A NEW ESTIMATION PROCEDURE FOR LINEAR COMBINATIONS OF EXPONENTIALS

by

Richard Garth Cornell, A.B., M.S.

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in candidacy for the degree of

DOCTOR OF PHILOSOPHY

in

STATISTICS

APPROVED:

Co-chairman, Advisory Committee

Co-chairman, Advisory Committee

May 1956

Blacksburg, Virginia

TABLE OF CONTENTS

I. INTRODUCTION

Many experimental problems in the natural sciences result in data which can best be represented by linear combinations of exponentials of the form

$$f(t) = \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t} \quad . \qquad (1.1.1)^*$$

Among such problems are those dealing with growth, decay, ion concentration, and survival and mortality. Also, in general, the solution to any problem which may be represented by linear differential equations with constant coefficients is a linear combination of exponentials. In most problems like those which have been mentioned, the parameters α_k and λ_k in equation (1.1.1) have biological or physical significance. Therefore, in fitting a function of the form (1.1.1) to the data it is not only necessary that the function approximate the data closely, but it is also necessary that the parameters of (1.1.1) be accurately estimated. Furthermore, a measure of the accuracy of the estimation of the parameters is required.

The present methods of estimating the parameters of a linear combination of exponentials are often inadequate. Some of these methods will be discussed in Chapter II. However, the primary purpose of this paper is to introduce a new estimation procedure which will overcome some

---

* (a, b, c) denotes the c^{th} equation in the b^{th} section of the a^{th} chapter.

of the present difficulties, at least for special cases. This new estimation procedure will be developed in Chapter III. Included will be a discussion of the basic model for which the method is derived. Chapter IV will be concerned with the limiting distribution of the estimators obtained from the new procedure. Then in Chapter V, the statistical properties of the estimators will be considered.

The small sample distribution of the estimators from the new procedure will be studied in Chapter VI. Results from some empirical sampling work will be reported in this chapter. Then in Chapter VII possible extensions of the method will be considered and ways will be described in which the new procedure may be applied to a greater number of experimental situations. Chapter VII will also contain several illustrations of the application of the new method as well as a limited empirical comparison of the new procedure with presently existing methods. Finally, Chapter VIII will be devoted to a critical evaluation of the new procedure relative to other estimation procedures for linear combinations of exponentials.

II. REVIEW OF LITERATURE

2.1 Iterative Maximum Likelihood Methods

Before turning to the development of the new procedure, let us look briefly at some of the estimation procedures now in existence. The first method that we will consider is an iterative procedure for calculating maximum likelihood estimates which has been presented by Fisher [7]* and illustrated by Koshal [20, 21]. A detailed discussion of this method and a few examples of its use are given by Garwood [9]. Although the presentation is applicable to fitting the parameters of any distribution, the particular application of the method to a linear combination of exponentials follows directly from the general development.

In general, let $y_1, y_2, \ldots, y_N$ be a sample drawn at random from a population of known form so that the sample has the joint density function $P(y, \theta)$, where θ represents a row vector of parameters $(\theta_1, \theta_2, \ldots, \theta_s)$ and y is a row vector of the observations. For example, suppose the variates y_i are independently and normally distributed with means $f(t_i)$, where the function f is given by (1.1.1), and with common variance σ^2. That is, let

$$y_i = f(t_i) + e_i \,, \quad i = 1, 2, \ldots, N \,, \tag{2.1.1}$$

---

* Numbers in square brackets refer to the bibliography.

where the errors e_i are independent, normally distributed variates. In this instance,

$$P(y, \theta) = (2 \pi \sigma^2)^{-\frac{N}{2}} \exp. \left[- \frac{1}{2\sigma^2} \sum_{i=1}^{N} (y_i - \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i})^2 \right] .$$

Now if we let $L(y, \theta)$ be the natural logarithm of P and $\hat{\theta}$ be the row vector of maximum likelihood estimators $\hat{\theta}_k$ of the parameters θ_k,

$$\frac{\partial}{\partial \theta_k} L(y, \theta) \Big|_{\theta = \hat{\theta}} = 0 \ , \ k=1, 2, \ldots, s \ . \qquad (2.1.2)$$

For our example*,

$$L(y, \theta) = - \frac{N}{2} \ln(2 \pi \sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{N} (y_i - \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i})^2 \ .$$

So if we let θ_k represent α_1, condition (2.1.2) implies that in this case

$$\frac{1}{\sigma^2} \sum_{i=1}^{N} e^{-\hat{\lambda}_1 t_i} (y_i - \sum_{k=1}^{p} \hat{\alpha}_k e^{-\hat{\lambda}_k t_i}) = 0 \ , \qquad (2.1.3)$$

where the $\hat{\alpha}_k$ and the $\hat{\lambda}_k$ are maximum likelihood estimators.

Now let us postulate, in the general development, that an approximation

---

* $\ln$ denotes a natural logarithm.

$$a_1 = (a_{11}, a_{21}, \ldots, a_{s1})$$

to $\hat{\theta}$ is available. Then each partial derivative given by (2.1.2) may be expanded in a Taylor series in terms of the elements of $\hat{\theta}$ about the corresponding elements of a_1. If terms which involve partial derivatives of L of order greater than two are ignored, the resulting linear equations may be solved for the vector

$$\delta_1 = (a_1 - \hat{\theta}) \quad .$$

Then a new vector of estimates

$$a_2 = (a_1 + \delta_1)$$

may be formed and the process repeated. If after, say, h iterations, δ_n is sufficiently close to a null vector, the resultant elements of a_{h+1} are taken to be the maximum likelihood estimates of the elements of θ .

