1.000     0.26857
```

Figure 6.5.4 – Prediction Analysis of a Chemostat Experiment

No attempt is made in this section to present results for a wide variety of combinations of the variables. The number of degrees of freedom for the general chemostat problem is large and the purpose of this section is to show how one would go about designing a specific chemostat experiment.

6.6 Astronomical Observations using Kepler's Laws

The history of astronomy is thousands of years old [KO95, KO04, HA00]. Observations of the sky include positions of stars and other heavenly bodies by cataloging angles from very distant objects. One of the earliest star catalogues was compiled by Hipparcos in about 190 BC. His catalogue included about 850 stars. Ptolomy developed the ***Almagest*** several centuries after Hipparchus. It is the only surviving ancient treatise on astronomy

and includes a catalogue of 1022 stars. He died in 168 AD. In the 15th century Ulugh Beg (1393 – 1449) compiled a catalogue of 1019 stars which is accurate to about 20 minutes of arc. In the following century Tycho Brahe (1546 to 1601) developed instruments capable of measuring star positions to accuracies of 15 to 35 seconds of arc (arcsecs). In 1729 James Bradley measured stellar parallax to less than one arcsec of accuracy.

Modern astronomy has made incredible strides in recent years. The use of CCD's (charge-coupled devices) has made possible observations with accuracies measured in milliarcsecs. In addition, satellite observations and ground-based telescopes with adaptive optics have reduced the distortions created by Earth's atmosphere and gravitation. The Hipparcos satellite was launched in 1989 by the European Space Agency and its observations over a 4 year period were used to create a new star catalogue called Tycho (named after Tycho Brahe). The positions of 118,218 stars were determined to accuracies in the 20 to 30 milliarcsec range.

One of the pivotal moments in astronomy was the publication of Johannes Kepler's three laws of planetary motion. Kepler was born in 1571 and died in 1630. He studied the motion of planets orbiting about the sun and derived a theory that would explain the orbits. He used the Tycho Brahe data to test his theory. Prior to Kepler, the prevailing assumption was that the motion should be circular but Kepler showed that in reality it is elliptical with one of the two foci of the ellipse deep within the sun. Kepler laws are stated in terms of planets revolving around the sun, however, Newton (1642 – 1727) generalized Kepler's laws to be applicable to rotation of any two bodies of any mass ratio revolving around each other.

The three laws of Kepler are as follows:

First law: The orbit of every planet is an ellipse with the sun at a focus. The generalized version of this law is that the motion of each body is an ellipse about the 2nd body with one of the foci at the barycenter of the two bodies. The barycenter is defined as the point at which the gravitational forces are equal. For planets revolving around the sun, the barycenter is located deep within the sun but not exactly at its center of mass. (Remember that the sun's gravity at its very center of mass is zero!)

Second law: The line joining a planet and the sun sweep out equal areas during equal intervals of time. The generalized version of this law is that the line joining one body and the barycenter sweeps out equal areas during equal time intervals.

Third law: The square of the orbital period of a planet is directly proportional to the cube of the semi-major axis of its orbit. In general this law is valid for any body revolving around any other body.

To determine orbits, astronomers collect data in the form of angular positions in the celestial sky of a body under observation as a function of time. The position in the sky is corrected to where it would have been on a particular date and time. The currently used standard is Jan 1, 2000 at 12:00 TT (Terestial Time). Two angles are measured: *right ascension* and *declination*. These angles are equivalent to longitude and latitude. Changes in *declination* indicate that the ellipse of the body under observation is tilted with respect to the plane in which earth rotates around the sun.

The mathematical treatment of Kepler's laws are best described using polar coordinates. Figure 6.6.1 is a schematic diagram of the motion of one body around a second body. Using the center of the polar coordinate system as one of the two foci of the ellipse, every point on the ellipse is denoted by two coordinates: r and Θ. The angular coordinate Θ ranges from 0 to 2π radians and the radial coordinate r ranges from r_{min} at angle $\Theta = 0$ to r_{max} at angle $\Theta = \pi$. Observations are not made directly in the polar coordinate system. What are measured are changes in the position of the rotating body from the initial position. Assuming that the initial observation is located at an unknown angle Θ_0 in the elliptic orbit, the subsequent observations can be transformed into the polar coordinate system as changes from the angle Θ_0. The length a is the separation between the two foci and is also half the length of the semi-major axis (i.e., $(r_{min} + r_{max}) / 2$). The length b is the semi-minor axis and is computed as $\mathbf{sqrt}(r_{min}r_{max}) = p / \mathbf{sqrt}(1-\varepsilon^2)$. The symbol ε denotes the eccentricity of the ellipse and is in the range 0 to less than one. The length p can be related to r_{min}, r_{max} and a as follows: $p = (1+\varepsilon)\, r_{min} = (1-\varepsilon)\, r_{max} = a(1-\varepsilon^2)$.

Kepler's first law:

$$r = \frac{p}{1 + \varepsilon \cos(\Theta)} \tag{6.6.1}$$

Kepler's second law can be restated as angular momentum l is constant**:**

$$l = r^2 \frac{d\Theta}{dt} = \left(\frac{p}{G(M+m)} \right)^{1/2} \quad \text{(where } l \text{ is constant)} \qquad (6.6.2)$$

The constants G, M and m are the gravitational constant, the mass of the sun and the mass of the body rotating about the sun.

Kepler's third law:

$$T^2 \propto a^3 \qquad (6.6.3)$$

In this equation T is the period (time for a complete rotation) and increases as $a^{3/2}$. For example, the value of a for earth is one AU (*Astronomical Unit*) which is the average distance between the earth and the sun) and is about 1.5237 for Mars. Thus Mars rotates about the sun with an orbital period of $1.5237^{1.5} = 1.8809$ earth years.

The changes in angles in the celestial sky can be used to compute changes in Θ (i.e., changes in the angles in the plane of rotation) from Θ_0. The values of the 3 unknown parameters p, ε and Θ_0 can be computed using the method of least squares based upon Equations 6.6.4 and 6.6.5:

Since $r(t) = \dfrac{p}{1+\varepsilon\cos(\Theta_0)} + \displaystyle\int_0^t \frac{dr}{d\Theta}\frac{d\Theta}{dt}dt$ we obtain:

$$r(t) = \frac{p}{1+\varepsilon\cos(\Theta_0)} + \sqrt{\frac{p}{G(M+m)}}\int_0^t \frac{p\varepsilon\sin(\Theta)}{(1+\varepsilon\cos(\Theta))^2 r^2}dt \qquad (6.6.4)$$

$$\Theta(t) = \Theta_0 + \sqrt{\frac{p}{G(M+m)}}\int_0^t \frac{1}{r^2}dt \qquad (6.6.5)$$

Note that these equations are recursive as Θ is required to compute r, and r is required to compute Θ.

As an example of the use of these equations in a prediction analysis let us consider the calculation of the orbit of Ceres. In 1801 Giuseppe Piazzi discovered a moving object in the sky which at first he thought might be a comet. He then made an additional 18 observations of the object over a 42

day period. Eventually he realized that the object was rotating around the sun and named it Ceres. During these 42 days the Earth rotated almost 42 degrees around the sun and at the end of this time period Ceres could no longer be observed. Subsequently Gauss analyzed the data and predicted the orbit of Ceres using the method of least squares which he developed as part of his analysis [GA57]. From this orbit he also successfully predicted when Ceres would reappear.

To develop a prediction analysis of an experiment similar to that of Piazzi, let us assume that the changes in Θ are measured to the accuracy of the instrument used by Piazzi (i.e., about 25 arcsecs = 0.006944 degrees). Assume that the 18 changes in angles are measured over a 42 day period with equal changes in time. (This differs from Piazzi's experiment because the times between his observations were not constant.) We know that r_{min} (called the Perihelion) of Ceres is 2.544 AU and r_{max} (the Aphelion) is 2.987 AU. The value of a is therefore (2.544 + 2.987) / 2 = 2.7655 AU. The eccentricity ε of Ceres's orbit is known to be about 0.080. From a we can compute $p = a(1-\varepsilon^2) = 2.7478$. We also know that the orbit that Ceres makes about the sun requires 4.599 years. A prediction analysis based upon Equations 6.6.4 and 6.6.5 with values of p=2.7478, ε=0.08 and Θ_0=45^0 is shown in Figure 6.6.2. The constant GM=0.02533 was determined by trial-and-error so that a complete revolution of earth around the sun is made in one years (using ε = 0.016710219 and a = 1 AU to describe the earth orbit). Equations 6.6.4 and 6.6.5 were modified so that angles are expressed in degrees rather than radians using the conversion constant **c** (from degrees to radians). The 18 time changes were each 42/365/18 = 0.006393 years. The constant CY1 = 0.006944 is the estimated uncertainty of the angular measurements in degrees (i.e., 25 arcsecs). Note that CY2 and CY3 (i.e., σ_Θ and σ_r) are set to zero because these are computed and not measured variables.

The results of the simulation in Figure 6.6.2 show the predicted value of σ_ε = 0.00556, the predicted value of σ_p = 0.01293 and the predicted value of $\sigma_{\Theta 0}$ = 0.00227. The starting point (i.e., first observation of Piazzi's measurements) was arbitrarily assumed to be 45^0 from the zero point of the Ceres ellipse. We can ask the following question: If we could choose any starting point, how will the resulting accuracies of the measured 3 parameters (i.e., p, ε and Θ_0) be effected? The prediction analysis was repeated for different values of Θ_0 and the results are in shown in Table 6.6.1. The value of $\Delta\Theta$ included in the table is the total angular change over the 42 day period. The most interesting parameter is p because it is directly re-

lated to the length of the major axis and the period of the orbit. We see from the results in the table that the most accurate measurements of $\boldsymbol{p}$ are achieved if the data is obtained in the region near $\boldsymbol{\Theta} = 90^0$. The values of $\boldsymbol{\Delta\Theta}$ in the table confirm that the change in angle is greatest near the point where $\boldsymbol{\Theta} = 0^0$ and least where $\boldsymbol{\Theta} = 180^0$. Based upon the known period of Ceres we can compute the average change for any 42 day period as $360^0*42/(365*4.6) = 9.005^0$.

Θ_0	σ_ε	σ_p	$\sigma_{\Theta 0}$	$\Delta\Theta$
0	0.00643	0.02223	0.00231	10.600
45	0.00556	0.01293	0.00227	10.054
90	0.00872	0.00222	0.00215	8.983
135	0.01370	0.04078	0.00217	8.031
180	0.00768	0.03121	0.00231	7.702

Table 6.6.1 Predicted accuracies of simulated experiment to determine the orbit of Ceres

Another interesting question relates to the total time of the observations (i.e., from the first to the last observation). In Table 6.6.2 results are included for times periods of 21, 42 and 84 days using $\boldsymbol{\Theta_0} = 45^0$ and $\boldsymbol{n} = 18$ (the number of angular changes taken during the time period). Increasing the time period of the observations will dramatically improve the measurements of $\boldsymbol{p}$ and $\boldsymbol{\varepsilon}$ but will have a very negligible effect on the determination of $\boldsymbol{\Theta_0}$. We can also vary $\boldsymbol{n}$. As expected, increasing $\boldsymbol{n}$ while maintaining the total time as a constant, the values of the predicted $\boldsymbol{\sigma}$'s are inversely proportional to $\boldsymbol{n^{1/2}}$. In other words, increasing $\boldsymbol{n}$ by a factor of 4 should reduce the resulting $\boldsymbol{\sigma}$'s by a factor of 2.

Total days	σ_ε	σ_p	$\sigma_{\Theta 0}$	$\Delta\Theta$
21	0.01460	0.03531	0.00229	5.052
42	0.00556	0.01293	0.00227	10.054
84	0.00178	0.00380	0.00222	19.894

Table 6.6.2 Predicted accuracies of simulated experiment to determine the orbit of the Ceres. Varying *n* for $\Theta_0 = 45^0$.

Example 6.6.1:

If Piazzi hadn't been ill during the 42 day period that he observed Ceres, and if the weather had been ideal so that he could have taken measurements on all 42 days, how much better would the resulting **σ**'s have been?

Over the 42 day period, 41 angle changes could have been measured and the **σ**'s would have been reduced by a factor of **sqrt**(41/18) = 1.51.

Example 6.6.2:

If measurements of the location of Ceres is taken daily over an 84 day period using modern-day instruments that are accurate to 25 milliarcsecs, how accurately would $\boldsymbol{p}$ be measured if $\Theta_0 = 45^0$?

From Table 6.6.2 we see that the predicted value of σ_p is 0.0038 if 18 changes in angle are measured over an 84 day period using an instrument with measurement accuracy of 25 arcsecs. Improving the accuracy of the measurements by a factor of 1000 and increasing $\boldsymbol{n}$ from 18 to 83, the accuracy should be improved by a factor of 1000 ***sqrt**(83/18) = 2150. Thus the value of σ_p should be about 0.0038 / 2150 = 0.00000177 AU. Since one AU is 149,597,871 Km, an error of this amount (i.e., σ_p) would be about 264 Km.

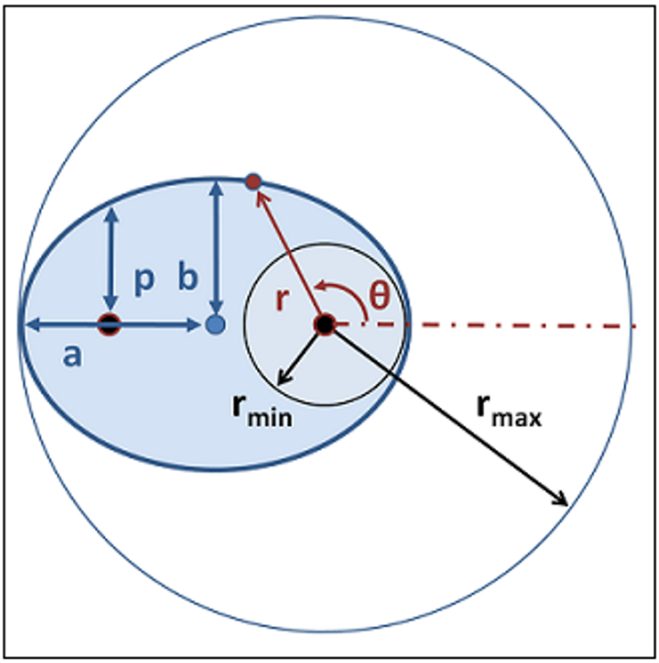

Figure 6.6.1 Diagram of the orbit of one body rotating around a second body (from Wikipedia: Kepler's laws of planetary motion)