The expansion of equation (2.1.2) in a truncated Taylor series gives rise to coefficients

$$L_{r,s;h} = \frac{\partial^2}{\partial \theta_r \, \partial \theta_s} L(y, \theta) \Bigg|_{\theta = a_h} \quad .$$

The $L_{r,s;h}$ are functions of the observations in the vector y as well as the approximations in the vector a_h. Garwood observes that the calculations are simplified if in the expression for the $L_{r,s;h}$ in each iteration the y_i observations are replaced by the values which they would be expected to have if in fact θ were equal to its current approximation a_h. For instance, in the example introduced earlier let us set $\alpha_h = \theta_{2h-1}$ and $\lambda_n = \theta_{2h}$, $h = 1, 2, \ldots, p$. An expansion of the partial derivative given by (2.1.3) in a truncated Taylor series about some vector a_h, say a_1, would include a coefficient of the form

$$L_{1,2;1} = \frac{\partial^2}{\partial\alpha_1\, \partial\lambda_1} L(y, \theta) \Bigg|_{\theta=a_1}$$

$$= -\frac{1}{\sigma^2} \sum_{i=1}^{N} t_i\, e^{-\lambda_1 t_i} (y_i - \alpha_1 e^{-\lambda_1 t_i} - \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i}) \Bigg|_{\theta=a_1}$$

$$= -\frac{1}{\sigma^2} \sum_{i=1}^{N} t_i\, e^{-a_{21} t_i} (y_i - a_{11} e^{-a_{21} t_i} - \sum_{h=1}^{p} a_{2h-1,1} e^{-a_{2h,1} t_i}) .$$

(2.1.4)

Now if θ were in fact equal to a_1, then α_h would equal $a_{2h-1,1}$ and λ_h would equal $a_{2h,1}$, $h = 1, 2, \ldots, p$. Furthermore, in each case y_i would have as its expectation

$$\mathcal{E}(y_i)\Big|_{\theta=a_1} = f(t_i)\Big|_{\theta=a_1} = \sum_{h=1}^{p} a_{2h-1,1}\, e^{-a_{2h,1}t_i} \; .$$

Thus, if the observations y_i are replaced by their expectations for θ equal to a_1, equation (2.1.4) reduces to

$$L_{1,2;1} = + \frac{a_{11}}{\sigma^2} \sum_{i=1}^{N} t_i\, e^{-2a_{21}t_i} \; . \qquad (2.1.5)$$

Two quite similar iterative methods have just been indicated for calculating maximum likelihood estimates for the parameters in a linear combination of exponentials. The first one involves carrying out the iterations described above without making any simplifying substitutions for the y_i and will be called Method 1. The second, designated Method 2, is the modification introduced when the approximations suggested for the y_i, such as those which led to (2.1.5), are utilized. In the nonlinear estimation examples which he tried, Garwood observed that Method 1 converged in fewer iterations than did Method 2, but that Method 2 entailed less work per iteration.

A more detailed exposition of these iterative methods may be found in Garwood's paper. However, a few more remarks are in order here. Firstly, if the observations y_i are postulated to have independent, normally distributed errors as in our example, these iterative procedures lead to the least squares estimates of the parameters as well as/to the maximum likelihood estimates. In this instance, Method 2 is the same as that developed for least squares estimation by Deming [3], even though the approach is different. Secondly, experience here in the Oak Ridge National Laboratory indicates that these methods, when applied to linear combinations of exponentials, are not in general amenable to calculation on desk computers because the convergence is too slow. This is especially true of Method 2. In fact, because the methods presented by Garwood have not been tried very extensively even though fast automatic computers are now available, little is known about their convergence properties.

2.2 The Prony Method

Another method for estimating the parameters of a linear combination of exponentials is presented by Prony [23] as a method of interpolation. Whittaker and Robinson [30] describe a modification of Prony's method while Householder [13] discusses the Whittaker and Robinson version of the method as an estimation procedure. Householder also suggests an extension of this estimation procedure.

The model underlying Prony's method is similar to that given by equation (2.1.1) for the observations y_i in our earlier example. However, the method requires that the y_i be taken at equally spaced intervals of time, so to the model given before we add the restriction that $t_i = Ki$ for constant K. Also, the errors are now only required to have zero means.

The first step in the Prony procedure is to set each y_i equal to its expectation. That is, set

$$y_i = \sum_{k=1}^{p} \alpha_k e^{-\lambda_k K i} , \quad i = 1, 2, \ldots, N . \qquad (2.2.1)$$

Now, when (2.2.1) holds, it is shown by Prony on the basis of some results from the calculus of finite differences that each y_i satisfies a p^{th} order difference equation of the form

$$y_{s+p} - y_{s+p-1} E_1 + \ldots + (-1)^{p-1} y_{s+1} E_{p-1} + (-1)^p y_s E_p = 0 , \qquad (2.2.2)$$

$$s = 0, 1, 2, \ldots, N - p ,$$

where the E_r functions are the elementary symmetric functions of the exponentials $\Lambda_k = \exp. (-\lambda_k K)$. That is, E_r equals the sum of all possible distinct products of the Λ_k taken r at a time. The next step in the Prony method is to solve the equations (2.2.2), which are linear in the functions E_r , for estimates $\hat{E}_r$ of these functions. It is at this point that Prony's original presentation differs from the estimation version given by Whittaker and Robinson. However, once estimates $\hat{E}_r$ are obtained, both versions proceed in the same manner. Since the E_r are the elementary symmetric functions of the exponentials $\exp.(-\lambda_k K)$, the estimates $\exp.(-\hat{\lambda}_k K)$ are determined by finding the p roots of the polynomial equation,

$$x^p - \hat{E}_1 x^{p-1} + \hat{E}_2 x^{p-2} \ldots + (-1)^p \hat{E}_p = 0 .$$

Finally, after the λ_k are estimated, estimates $\hat{\alpha}_k$ of the coefficients in (2.1.1) are determined by least squares.