```
PARAMETERS USED IN REGRESS ANALYSIS: Sun Nov 01, 2009
  INPUT PARMS FILE: cerces.par
  INPUT DATA  FILE: cerces.par
  REGRESS  VERSION: 4.27, May 5, 2009

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  :     18
  P, E, and THETA0 VALUES - computed using R & THETA
  SIGMA DTH = CY1                      CY1: 0.006944
  SIGMA THETA = CY2                    CY2:     0.000
  SIGMA R = CY3                        CY3:     0.000
  T Values - computed using interpolation table
       NP1=18  X01=0.006393  DELX1=0.006393
  MAXDEPTH - Max depth in INT scheme       :     10
  INTEPS - Integration converge criterion 0.00100

  Function DTH:     SQRT(P/GM)*INT(1/R^2,0,T)/C
  Function THETA:   THETA0 + DTH
  Function R:   P/(1 + E*COS(C*THETA0)) + SQRT(P/GM)*
    INT(P*E*SIN(C*THETA)/(R^2*(1+E*COS(C*THETA))^2),0,T)
    K         CONSTANT          VALUE
    1               GM        0.02533
    2                C    0.017453292 (degrees to radians)
 REC           T          DTH       THETA          R
   1     0.00639      0.56383     45.56383     2.60206
   2     0.01279      1.12707     46.12707     2.60345
   3     0.01918      1.68970     46.68970     2.60486
   4     0.02557      2.25172     47.25172     2.60627
   5     0.03197      2.81313     47.81313     2.60770
   6     0.03836      3.37393     48.37393     2.60915
   7     0.04475      3.93410     48.93410     2.61060
   8     0.05114      4.49364     49.49364     2.61207
   9     0.05754      5.05255     50.05255     2.61355
  10     0.06393      5.61083     50.61083     2.61504
  11     0.07032      6.16847     51.16847     2.61655
  12     0.07672      6.72547     51.72547     2.61806
  13     0.08311      7.28182     52.28182     2.61959
  14     0.08950      7.83752     52.83752     2.62113
  15     0.09589      8.39257     53.39257     2.62268
  16     0.10229      8.94696     53.94696     2.62424
  17     0.10868      9.50068     54.50068     2.62581
  18     0.11507     10.05375     55.05375     2.62740

 PARAMETER  INIT_VALUE        VALUE  PRED_SIGMA
         E     0.08000      0.08000     0.00556
         P     2.74780      2.74780     0.01293
    THETA0    45.00000     45.00000     0.00227
```

Figure 6.6.2 Prediction Analysis of an experiment which simulates Piazzi's observations of Ceres.

Chapter 7 RANDOM DISTRIBUTIONS

7.1 Introduction

For many of the experiments discussed in the preceding chapters there was an implied assumption: the experimenter could control the values of the independent variables. Often the assumption was that the independent variable x could be varied from x_1 to x_n with a constant value of $\Delta x = x_{i+1} - x_i$. For experiments with more than one independent variable it was usually assumed that the independent variable space could be spanned in a similar manner. For many experiments this is a reasonable assumption but there are innumerable experiments where this assumption is unrealistic. For example, experiments based upon the measurement of the effects of temperature upon some variable might or might not be controllable. If, for example, the temperature is the atmospheric temperature at the time the variable is measured then this temperature can only be observed. It cannot be controlled. Examples from the medical field are experiments related to the effects of cholesterol upon a variable that is a measure of heart function. The amount of cholesterol in a subject of the experiment is measured but not controlled. All that can be said for the subjects when taken as a group is that the amount of cholesterol in the blood of the subjects is distributed according to some known (or observed) distribution. If the experiment is designed properly, then the subjects taken as a group are representative of similar groups of subjects that are being modeled.

The subject of distributions was discussed in Section 2.4. Many distributions of the independent variable x can be approximated by a normal distribution when the number of samples n is large. Another simple approximation is a uniform distribution of the value of x within the range $xmin$ to $xmax$ (i.e., all values of $\Phi(x)$ are equal within this range). To design experiments based upon randomly distributed independent variables, one must generate values of the independent variable (or variables) based upon the representative distribution. For example, if we can assume that x is nor-

J. Wolberg, *Designing Quantitative Experiments*, DOI 10.1007/978-3-642-11589-9_7,

mally distributed then values of $\boldsymbol{x}$ conforming to this distribution can be generated and the prediction analysis can then be performed.

The REGRESS program includes the ability to generate randomly distributed variables. This feature is used in the following sections to perform prediction analyses on several common types of experiments. In Section 7.2 the multiple linear regression problem is revisited but this time with randomly distributed independent variables. In Section 7.3 experiments based upon the bivariate normal distribution are discussed. In Section 7.4 the value of orthogonally distributed $\boldsymbol{x}$'s is considered.

7.2 Revisiting Multiple Linear Regression

The multiple linear regression problem was discussed in Section 4.10. The analysis was based upon the assumption that independent variables $\boldsymbol{x_j}$ could be controlled and values of $\boldsymbol{x_{ji}}$ were generated using $\boldsymbol{x_{j1}}$ and $\boldsymbol{\Delta x_j}$ and Equation 4.10.2. In this section we assume that the values of $\boldsymbol{x_{ji}}$ cannot be controlled but they are generated from known random distributions.

Clearly if the assumption is that all values within the specified ranges are equally probable, the predicted values of all the $\boldsymbol{\sigma}$'s will approach the same values (for large $\boldsymbol{n}$) as obtained in Section 4.10 for the same volumes. In Figure 7.2.1 we see results from a prediction analysis similar to the analysis shown in Figure 4.10.1. The only difference between these two analyses is that in 4.10.1 the $\boldsymbol{x_{ji}}$ are controlled and in 7.2.1 they are generated randomly within the same volume as in Figure 4.10.1. The probability of a data point falling on any point within the volume is equally probable. Comparing the results, as expected, the values of the $\boldsymbol{\sigma_{aj}}$'s (i.e., PRED_SA(K)) are very close.

If the random distributions are normal, then the results are different than those obtained for a uniform distribution. We start the analysis using the same mathematical model as in Section 4.10:

$$\boldsymbol{y = a_1 + a_2 x_1 + ... + a_{d+1} x_d} \qquad (4.10.1)$$

```
PARAMETERS USED IN REGRESS ANALYSIS: Sun Mar 22, 2009

  INPUT PARMS FILE: fig721.par
  INPUT DATA  FILE: fig721.par
  REGRESS  VERSION: 4.25, Mar 18, 2009

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  : 10000
  Y VALUES - computed using A0 & X values
  SYTYPE1 - Sigma type for Y              :     1
     TYPE 1:  SIGMA Y = 1
  M - Number of independent variables     :     3
  X1 Values - Random Distribution:  0.000 to  1.000
  X2 Values - Random Distribution:  0.000 to  2.000
  X3 Values - Random Distribution:  0.000 to  1.000