In his presentation of this method, Prony requires that as many parameters be included in the model as there are observations available. In this case (2.2.2) leads to as many linear equations in the function E_r as there are such functions. Hence, the equations (2.2.2) can be solved exactly for the estimates $\hat{E}_r$. From an interpolation point of view Prony's approach does not impose a severe limitation on the method, but it is not acceptable from a statistical viewpoint. Then it is usually desirable to make many more observations than there are parameters to be estimated. Also, for large numbers of observations Prony's original method becomes too cumbersome computationally. However, the adaptation of Prony's method given in [13] and [30] is suitable for statistical estimation purposes. In this version of the method the number of observations is allowed to be larger than the number of parameters to be estimated. Thus, equation (2.2.2) yields more equations for the functions E_r than there are such functions. But these equations are regarded as equations of condition, and from them estimates of the E_r are obtained by least squares calculations.

The estimates of the E_r in the estimation version of Prony's method are not the same as the estimates which would result if the least squares technique were applied directly to the model for a linear combination of exponentials. As pointed out by Householder, both the coefficients y_i and the E_r of the set of equations (2.2.2) are subject to error, while in the usual least squares situation only the E_r would be subject to error. Therefore, we cannot attribute any of the usual least squares properties to estimators obtained by the Prony method, and, in fact, little is known regarding the properties of these estimators. In particular, we have no measure of the variances of such estimators.

Householder modifies the estimation version of the Prony method so that valid least squares estimates may be obtained. He essentially applies the method as it is currently used to obtain initial estimates of the exponents, and then he goes through an iterative procedure to arrive at least squares estimates. The iterative method is that given by Deming, which, as we have already observed, is the same as Method 2 in Garwood's paper if the observations are subject to normally distributed errors. Householder also incorporates a test to determine how many exponentials are needed to adequately represent the data in his modification. Unfortunately, Householder's adaption of Prony's method not only fails to converge sometimes, but it also has been known to converge to unreasonable estimates. Whether this difficulty is inherent in the iterative least squares method or is due to a failure of the Prony estimation procedure to produce satisfactory initial estimates is not known.

2.3 A Graphical Procedure

Perhaps the most common way of fitting linear combinations of exponentials is a graphical "peeling off" procedure applied to a plot of the logarithms of the data against time. For the simplest case of one exponential term, this procedure reduces to fitting a straight line, usually by least squares, to such a semi-logarithmic plot of the data. This method of fitting a single exponential with its coefficient requires the assumption that the errors in the logarithms of the data are homogeneous. For more than one exponential term, the data must be such that a plot of the logarithms of the last few observations is essentially linear. This situation often obtains in a linear combination of exponentials when one of the exponents λ_k in (1.1.1) is appreciably smaller than any of the others, for beyond a certain point that exponential would be expected to be the only one contributing markedly to the total.

The first step in the graphical procedure is to determine the linear relationship which exists on the tail of the semi-logarithmic curve from any two points on the tail. From this linear relationship one of the terms in the linear combination of exponentials is found by taking into account the linearity of the data on a semi-logarithmic plot. Next a semi-logarithmic plot is made of the difference between the experimental points and the corresponding calculated points determined from the term which has already been found. Then another linear fit is made on the tail of this plot in order to obtain another term of the linear combination of exponentials. The procedure is repeated until all the experimental points are included in the partitioning process.

Feurzeig and Tyler [4] note that although this method of fitting linear combinations of exponentials is well known, it has received very little attention in the literature. Therefore, they give a detailed description of the method with several illustrations. Previous to Feurzeig and Tyler's paper, Smith and Morales [26] presented an application of the method and Running [24] discussed the special case of a linear combination of two exponentials. This graphical method is computationally easy and it frequently gives a good fit to the data.

However, no indication of the accuracy of the estimation of the parameters is available. In fact, the number of exponential terms included and the values of the estimates obtained depend greatly on the judgment of the statistician in partitioning the data to obtain linear relationships on the semi-logarithmic plots. Although the method may be carried out easily on desk computers, it is not easily adaptable to calculation on automatic computers because of the judgment decisions required.

III. THE NEW ESTIMATION PROCEDURE

3.1 The Model

A new noniterative method for fitting linear combinations of exponentials will be developed in this chapter. The new method leads to relatively simple estimators of the parameters of the model presented in this section. Under this model observations y_{ij} are specified such that

$$y_{ij} = \alpha_1 e^{-\lambda_1 t_i} + \alpha_2 e^{-\lambda_2 t_i} + \ldots + \alpha_p e^{-\lambda_p t_i} + e_{ij} ; \qquad (3.1.1)$$

$$i = 0, 1, 2, \ldots, 2n\,p-1 ; \quad j = 1, 2, \ldots, m .$$

Also, we require that $t_i = Ki$ where K is a constant. Thus we have a linear combination of p exponentials with observations taken at $2n\ p$ equal intervals of length K, and with m observations made at each point Ki. The total number of parameters is $2p$, while the number of points at which observations are made in $2n\ p$, an integral multiple of the number of parameters.

It is required in the model that $\alpha_k \neq 0$ and $\lambda_k > 0$, $k = 1, 2, \ldots, p$, and this is realistic in most practical situations. However, a procedure like that presented in this paper could be developed for negative λ_k. The λ_k are further restricted so that $\lambda_r \neq \lambda_s$ when $r \neq s$. Moreover, the errors e_{ij} are assumed to be identically and independently distributed with mean zero and common variance σ^2. Actually, in order to carry out the estimation procedure developed in this chapter, it is only necessary to assume that the errors e_{ij} each have zero mean and a finite variance. However, the additional conditions given here for the e_{ij} make it possible to study the distributions and properties of the estimators of the new procedure in later chapters.

The specification in the model that the number of observation points t_i be an integral multiple of the number of parameters imposes a severe practical limitation on the new method if it is to be applied exactly. This specification essentially requires that the experimenter decide how many exponential terms to fit before conducting his experiment. Otherwise, the number of observation points might not be an integral multiple of $2p$. However, approximate methods of circumventing this limitation will be introduced in Chapter VII. Two other conditions of the model which are subject to criticism are the requirements that the observation points be evenly spaced and that the same number, m, of observations be made at each observation point. Experimenters naturally tend to take more observations in intervals where the data seem more variable than in intervals where the data appear to level off near an asymptote. The model will be modified in Chapter VII so that the value of m may vary to some extent during an experiment and an approximate solution will be given for some situations in which the t_i are not evenly spaced.

Note that since the errors e_{ij} are only specified to be identically and independently distributed, negative observations are not precluded by our model. In most practical applications of exponential fitting, only positive observations are possible. However, we may want to specify a distribution for the e_{ij} admitting negative observations, and that is not too unrealistic, for if the error variance σ^2 is reasonably small, very few negative observations would be expected. Observe that the model also requires that the y_{ij} have homogeneous variances. This is in contrast to the assumption that the logarithms of the y_{ij} have homogeneous errors. As mentioned in the second chapter, the latter assumption is made when a single exponential is fitted by fitting a straight line to a semi-logarithmic plot of the data. Furthermore, the assumption in the model presented in this section that the e_{ij} have zero means gives the expectation,

$$\mathcal{E}(y_{ij}) = \alpha_1 e^{-\lambda_1 t_i} + \alpha_2 e^{-\lambda_2 t_i} + \ldots + \alpha_p e^{-\lambda_p t_i} \quad . \qquad (3.1.2)$$

In order to simplify the exposition of the new estimation procedure, three special cases of the basic model as well as three slight modifications of these three cases are differentiated throughout this paper. Case 1 refers to a single exponential with its coefficient while Case 2 denotes a linear combination of two exponentials. The general case, where p may be any positive integer, is referred to as Case 3. Thus the models for y_{ij} for these three cases are as follows:

Case 1: $$y_{ij} = \alpha e^{-\lambda t_i} + e_{ij} \; ; \quad i = 0, 1, 2, \ldots, 2n-1 \; ; \qquad (3.1.3)$$

Case 2: $$y_{ij} = \alpha_1 e^{-\lambda_1 t_i} + \alpha_2 e^{-\lambda_2 t_i} + e_{ij} \; ; \; i = 0, 1, 2, \ldots, 4n-1 \; ; \qquad (3.1.4)$$

Case 3: $y_{ij} = \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i} + e_{ij}$; $i = 0, 1, 2, \ldots, 2n\,p-1$.

(3.1.5)

In addition to the cases already defined, the new estimation procedure may be applied to what Keeping [17] calls a modified exponential function. The model for this function is the same as that for Case 1 except for the addition of a constant term. A modified linear combination of exponentials may be defined similarly for p greater than one. Thus, corresponding to Cases 1, 2, and 3, we have Cases 4, 5, and 6 respectively, where the models for y_{ij} are as follows:

Case 4: $y_{ij} = \alpha_0 + \alpha_1 e^{-\lambda t_i} + e_{ij}$; $i = 0, 1, 2, \ldots, 3n-1$;

(3.1.6)

Case 5: $y_{ij} = \alpha_0 + \alpha_1 e^{-\lambda_1 t_i} + \alpha_2 e^{-\lambda_2 t_i} + e_{ij}$;

$i = 0, 1, 2, \ldots, 5n-1$; (3.1.7)

Case 6: $y_{ij} = \alpha_0 + \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i} + e_{ij}$;

$i = 0, 1, 2, \ldots, (2p+1)n-1$. (3.1.8)

For each of the six cases defined above, the subscript j ranges from one to m . The conditions presented in this section for the coefficients

α_k , the exponents λ_k , the observation points t_i and the errors e_{ij} apply, of course, to each of the six special cases and are the same throughout this paper unless otherwise stated.

3.2 The Estimation Procedure for Case 3

The new estimation procedure associated with the model presented in Section 3.1 is conceptually simple. First the domain of the observation points t_i is divided into as many intervals of equal length T as there are parameters in the model. Thus, for Case 3 the t_i are separated into $2p$ groups. Since there are $2p\,n$ equally spaced points t_i , each such group will contain n points t_i . Included in the first group will be t_0 , t_1 , ..., t_{n-1} , included in the second will be t_n , t_{n+1}, ..., t_{2n-1}, and in general $t_{(q-1)n}$, $t_{(q-1)n+1}$, ..., t_{qn-1} will be included in the q^{th} group. There are several other ways in which the t_i could be grouped without essentially changing the estimation procedure, but in Section 5.5 it will be shown that the grouping given here is in certain respects optimum.