Analysis for Set 1
  Function Y:  A1+A2*X1+A3*X2+A4*X3

   K       A0(K)         A(K)  PRED_SA(K)
   1     1.00000      1.00000     0.03173
   2     2.00000      2.00000     0.03463
   3     3.00000      3.00000     0.01748
   4     0.00000      0.00000     0.03454
```

Figure 7.2.1 - Prediction Analysis of 3D plane model (Note how close these results are to Figure 4.10.1)

The assumption of normally distributed x_j's requires two parameters for each x_j : μ_j and σ_j. In Section 4.10 Equations 4.10.5 and 4.10.6 were used to define the dimensionless groups used to present the results. Equation 4.10.5 can also be used for the case of normally distributed x's but Equation 4.10.6 is replaced by 7.2.1:

$$\Phi_1 \equiv \frac{\sigma_{a_1} n^{1/2}}{\sigma_y} \tag{4.10.5}$$

$$\Phi_{j+1} \equiv \frac{\sigma_{a_{j+1}} n^{1/2} \sigma_j}{\sigma_y} \qquad j = 1 \text{ to } d \tag{7.2.1}$$

Once again the mathematics is simplified if we consider the cases in which all the values of μ_j are zero. Any set of experimental data can be transformed to this specific case by subtracting out the values of μ_j from the

values of the x_j 's. For these cases we can show that the inverse C matrix is:

$$C^{-1} = \frac{1}{n}\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1/\sigma_1^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1/\sigma_d^2 \end{bmatrix} \tag{7.2.2}$$

From Equation 4.2.1 we see that the values of $\sigma_{a_j}^2$ are just the diagonal terms of this inverse matrix. Substituting these terms into Equations 4.10.5 and 7.2.1 we get the following:

$$\Phi_1 \equiv \frac{\sigma_{a_1} n^{1/2}}{\sigma_y} = 1 \quad \text{(if all the } \mu_j \text{'s are zero)} \tag{7.2.3}$$

$$\Phi_{j+1} \equiv \frac{\sigma_{a_{j+1}} n^{1/2} \sigma_j}{\sigma_y} = 1 \qquad j = 1 \text{ to } d \tag{7.2.4}$$

If the μ_j 's are not zero, only Φ_1 changes:

$$\Phi_1 \equiv \frac{\sigma_{a_1} n^{1/2}}{\sigma_y} = 1 + \sum_{j=1}^{j=d} (\mu_j / \sigma_j)^2 \tag{7.2.5}$$

If the purpose of the experiment is to compute the values of y at any point in the d dimensional space, the values of σ_f^2 are computed using Equation 4.2.2. It can be shown that this equation reduces to the following simple form:

$$\frac{\sigma_f^2 n}{\sigma_y^2} = 1 + \sum_{j=1}^{j=d} \frac{(x_j - x_{javg})^2}{(\sigma_j)^2} \tag{7.2.6}$$

Example 7.2.1:

Assume that we wish to perform a multiple linear regression on a 3 dimensional space. Assume that Equation 4.10.1 is a realistic representation of the dependence of the values of y upon the values of x_1, x_2 and x_3. Assume that we can measure each value of y to an accuracy of σ_y and that we can assume that the values of the x's are error free. The purpose of the experiment is to predict values of y for any point within the space. We want to design the experiment so that within the range $|x_{ji} - x_{javg}| <= 2\sigma_j$ all values of y will be computed to an accuracy of 0.01 (i.e., $\sigma_f <= 0.01$).

From Equation 7.2.6, within the design region, the maximum values of σ_f are :

$$\frac{\sigma_{f\,max}^2 n}{\sigma_y^2} = 1 + \sum_{j=1}^{j=3} \frac{(x_{j\,max} - x_{javg})^2}{(\sigma_j)^2} = 1 + \sum_{j=1}^{j=3} 2^2 = 13$$

Setting σ_{fmax} to 0.01 we get the following design requirement:

$$\frac{\sigma_{f\,max}^2 n}{\sigma_y^2} = 13 = \frac{0.01^2 n}{\sigma_y^2} \Rightarrow \frac{n}{\sigma_y^2} = 130{,}000$$

From this equation we see that if we can only measure the values of y to an accuracy of σ_y = 1, then we would require 130,000 randomly distributed points to satisfy the design requirement. However, if we could measure the values of y to an accuracy of σ_y = 0.1, then the number of points required is reduced to 1300. If we relax our requirement regarding σ_f from 0.01 to 0.02 then the value of n is reduced by an additional factor of 4.

Example 7.2.2:

In Example 7.2.1 with the requirement σ_{fmax} = 0.01 it was seen that 1300 points were required to achieve the desired result if the values of y could be measured to an accuracy of σ_y = 0.1. Assuming that the values of the x's are adjusted so that their means are zero, what is the predicted accuracy

for the coefficients in Equation 4.10.1? From Equations 7.2.3 and 7.2.4 we get:

$$\Phi_1 \equiv \frac{\sigma_{a_1} n^{1/2}}{\sigma_y} = 1 \Rightarrow \sigma_{a_1} = \frac{0.1}{1300^{1/2}} = 0.0028$$

$$\Phi_{j+1} \equiv \frac{\sigma_{a_{j+1}} n^{1/2} \sigma_j}{\sigma_y} = 1 \Rightarrow \sigma_{a_{j+1}} = \frac{0.0028}{\sigma_j} \qquad j = 1 \text{ to } d$$

It should be remembered that these results are based upon the assumption that Equation 4.10.1 is a realistic model for y as a function of x_1, x_2 and x_3. When the experiment is performed, this assumption can be tested using the Goodness-of-Fit tools discussed in Section 3.9.

7.3 Bivariate Normal Distribution

The bivariate normal distribution is useful for modeling many experiments in which the dependent variable is a function of 2 independent variables (which are often spatial dimensions). For example, the dependent variable y might represent the response from a signal at a location x_1, x_2. The bivariate normal distribution is:

$$f(x_1, x_2, \rho) = \frac{1}{2\pi(1-\rho^2)^{1/2}} exp\left(\frac{-x_1^2 + 2\rho x_1 x_2 - x_2^2}{2(1-\rho^2)}\right) \qquad (7.3.1)$$

This equation assumes that x_1 and x_2 have been normalized by subtracting their means μ_1 and μ_2 and then dividing by their standard deviations σ_1 and σ_2. The parameter ρ is the correlation coefficient between x_1 and x_2. The parameter ρ can be treated as know or unknown. An experiment based upon this distribution might be based upon the following model:

$$y = af(x_1, x_2, \rho) \qquad (7.3.2)$$

If ρ is known than there is only one unknown parameter (i.e., a), otherwise both ρ and a are the unknowns that will be determined by the experiment. Other variables in the experiment include the number of data points n, the distributions of x_1 and x_2, and the uncertainty model for the standard deviations of the values of y. The simplest uncertainty model is that all values

of σ_y are equal. Another useful model is that the values of $\sigma_y = C_y\, y^{1/2}$. This second model is used when y represents the number of counts within a region of the x_1 - x_2 plane. For example, in a ballistics experiment, a weapon is fired repeatedly at a target and the numbers of "hits" within each region is counted.

To apply prediction analysis to experiments based upon Equation 7.3.2 it is necessary to be able to generate data sets with any given value of ρ. The method for doing this is to generate two sets of random numbers (x_1 and r). Two popular methods for generating random numbers are numbers that are equally probably within a given range or numbers that are normally distributed with given values of the mean and standard deviation. To create the series x_2 with a correlation coefficient ρ with x_1 the following transformation is made:

$$x_2 = \frac{cx_1 + r\sigma_{x_1} / \sigma_r}{(1 + c^2)^{1/2}} \tag{7.3.3}$$

When both r and x_1 are generated from the same random distributions then $\sigma_{x_1} = \sigma_r$ and Equation 7.3.3 is simplified. The n values of x_2 created in this manner have the value of σ_{x_2} equal to σ_{x_1}. The value of c is related to ρ by the following equation:

$$c = \frac{\rho}{(1 - \rho^2)^{1/2}} \tag{7.3.4}$$

To demonstrate a prediction analysis of an experiment based upon Equation 7.3.2 with two unknowns (ρ and a), the two independent variables are x_1 and r. The variable x_2 (as well as y) is treated as a dependent variable. The parameter c is a constant. The results are included in Figure 7.3.1 using the uncertainty model $\sigma_y = 1$. In Figure 7.3.2 the uncertainty model $\sigma_y = y^{1/2}$ is used. For both cases both x_1 and r are generated from standard normal distributions ($\mu = 0$ and $\sigma = 1$) and x_2 is generated so that $\rho = 0.5$. From Equation 7.3.4 the values of $c = 0.5/(3/4)^{1/2} = 1/3^{1/2} = 0.57735$. Notice that the assumed uncertainties in the values of x_2 are zero because x_2 is an observed and not measured variable. In the proposed experiment it is assumed that x_1 and x_2 are observed without error and only y is measured to assumed uncertainty σ_y.