The next step in the estimation procedure is to let S_q be the sum of all the observations made at the points t_i included in the q^{th} group. That is, set

$$S_q = \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} y_{ij} , \quad q = 1, 2, \ldots, 2p , \qquad (3.2.1)$$

where according to the model for Case 3 given by (3.1.5),

$$y_{ij} = \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i} + e_{ij} .$$

Since each e_{ij} has mean zero in accordance with the assumptions listed for our model in Section 3.1, the expectation of y_{ij} is given by

$$\mathcal{E}(y_{ij}) = \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i} . \tag{3.2.2}$$

But $t_i = Ki$ in our model. Therefore, substitution of (3.2.2) into (3.2.1) yields

$$\mathcal{E}(S_q) = m \sum_{k=1}^{p} \sum_{i=(q-1)n}^{qn-1} \alpha_k \Lambda_k^i , \tag{3.2.3}$$

where

$$\Lambda_k = e^{-\lambda_k K} . \tag{3.2.4}$$

The right side of equation (3.2.3) is a geometric series which may be summed to give

$$\mathcal{E}(S_q) = m \sum_{k=1}^{p} \alpha_k \Lambda_k^{(q-1)n} \frac{(1 - \Lambda_k^n)}{1 - \Lambda_k} , \quad q = 1, 2, \ldots, 2p . \tag{3.2.5}$$

The equation (3.2.5) is actually a set of 2p equations for the 2p parameters of our model in terms of the expectations $\mathcal{E}(S_q)$. In order that estimators for the parameters may be obtained from these equations, we set the observation sums S_q equal to their expected values $\mathcal{E}(S_q)$. The result is the set of equations,

$$S_q = m \sum_{k=1}^{p} \hat{\alpha}_k \hat{\Lambda}_k^{(q-1)n} \frac{(1-\hat{\Lambda}_k^n)}{1-\hat{\Lambda}_k}, \quad q = 1, 2, \ldots, 2p, \qquad (3.2.6)$$

which define the estimators $\hat{\alpha}_k$ and $\hat{\Lambda}_k = \exp.(-\hat{\lambda}_k K)$, $k = 1, 2, \ldots, p$. It is interesting to note that our procedure of reducing the observations to as many sums as there are parameters in the model, substituting observations for their expectations, and then solving for estimates of the parameters is similar to Wald's method [29] for fitting a linear regression with error in both of the variates.

In order to complete the new estimation procedure we need only solve the set of equations (3.2.6) for $\hat{\alpha}_k$ and $\hat{\lambda}_k$. In the next section we shall represent (3.2.6) in matrix notation and proceed with a direct solution for Case 3 which makes extensive use of properties of certain symmetric functions. Then in Section 3.4 a much shorter solution will be given which utilizes certain results from the calculus of finite differences. Illustrations of the procedure for the special Cases 1 and 2 will be presented in Section 3.5 before the development of the procedure for Case 6 is given in Section 3.6. Section 3.7 will be concerned with the relationship of the new estimation procedure to the Prony method outlined in Chapter II.

3.3 A Solution for Case 3 Estimators

To facilitate the solution of the set of equations (3.2.6) for estimates of λ_k and α_k, we shall represent the equations in matrix notation. Also, in order to simplify the writing of this section, the carets which designate the estimators in equation (3.2.6) will be dropped until the final steps in the solution of (3.2.6) for the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$.

Let α be a column vector of the coefficients α_k, so that*

$$\alpha^T = (\alpha_1, \alpha_2, \ldots, \alpha_p) .$$

Also, define column vectors $\bar{s}$ and $\underline{s}$ in such a way that

$$\bar{s}^T = (S_1 , S_2 , \ldots, S_p) ,$$

$$\underline{s}^T = (S_{p+1} , S_{p+2} , \ldots, S_{2p}) .$$

Let L be a p by p matrix with elements**

$$\ell_{rs} = \Lambda_s^{(r-1)n} .$$

Then define two p by p diagonal matrices W and V with elements

$$w_{rs} = \frac{m(1 - \Lambda_r^n)}{1 - \Lambda_r} \delta_{rs} ,$$

$$v_{rs} = \Lambda_r^{pn} \delta_{rs} ,$$

---

* A superscript T denotes the transpose of the matrix indicated.

** Unless otherwise indicated, when an element of a matrix or a determinant is defined in this paper, r refers to the row and s to the column in which the element is located.

where $\delta_{rs} = 1$, $r = s$, and $\delta_{rs} = 0$, $r \neq s$. Now we may represent the equations (3.2.6) which involve $S_1, S_2, \ldots, S_p$ as

$$LW\alpha = \bar{s} \quad . \tag{3.3.1}$$

Moreover, the equations involving $S_{p+1}, S_{p+2}, \ldots, S_{2p}$ may be written as

$$LWV\alpha = \underline{s} \quad . \tag{3.3.2}$$

Now the inverse of W is a diagonal matrix with diagonal elements $(1 - \Lambda_k)/m(1 - \Lambda_k^n)$, $k = 1, 2, \ldots, p$, and hence the inverse of W exists provided that $\Lambda_k \neq 1$. This is also a necessary condition for W itself to be defined. However, $\Lambda_k = \exp.(-\lambda_k K)$ and cannot equal one in accordance with our model since we have specified that the exponent $\lambda_k K > 0$. Therefore, both W and W^{-1} exist. Furthermore, the matrix L is an alternant matrix. Its inverse also exists under our model and is given below by (3.3.8). Since both W^{-1} and L^{-1} exist, it is permissible to premultiply both sides of equation (3.3.1) by $W^{-1} L^{-1}$ to obtain

$$\alpha = W^{-1} L^{-1} \bar{s} \quad .$$

This result, substituted in (3.3.2), gives

$$LWVW^{-1} L^{-1} \bar{s} = \underline{s} \quad . \tag{3.3.3}$$

But $WVW^{-1} = WW^{-1}V = V$ because diagonal matrices are commutable in multiplication. Hence equation (3.3.3) reduces to

$$LVL^{-1}\,\bar{s} = \underline{s} \quad . \qquad (3.3.4)$$

Since the matrix W does not appear in (3.3.4), we no longer have to deal with terms of the form $(1 - \Lambda_k^n)/(1 - \Lambda_k)$ in the solution for the λ_k , and each time a Λ_k appears in (3.3.4) it has a multiple of n as an exponent. Therefore, as a further simplification, we let

$$x_k = \Lambda_k^n = e^{-\lambda_k Kn} \quad . \qquad (3.3.5)$$

Now the elements of the matrices L and V of (3.3.4) may be written respectively as

$$\ell_{rs} = x_s^{r-1} \quad ,$$

$$v_{rs} = x_r^p \, \delta_{rs} \quad .$$

To solve (3.3.4) for estimates of the λ_k , we need to know L^{-1} . But since L is an alternant matrix, the form of its inverse is well known and is indicated, for instance, by Aitken [1, pg. 118]. However, in order to write L^{-1} in a concise form, we must first define several terms. Denote by C the set of elements $x_1, x_2, \ldots, x_p$ and by C_i the set C with the element x_i deleted. Then let $E_r(C)$ and $E_r(C_i)$

be the sums of all possible distinct products of the elements of the sets C and C_i respectively taken r at a time. That is, define $E_r(C)$ and $E_r(C_i)$ to be the r^{th} order elementary symmetric functions of C and C_i respectively. E_r was defined previously in Section 2.2, but with respect to a different set of elements than those considered here. An elementary symmetric function of order zero is defined to be one. Now, if we let

$$D = \prod_i \prod_j (x_j - x_i) \tag{3.3.6}$$

and

$$D_h = \prod_{\substack{i \\ i,j \neq h}} \prod_j (x_j - x_i) \, , \tag{3.3.7}$$

we can show as an extension of Aitken's discussion that L^{-1} has elements

$$\ell^{rs} = (-1)^{r+s} \frac{D_r}{D} E_{p-s} (C_r) \, . \tag{3.3.8}$$

Continuing with our solution of (3.3.4), let us consider the elements u_{rs} of the p by p matrix $U = LVL^{-1}$. Carrying out the matrix multiplication indicated by the definition of U , we find that

$$u_{rs} = \frac{1}{D} \sum_{i=1}^{p} (-1)^{i+s-2} D_i x_i^{p+r-1} E_{p-s}(C_i); \; r,s=1,2,\ldots, p \, . \tag{3.3.9}$$

We wish to simplify the expression (3.3.9) for these elements u_{rs} . To do this, we shall make extensive use of two mathematical properties.

The first property with which we are concerned is given by the equation

$$x_i E_r (C_i) = E_{r+1} (C) - E_{r+1} (C_i) , \qquad (3.3.10)$$

and it follows directly from the definitions of C , C_i and E_r . The second property is concerned with the quotient

$$Q_k = \frac{(-1)^{p-1}}{D} \sum_{i=1}^{p} (-1)^{i-1} x_i^k D_i , \qquad (3.3.11)$$

where k is a nonnegative integer. It is shown below that

$$\begin{aligned} Q_k &= H_{k-p+1} , \qquad k \geq p-1 \\ &= 0 \qquad , \qquad k < p-1 , \end{aligned} \qquad (3.3.12)$$

given that

$$H_r = \sum_{\substack{\sum_{i=1}^{p} \alpha_i = r \\ \alpha_i \geq 0}} \prod_{i=1}^{p} x_i^{\alpha_i} , \qquad (3.3.13)$$

where the summation is over all possible permutations of the nonnegative integers α_i for which the sum of the α_i is r. H_r is called the complete homogeneous symmetric function of degree r, and H_0 is defined to be one.

To prove the relationship given by (3.3.12) in conjunction with (3.3.13), we first note that D, as defined by (3.3.6), equals the alternant $|L|$. This result is given, for instance, in [1, pg. 112]. Also, each minor of $|L|$ is an alternant determinant. In particular, the cofactor of any element ℓ_{ps} in the last row of $|L|$ is an alternant of order $(p - 1)$ and equals D_r as defined by (3.3.7). Now let us define another alternant $|L_k|$ which is like $|L|$ except that the elements in the last row of $|L_k|$ are raised to the k^{th} power instead of the $(p - 1)^{st}$ power. If we expand $|L_k|$ in terms of the elements of its last row, each element has the same cofactor as the corresponding element of $|L|$, and therefore

$$|L_k| = x_p^k D_p - x_{p-1}^k D_{p-1} + \dots + (-1)^{p-1} x_1^k D_1 \qquad (3.3.14)$$

$$= (-1)^{p-1} \sum_{i=1}^{p} (-1)^{i-1} x_i^k D_i \ .$$

Hence, since $D = |L|$, a glance at (3.3.11) reveals that

$$Q_k = \frac{|L_k|}{|L|} \ . \qquad (3.3.15)$$

But when $k < (p - 1)$, two rows of $|L_k|$ are the same, and therefore $|L_k| = Q_k = 0$, $k < (p - 1)$. And when $k \geqslant (p - 1)$,

$$\frac{|L_k|}{|L|} = H_{k-p+1} \quad , \tag{3.3.16}$$

a result given in [28, pg. 150]. Therefore, referring to (3.3.15), we conclude that $Q_k = H_{k-p+1}$, $k \geq p - 1$. Thus equation (3.3.12) is correct.