```
PARAMETERS USED IN REGRESS ANALYSIS: Tue Apr 07 14:06:50 2009
  INPUT PARMS FILE: fig731.par
  INPUT DATA  FILE: fig731.par
  Prediction Analysis Option  (MODE='P')
  STARTREC - First record used            :      1
  N - Number of recs used to build model  : 10000

  Y VALUES - computed using A & RHO values
  SYTYPE1 - Sigma type for X2             :      2
     TYPE 2:  SIGMA X2 = CY1           CY1: 0.000
  SYTYPE2 - Sigma type for Y              :      2
     TYPE 2:  SIGMA Y = CY2            CY2: 1.000
  X1 Values - Normal Distribution: mean= 0.000 sigma= 1.000
  R Values - Normal Distribution: mean= 0.000 sigma= 1.000

Analysis for Set 1
  Function X2: (R + C*X1) / SQRT(1/(1+C^2))
  Function Y:  A * EXP(-(X1^2-2*RHO*X1*X2+X2^2)/(2*(1-RHO^2)))/
               (2*PI*SQRT(1-RHO^2))

    K        CONSTANT        VALUE
    1              C      0.57735

   PARAMETER  INIT_VALUE       VALUE  PRED_SIGMA
         RHO     0.50000     0.50000     0.00194
           A   100.00000   100.00000     0.13372
```

Figure 7.3.1 - Prediction Analysis of a Bivariate model, $\sigma_y = 1$

```
PARAMETERS USED IN REGRESS ANALYSIS: Tue Apr 07 14:47:34 2009
  INPUT PARMS FILE: fig732.par
  INPUT DATA  FILE: fig732.par
    Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  : 10000
  Y VALUES - computed using A & RHO values
  SYTYPE1 - Sigma type for X2             :      2
     TYPE 2:  SIGMA X2 = CY1           CY1: 0.000
  SYTYPE2 - Sigma type for Y              :      4
     TYPE 4: SIGMA Y = CY2 * sqrt(Y)   CY2: 1.000
  X1 Values - Normal Distribution: mean= 0.000 sigma= 1.000
  R Values - Normal Distribution: mean= 0.000 sigma= 1.000

Analysis for Set 1
  Function X2: (R + C*X1) / SQRT(1/(1+C^2))
  Function Y:  A * EXP(-(X1^2-2*RHO*X1*X2+X2^2)/(2*(1-RHO^2)))/
               (2*PI*SQRT(1-RHO^2))