Now that the properties given by equations (3.3.10) and (3.3.12) have been established, let us proceed with our work on the elements u_{rs} of the U matrix, as given by (3.3.9). Applying (3.3.10) and (3.3.12) to (3.3.9) repeatedly, we deduce that

$$u_{rs} = (-1)^{p-s} E_{p-s+1}(C)\, Q_{p+r-2} + (-1)^{p-s+1} E_{p-s+2}(C)\, Q_{p+r-3} + \ldots$$

$$\ldots + (-1)^{p-1} E_p(C)\, Q_{p+r-s-1} \; ; \quad r,s=1,2, \ldots, p \; . \tag{3.3.17}$$

Let us refer to $E_r(C)$ simply as E_r . Now we may bring equation (3.3.12) into play to evaluate the Q_k factors in (3.3.17) giving the general result that

$$u_{rs} = (-1)^{p-s} E_{p-s+1} H_{r-1} + (-1)^{p-s+1} E_{p-s+2} H_{r-2} + \ldots$$

$$\ldots + (-1)^{p-1} E_p H_{r-s} \; ; \; r,s=1,2, \ldots, p \; ; \tag{3.3.18}$$

where a function E_r with $r > p$ or a function H_r with $r < 0$

is defined to be zero.

Each term in the expression (3.3.18) is the product of an E_r function and an H_r function. Furthermore, for any element u_{rs} in the r^{th} row of U, the H_r function in the first term has subscript $(r - 1)$, and this subscript decreases by one with each successive term until for some term either H_r or E_r vanishes. Similarly, the subscripts on the E_r functions increase by one with each successive term. Hence, the matrix U , which has the u_{rs} as elements, may be represented as the product of two triangular matrices. For if we define p by p triangular matrices $\bar{E}$ and H with elements $(-1)^{r-s+1} E_{p+r-s}$ and H_{r-s} respectively, we can see that

$$H\bar{E} = U \quad . \tag{3.3.19}$$

Since $U = LVL^{-1}$, (3.3.4) now may be written as

$$H \bar{E} \bar{s} = \underline{s} \quad . \tag{3.3.20}$$

But the matrix H has a simple inverse [1, pg. 115] , namely the p by p triangular matrix $\underline{E}$ with elements $(-1)^{r-s} E_{r-s}$. Thus, premultiplying both sides of (3.3.20) by $\underline{E}$, we have

$$\bar{E} \bar{s} = \underline{E} \, \underline{s} \tag{3.3.21}$$

The matrix equation (3.3.21) when multiplied out yields a set of p linear equations in the p elementary symmetric functions E_r .

These equations are

$$E_p S_{w+1} - E_{p-1} S_{w+2} + E_{p-2} S_{w+3} - \dots + (-1)^{p-1} E_1 S_{w+p} = (-1)^{p-1} S_{w+p+1} ,$$

$$w = 0, 1, 2, \dots, p-1 . \qquad (3.3.22)$$

The sums S_q are known functions of the observations and therefore the equations (3.3.22) may be solved by elementary means for an estimator $\hat{E}_r$ of any E_r . This solution may be presented as

$$\hat{E}_r = \frac{|R_{p-r+1}|}{|R|} , \quad r = 0, 1, 2, \dots, p , \qquad (3.3.23)$$

where $|R|$ is a p by p persymmetric determinant with elements S_{r+s-1} and $|R_k|$ is a p by p determinant with elements S_{r+s-1} , $r \leq k - 1$, and S_{r+s} , $r > k$.

Since the functions E_r are the elementary symmetric functions of the x_k , each x_k is a root of the equation

$$x^p - E_1 x^{p-1} + E_2 x^{p-2} - \dots + (-1)^{p-1} E_{p-1} x + (-1)^p E_p = 0 . \qquad (3.3.24)$$

Thus, after substitution of $\hat{E}_r$ for E_r in this equation for $r = 1, 2, \dots, p$, estimators $\hat{x}_k$ of the x_k may be obtained by finding the p roots of the equation

$$x^p - \hat{E}_1 x^{p-1} + \hat{E}_2 x^{p-2} - \dots + (-1)^{p-1} \hat{E}_{p-1} x + (-1)^p \hat{E}_p = 0 . \qquad (3.3.25)$$

Because of the symmetry of the coefficients in equation (3.3.25), the subscripts of the estimators $\hat{x}_k$ may be assigned in any order.

During the preceding exposition we have seen how the new estimation procedure leads to estimators $\hat{x}_k$ of the exponentials x_k. In the remainder of this chapter we shall assume that these $\hat{x}_k$ solutions yielded by equation (3.3.25) are admissible estimators. That is, we will assume that each root $\hat{x}_k$ is real with $0 < \hat{x}_k < 1$, the range dictated by the model for the parameters x_k. Conditions which are necessary for equation (3.3.25) to have such admissible roots are discussed in Section 5.6.