    K        CONSTANT        VALUE
    1              C      0.57735

   PARAMETER  INIT_VALUE       VALUE  PRED_SIGMA
         RHO     0.50000     0.50000     0.00422
           A   100.00000   100.00000     0.37826
```

Figure 7.3.2 – Same as 7.3.1 except $\sigma_y = y^{1/2}$

For the case of constant σ_y the results are presented using the following dimensionless groups:

$$\Psi_a \equiv \frac{\sigma_a n^{1/2}}{\sigma_y} \tag{7.3.5}$$

$$\Psi_\rho \equiv \frac{\sigma_\rho n^{1/2}}{\sigma_y / a} \tag{7.3.6}$$

Results for values of ρ from 0.0 to 0.95 are included in Table 7.3.1. The results show that the predicted value of Ψ_ρ deceases and Ψ_a increases as ρ increases but as ρ approaches one the analysis becomes ill-conditioned and both approach infinity.

ρ	c	Ψ_ρ	Ψ_a
0.0	0.00000	32.66	10.89
0.1	0.10050	32.05	11.10
0.2	0.20412	30.21	11.63
0.3	0.31449	27.32	12.30
0.4	0.43644	23.62	12.93
0.5	0.57735	19.43	13.37
0.6	0.75000	15.07	13.58
0.7	0.98020	10.85	13.60
0.8	1.33333	6.95	13.68
0.9	2.06474	3.61	14.89
0.95	9.74359	5.41	52.35

Table 7.3.1 – Prediction Analysis for constant σ_y

Figure 7.3.2 shows a prediction analysis for a case in which the values of σ_y are equal to $\sigma_y = C_y y^{1/2}$. The proportionality constant C_y is in units of $y^{1/2}$ and typically is equal to one unless a and y are expressed in different units (for example, a is expressed in counts/cm^2 and y will be measured in counts within a grid location in which the grid size is not one cm^2). The dimensionless groups used to present the results are:

$$\Phi_a \equiv \frac{\sigma_a n^{1/2}}{C_y a^{1/2}} \tag{7.3.7}$$

$$\Phi_\rho \equiv \frac{\sigma_\rho n^{1/2}}{C_y / a^{1/2}} \tag{7.3.8}$$

ρ	c	Φ_ρ	Φ_a
0.0	0.00000	7.11	3.54
0.1	0.10050	6.97	3.56
0.2	0.20412	6.56	3.61
0.3	0.31449	5.93	3.68
0.4	0.43644	5.13	3.73
0.5	0.57735	4.22	3.78
0.6	0.75000	3.27	3.83
0.7	0.98020	2.34	3.94
0.8	1.33333	1.51	4.32
0.9	2.06474	0.87	5.76
0.95	9.74359	1.54	25.47

Table 7.3.2 – Prediction Analysis for $\sigma_y = C_y y^{1/2}$

Results for values of ρ from 0.0 to 0.95 are included in Table 7.3.2 and are similar to the results shown in Table 7.3.1. The results show that the predicted value of Φ_ρ deceases and Φ_a increases as ρ increases but as ρ approaches one the analysis becomes ill-conditioned and both approach infinity.

Example 7.3.1:

We want to design an experiment to determine the amplitude $\boldsymbol{a}$ of a bivariate distribution to one percent accuracy. An initial analysis suggests that the values of $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ are correlated with a value of the correlation coefficient $\boldsymbol{\rho}$ equal to approximately 0.5. We plan to measure values of $\boldsymbol{y}$ using a grid in the $\boldsymbol{x_1}$ - $\boldsymbol{x_2}$ plane of size one $\boldsymbol{mm^2}$. The units of $\boldsymbol{a}$ are to be expressed in counts per $\boldsymbol{cm^2}$ = 100 * counts per $\boldsymbol{mm^2}$ so $\boldsymbol{C_y}$ = 10. For example if the number of counts within a particular 1 $\boldsymbol{mm^2}$ square in the grid is 100 counts, then this number is accurate to $\pm 100^{1/2} = \pm 10$ counts. Converting to counts per $\boldsymbol{cm^2}$ the count rate is 10000 but the accuracy remains 10% (i.e., 10/100 = 1000/10000 = $\boldsymbol{C_y}$*$10000^{1/2}$/10000 = $\boldsymbol{C_y}$/100 and therefore $\boldsymbol{C_y}$ = 10).

From Equation 7.3.7 and Table 7.3.2 we establish the following relationship:

$$\Phi_a \equiv \frac{\sigma_a n^{1/2}}{C_y a^{1/2}} = 4.22 = \frac{\sigma_a}{a}\frac{a n^{1/2}}{C_y a^{1/2}} = 0.01\frac{a^{1/2} n^{1/2}}{C_y}$$

$$0.01\frac{a^{1/2} n^{1/2}}{C_y} = 4.22 \Rightarrow (an)^{1/2} = \frac{10 * 4.22}{0.01} = 4220$$

This relationship can now be used to design the experiment. For example if we set $\boldsymbol{n}$ = 10000 points, the value of $\boldsymbol{a^{1/2}}$ should be about 42.2 and therefore we require a count rate of $\boldsymbol{a}$ equal to approximately 1780 counts per $\boldsymbol{cm^2}$ to satisfy the accuracy requirements of the experiment.

7.4 Orthogonality

The subject of experimental design has long recognized the value of orthogonality in the independent variables for multidimensional models. An excellent book that summarizes the fundamentals of experimental design is the 2005 book by G.E.P. Box, J.S. Hunter and W.G. Hunter: ***Statistics for Experimenters*** [BO05]. On the subject of orthogonality for multidimensional models they state: "In general, when the $\boldsymbol{x}$'s are orthogonal, calculations and understanding are greatly simplified. Thus orthogonality is a property that is still important even in this age of computers."

In our discussions of multidimensional models we have often assumed an orthogonal distribution of x's. For example in Section 4.10 in the discussion of Multiple Linear Regression, the values of the x's were assumed to be distributed according to Equation 4.10.2:

$$x_{ji} = x_{j1} + (i-1)\Delta x_j \qquad i = 1 \text{ to } n_j \tag{4.10.2}$$

This assumption insures an orthogonal distribution of the independent variables because it satisfies the following criterion for orthogonality:

$$\sum_{i=1}^{i=n} (x_{ji} - x_{javg})(x_{ki} - x_{kavg}) = 0 \quad \text{for } j \neq k \tag{7.4.1}$$

In Section 5.5 the discussion of Bivariate Separation is also based upon the assumption that the values of x_1 and x_2 are orthogonal. This orthogonality is seen in Figures 5.5.1 and 5.5.2.

In this section we consider the problem of a lack of orthogonality in the independent variables. Is there a price to pay for a lack of orthogonality? To study this problem, we can use random distributions to generate bivariate distributions with a specified amount of correlation between x_1 and x_2. The technique for accomplishing this is described in Section 7.3. Prediction analyses can then be performed to observe the effect of a lack of orthogonality in the independent variables.

One of the basic problems considered in the design of experiments is based upon the simple mathematical model:

$$y = a_1 x_1 + a_2 x_2 \tag{7.4.2}$$

Box and Hunter use an example from chemical engineering to discuss experiments based upon this equation: the rate of formation on an undesirable impurity y as a function of the concentration of a monomer x_1 and the concentration of a dimmer x_2 [BO05]. The rate of formation is zero when both components x_1 and x_2 are zero.

To illustrate the effect of a lack of orthogonality in the x variables, the following simple example based upon Equation 7.4.2 shows the deterioration in the predicted accuracy of a_1 and a_2 as the correlation coefficient ρ in-

creases. Consider an experiment in which there are only 4 data points distributed in two dimensions as shown in Figure 7.4.1.

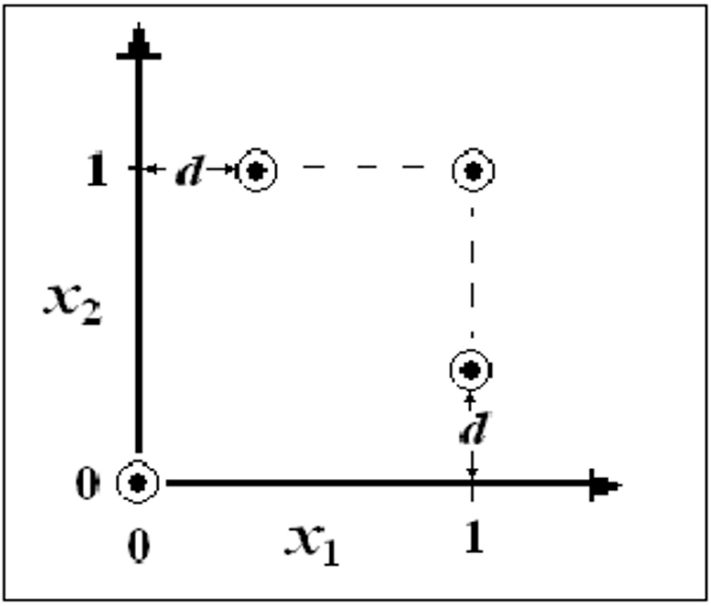

Figure 7.4.1 – Four data points for 2D experiment

For the case where $d = 0$ the correlation coefficient ρ is zero (i.e., the x values are orthogonal) but as d approaches one ρ approaches one. In Table 7.4.1 we see values of σ_{a_1} and σ_{a_2} for values of d from zero (orthogonal) to 0.95 (highly correlated). These results were obtained using prediction analyses based upon Equation 7.4.2 with coefficient a_1 and a_2 both equal to 1. (However, the same results are obtained regardless of the values of a_1 and a_2.) We see that as d (and therefore ρ) increases the predicted values of σ_{a_1} and σ_{a_2} increase. If the purpose of the experiment is to measure a_1 and a_2, then we see that if the x's are highly correlated, the predicted values of σ_{a_1} and σ_{a_2} are much greater than if we can choose values for the x's that are uncorrelated or close to being uncorrelated (i.e., orthogonal).

d	ρ	σ_{a_1}	σ_{a_2}
0.0	0.000	0.816	0.816
0.1	0.107	0.879	0.879
0.2	0.229	0.963	0.963
0.3	0.362	1.075	1.075
0.4	0.500	1.231	1.231
0.5	0.636	1.455	1.455
0.6	0.761	1.798	1.798
0.7	0.865	2.379	2.379
0.8	0.941	3.549	3.549
0.9	0.985	7.077	7.077
0.95	0.997	14.145	14.145

Table 7.4.1 – Results for a simple experiment with 4 data points.

When the independent variables x_1 and x_2 are uncorrelated, we can proceed to a simple analytical solution using a procedure similar to the procedure followed in Section 4.3 for a straight line. Assuming that all values of y are measured to the same accuracy (i.e., σ_y = *constant*), the values of w_i are all equal to $1 / \sigma_y^2$ and we can then express the terms of the C matrix as:

$$C = \frac{1}{\sigma_y^2}\begin{bmatrix} \sum_{i=1}^{i=n} x_{1i}^2 & \sum_{i=1}^{i=n} x_{1i}x_{2i} \\ \sum_{i=1}^{i=n} x_{1i}x_{2i} & \sum_{i=1}^{i=n} x_{2i}^2 \end{bmatrix} \tag{7.4.3}$$

If we assume that the values of x_1 and x_2 are uncorrelated and randomly distributed within the volume $-R \le x_1 \le R$, $-R \le x_2 \le R$ we can show that the number of data points n becomes large, the C matrix approaches:

$$C = \frac{R^2 n}{\sigma_y^2}\begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \tag{7.4.4}$$

Inverting the matrix we obtain:

$$C^{-1} = \frac{3\sigma_y^2}{R^2 n}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{7.4.5}$$

Thus the predicted values of σ_{a_1} and σ_{a_2} are:

$$\sigma_{a_k} = (C_{kk}^{-1})^{1/2} = \frac{3^{1/2}\sigma_y}{\mathbf{R}n^{1/2}} = 1.732\frac{\sigma_y}{\mathbf{R}n^{1/2}} \tag{7.4.6}$$

If the values of x_1 and x_2 are correlated ($\rho \neq 0$) we can run prediction analyses to observe the effect upon the predicted values of σ_{a_1} and σ_{a_2}. A prediction analysis for $\rho = 0.5$ is shown in Figure 7.4.1. To generate data sets using Equation 7.4.2 with a specified value of ρ using Equations 7.3.3 and 7.3.4, random values of x_1 and r were generated in the range -1 to 1 (i.e., $\mathbf{R} = 1$). The computed values of x_2 are not necessarily within the same range. For example, for the 10,000 points generated for the calculation shown in Figure 7.4.2, about 4% of the values of x_2 were less than -1 and about 4% were greater than 1. Results for $\rho = 0$ to 0.95 are shown in Table 7.4.2. The results are presented in terms of the dimensionless groups Ψ_1 and Ψ_2:

$$\Psi_k \equiv \frac{\sigma_{a_k}\mathbf{R}n^{1/2}}{\sigma_y} \quad k = 1 \;\&\; 2 \tag{7.4.7}$$

```
PARAMETERS USED IN REGRESS ANALYSIS: Wed Jun 24, 2009

  INPUT PARMS FILE: fig742.par
  INPUT DATA  FILE: fig742.dat
  REGRESS  VERSION: 4.27, May 5, 2009

  Prediction Analysis Option  (MODE='P')
  N - Number of recs used to build model  : 10000
  Y VALUES - computed using A0 & X values
  SYTYPE1 - Sigma type for Y              :     1
     TYPE 1:  SIGMA Y = 1
  M - Number of independent variables     :     2
  Column for X1                           :     1
  SXTYPE1 - Sigma type for X1             :     0
  Column for X2                           :     2
  SXTYPE2 - Sigma type for X2             :     0

Analysis for Set 1
  Function Y:  A1*X1 + A2*X2

   PARAMETER  INIT_VALUE       VALUE  PRED_SIGMA
          A1     1.00000     1.00000     0.02004
          A2     1.00000     1.00000     0.02000
```

Figure 7.4.2 – Prediction Analysis for constant σ_y, $\rho = 0.5$

Notice that the observed values of Ψ_1 and Ψ_2 are not exactly the same in Table 7.4.2 for the various values of ρ. These small differences are due to the random nature of the values of x_1 and x_2. Increasing n decreases these differences. What is clear from the table is that increasing the correlation between the x's does increase the predicted values of σ_{a_1} and σ_{a_2}. The effect, however, is not dramatic but as expected, as ρ approaches one, the predicted values approach infinity.

ρ	c	Ψ_1	Ψ_2
0.0	0.00000	1.728	1.732
0.1	0.10050	1.738	1.741
0.2	0.20412	1.767	1.768
0.3	0.31449	1.817	1.816
0.4	0.43644	1.893	1.890
0.5	0.57735	2.004	2.000
0.6	0.75000	2.171	2.165
0.7	0.98020	2.433	2.426
0.8	1.33333	2.895	2.887
0.9	2.06474	3.981	3.974
0.95	9.74359	5.551	5.548

Table 7.4.2 – Prediction Analysis for constant σ_y, varying ρ

A slightly more sophisticated model than Equation 7.4.2 includes an interaction term between the two independent variables:

$$y = a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 \tag{7.4.8}$$

This model requires an additional dimensionless group:

$$\Psi_{12} \equiv \frac{\sigma_{a_{12}} R^2 n^{1/2}}{\sigma_y} \tag{7.4.9}$$

Results for this model are summarized in Table 7.4.3. Note that the results for Ψ_1 and Ψ_2 are exactly the same as for the simpler model and the value of Ψ_{12} decreases with increasing ρ. Clearly, increasing correlation implies that the interaction term becomes increasingly important.

ρ	Ψ_1	Ψ_2	Ψ_{12}
0.0	1.728	1.732	3.000
0.1	1.738	1.741	2.964
0.2	1.767	1.768	2.926
0.3	1.817	1.816	2.869
0.4	1.893	1.890	2.797
0.5	2.004	2.000	2.713
0.6	2.171	2.165	2.621
0.7	2.433	2.426	2.525
0.8	2.895	2.887	2.427
0.9	3.981	3.974	2.330
0.95	5.551	5.548	2.283

Table 7.4.3 – Prediction Analysis for Equation 7.4.8, varying ρ

Example 7.4.1:

We want to design an experiment based upon Equation 7.4.8 and we wish to determine all three coefficients to an accuracy of 0.02. The values of the y's will be measured to an accuracy of 0.1 (i.e., $\sigma_y = 0.1$). A preliminary analysis indicates that the correlation coefficient between the x's is about 0.5 (i.e., $\rho = 0.5$). How many points in the range $-1 \le x_1 \le 1$ are required to satisfy the experimental requirement?

From Table 7.4.3 we see that the most difficult requirement is the measurement of a_{12}:

$$\Psi_{12} \equiv \frac{\sigma_{a_{12}} R^2 n^{1/2}}{\sigma_y} = 2.713 = \frac{0.02 * 1^2 * n^{1/2}}{0.1}$$

$$n = (5 * 2.713)^2 = 184$$

The results included in Tables 7.4.2 and 7.4.3 are based upon a choice of the range of values of x_1 from -1 to 1. Changes in the range have a very important effect upon these results. If the range is changed from $-1 \le x_1 \le 1$ to $-R \le x_1 \le R$ the values of σ_{a_1} and σ_{a_2} decrease inversely with increas-

ing $\mathbf{R}$ and $\sigma_{a_{12}}$ decreases inversely with increasing $\mathbf{R^2}$. For example, doubling the range from $-1 \le x_1 \le 1$ to $-2 \le x_1 \le 2$ will result in a decrease in the values of σ_{a_1} and σ_{a_2} by a factor of 2 and $\sigma_{a_{12}}$ by a factor of 4.

Example 7.4.2:

We wish to design the same experiment as in Example 7.4.1 but can perform the experiment by choosing values of x_1 in the range $-2 \le x_1 \le 2$. For this modified experiment, the most difficult requirement is the measurements of a_1 and a_2:

$$\Psi_k \equiv \frac{\sigma_{a_k} \mathbf{R} n^{1/2}}{\sigma_y} = 2.000 = \frac{0.02 * 2n^{1/2}}{0.1} \Rightarrow n = (2.5 * 2.000)^2 = 25$$

What value of the range parameter $\mathbf{R}$ is required so that $n = 100$ satisfies the experimental requirement?

$$\Psi_k \equiv \frac{\sigma_{a_k} \mathbf{R} n^{1/2}}{\sigma_y} = 2.000 = \frac{0.02\mathbf{R}100^{1/2}}{0.1} \Rightarrow \mathbf{R} = \frac{2.000 * 0.1}{0.02 * 10} = 1$$

$$\Psi_{12} \equiv \frac{\sigma_{a_{12}} \mathbf{R}^2 n^{1/2}}{\sigma_y} = 2.713 \Rightarrow \mathbf{R} = \left(\frac{2.713 * 0.1}{0.02 * 10}\right)^{1/2} = 1.16$$

The requirement on the a_{12} term is the more difficult to satisfy so to measure all 3 terms to an accuracy of at least 0.02 using only 100 data points, we must extend the range to $-1.16 \le x_1 \le 1.16$.

References

[**AB64**] M. Abramowitz, I. Stegen, ***Handbook of Mathematical Functions***, NBS, 1964.

[**AZ94**] M. Azoff, ***Neural Network Time Series Forecasting of Financial Markets***, Wiley, 1994.

[**BA74**] Y. Bard, ***Nonlinear Parameter Estimation***, Academic Press, 1974.

[**BA94**] R. J. Bauer, ***Genetic Algorithms and Investment Strategies***, Wiley, 1994.

.[**BE06**] Y. Ben-Haim, ***Info-Gap Theory: Decisions Under Severe Uncertainty****, 2nd Edition*, Academic Press, 2006.

[**BR01**] F. Brauer, C. Castillo-Chavez, ***Mathematical Models in Population Biology and Epidemiology,*** Springer Verlag, 2001.

[**BO05**] G.E.P. Box, J.S.Hunter, W.G. Hunter, ***Statistics for Experimenters,*** Wiley, 2005.

[**BU81**] W. K. Buler, ***Gauss – A Biographical Study***, Springer Verlag, 1981.

[**DA95**] M. Davidian, D. Giltinan, ***Nonlinear Models for Repeated Measurement Data,*** Chapman and Hall, 1995.

[**DE43**] W.E. Deming, ***Statistical Adjustment of Data,*** Wiley, 1943.

J. Wolberg, *Designing Quantitative Experiments*, DOI 10.1007/978-3-642-11589-9,

[**DR66**] N.R. Draper, H. Smith, ***Applied Regression Analysis,*** Wiley, 1966.

[**ED88**] L. Edelstein-Keshet, ***Mathematical Models in Biology***, Random-House, 1988.

[**FO57**] G.E. Forsythe, ***Generation and Use of Orthogonal Polynomials for Data-Fitting,*** Journal of Soc. Of Applied Math, Vol 5, No 2, June 1957.

[**FR80**] H. Freedman, ***Deterministic Models in Population Biology,*** Marcel Dekker, 1980.

[**FR92**] J. Freund, ***Mathematical Statistics*** *(5th Edition)*, Prentice Hall, 1992.

[**GA57**] C.F. Gauss, C.H. Davis, ***Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections: A Translation of Gauss's "Theoria Motus",*** Little Brown, 1857.

[**GA84**] T. Gasser, H. G. Muller, W. Kohler, L. Molianari, and A. Prader, ***Nonparametric Regression Analysis of Growth Curves***, Annals of Statistics, 12, 210-229, 1984.

[**GA92**] P. Gans, ***Data Fitting in the Chemical Sciences***, Wiley, 1992.

[**GA95**] E. Gately, ***Neural Networks for Financial Forecasting***, Wiley, 1995.

[**GR09**] M. Greenberg, ***Applied Differential Equations with Boundary Value Problems***, Prentice Hall, 2009.

[**HA00**] W. Hans, ***Astrometry of Fundamental Catalogues: the Evolution from Optical to Radio Reference Frames***, Springer, 2000.

[**HA01**] T. Hastie, R. Tibshirani, J. Friedman, ***The Elements of Statistical Learning***, Springer Series in Statistics, 2001.

[**HA90**] W. Hardle, ***Applied Nonparametric Regression***, Cambridge University Press, 1990.

[**HO04**] T. Ho, S. B. Lee, ***The Oxford Guide to Financial Modeling***, Oxford University Press, 2004.

[**HO06**] F. Hoppensteadt, ***Predator – Prey Model***, Scholarpedia 1 (10): 1563, 2006.

[**JO70**] N.I. Johnson, S. Kotz, ***Continuous Univariate Distributions***, Houghton Mifflin, 1970.

[**KO95**] J. Kovalevsky, ***Modern Astronomy***, Springer, 1995.

[**KO04**] J. Kovalevsky, P. Seidelman, P. Kenneth, ***Fundementals of Astronomy***, Cambridge University Press, 2004.

[**MA04**] B. Mandelbrot, R.L. Hudson, ***The (Mis) Behavior of Markets***, Basic Books, 2004.

[**MC94**] J. McClave, P. Benson, ***Statistics for Business and Economics***, Macmillian, 1994.

[**MC05**] P. McNelis, ***Neural Networks in Finance***, Academic Press Advanced Finance, 2005.

[**ME92**] W. Mendenhall, T. Sincich, ***Statistics for Engineering and the Sciences***, Maxwell MacMillan, 1992.

[**ME84**] M. Merriman, ***A Text-book on the Method of Least Squares***, **Wiley**, 1884.

[**RE95**] A.P. Refenes, ***Neural Networks in the Capital Markets***, Wiley, 1995.

[**ST86**] S. M. Stigler, ***The History of Statistics: the Measurement of Uncertainty before 1900***, Harvard University Press, 1986.

[**ST03**] ***Statistica Electronic Textbook***, www.statsoft.com/textbook/stathome.html, Distribution Tables, StatSoft Inc, 2003.

[**TV04**] J. Tvrdik, I. Krivy, ***Comparison of Algorithms for Nonlinear Regression Estimates,*** Compstat'2004 Symposium, Physcia-Verlag/Springer 2004.

[**VE02**] W. Venables, B. Ripley, ***Modern Applied Statistics with S***, 4th Edition, Springer-Verlag, 2002.

[**VE**38] P.F. Verhulst, ***Notice sur la loi que la population suit dans son accroissment***, Corr. Math. Et Phys., 10:113-121, 1838.

[**WA93**] R. E. Walpole, R. H. Meyers, ***Probability and Statistics for Engineers and Scientists****, (5th edition),* McMillan, 1993.

[WO62] J. R. Wolberg, T.J. Thompson, I. Kaplan, ***A Study of the Fast Fission Effect in Lattices of Uranium Rods in Heavy Water,*** M.I.T. Nuclear Engineering Report 9661, February, 1962.

[WO67] J. R. Wolberg, ***Prediction Analysis,*** Van Nostrand, 1967.

[**WO00**] J. R. Wolberg, ***Expert Trading Systems: Modeling Financial Markets with Kernel Regression,*** Wiley , 2000.

[WO06] J. R. Wolberg, ***Data Analysis using the Method of Least Squares,*** Springer, 2006.

Index