Now we wish to make use of the $\hat{x}_k$ to compute estimators of the parameters in our model. Since by definition, $x_k = \exp.(-\lambda_k Kn)$, we take as our estimators of the λ_k,

$$\hat{\lambda}_k = -\frac{1}{Kn} \ln \hat{x}_k \,, \quad k = 1, 2, \ldots, p \,. \qquad (3.3.26)$$

Then substitution of the $\hat{x}_k$ in (3.3.1) determines a set of linear equations in α_k which may be solved easily for estimators $\hat{\alpha}_k$ of the coefficients α_k. In summation notation, these equations are

$$m \sum_{k=1}^{p} \hat{x}_k^{q-1} \frac{(1 - \hat{x}_k)}{1 - \hat{x}_k^{1/n}} \hat{\alpha}_k = S_q \,, \quad q = 1, 2, \ldots, p \,. \qquad (3.3.27)$$

3.4 An Alternative Solution for Case 3 Estimators

In this section the equation (3.2.6) will be transformed into the set of equations (3.3.22) by another method. Suppose we let

$$G_k = m\,\hat{\alpha}_k\;\frac{(1-\hat{\Lambda}_k^n)}{1-\hat{\Lambda}_k}\quad . \qquad (3.4.1)$$

Now recalling the definition of x_k given by equation (3.3..5), we may write equation (3.2.6) as

$$S_q = \sum_{k=1}^{p} G_k\,\hat{x}_k^{q-1}\ , \quad q = 1, 2, \ldots, 2p\ . \qquad (3.4.2)$$

But functions, such as the S_q , of a discontinuous variable which, like q , takes on only integral values are known to satisfy a difference equation which is of the same order as the number of unknown parameters in each function (see [16]). Therefore, if we consider the $\hat{x}_k$ as known, the S_q satisfy a p^{th} order difference equation which may be used to find the estimators $\hat{x}_k$. Moreover, for a polynomial such as (3.4.2), Householder [14] shows that this difference equation is of the form (3.3.22) provided that the $\hat{x}_k$ are all distinct. Now the parameters x_k are all distinct since our model specifies that $\lambda_r \neq \lambda_s$, $r \neq s$. Therefore, if the $\hat{x}_k$ are admissible estimators of the x_k , they are also distinct, and equation (3.3.22) is correct.

Now that we have again arrived at the difference equation (3.3.22), the estimators $\hat{x}_k$, $\hat{\lambda}_k$, and $\hat{\alpha}_k$ may be obtained in the same manner as before. However, the calculus of finite differences [16] does lead to the further results that the $\hat{x}_k$ are linearly independent and that both the $\hat{x}_k$ and the $\hat{\alpha}_k$ are unique provided that the $\hat{x}_k$ are all distinct.

3.5 Estimation for Cases 1 and 2

We shall now illustrate the new estimation procedure for Cases 1 and 2. Additional examples are given in Chapter VII. The models for these two cases, as defined in Section 3.1, may be obtained from the model for Case 3 by setting p equal to one and two respectively.

For Case 1 we have estimators $\hat{x}$, $\hat{\lambda}$ and $\hat{\alpha}$ to calculate for the parameters given in equation (3.1.3). Note that the subscript k is dropped for Case 1 because it always equals one in this instance. Now there are two parameters in the model for Case 1, so we need to calculate two sums S_q. These sums are

$$S_1 = \sum_{i=0}^{n-1} \sum_{j=1}^{m} y_{ij} \quad , \qquad (3.5.1)$$

$$S_2 = \sum_{i=n}^{2n-1} \sum_{j=1}^{m} y_{ij} \quad . \qquad (3.5.2)$$

Also, for Case 1 there is only one elementary symmetric function, namely E_1, and therefore $\hat{E}_1 = \hat{x}$. So, substitution in equation (3.3.23) with $p = 1$ yields

$$\hat{x} = \frac{S_2}{S_1} \quad . \qquad (3.5.3)$$

Furthermore, from equations (3.3.26) and (3.3.27) we have that

$$\hat{\lambda} = -\frac{1}{Kn} \ln\left(\frac{S_2}{S_1}\right) \quad , \qquad (3.5.4)$$

$$\hat{\alpha} = \frac{\left[1 - \left(\frac{S_2}{S_1}\right)^{1/n}\right] S_1^2}{m(S_1 - S_2)} \quad . \qquad (3.5.5)$$

Let us apply these Case 1 estimation equations to the data in Table 1. This data is not actual experimental data, but it is typical of data, say, for successive determinations of the activity present in a solution containing a pure radioactive substance. There are an even number of points t_i , so the number of observation points is an integral multiple of the number of parameters as required by our model. This integral multiple, n , is equal to four. K , the constant length of the

Table 1

ACTIVITY DETERMINATIONS VERSUS TIME

Time t_i	0	1	2	3	4	5	6	7
Dosage y_i	6.81	4.70	3.23	2.24	1.55	1.07	0.74	0.51

intervals between successive t_i , is one. Also, m , the number of observations for each t_i, is one, so we drop the subscript j . Now if we group the t_i as suggested in Section 3.2, we shall put t_0 , t_1 , t_2, and t_3 in the first group and t_4 , t_5 , t_6 , and t_7 in the second. The corresponding observation sums, as given by (3.5.1) and (3.5.2), are

$$S_1 = 6.81 + 4.70 + 3.23 + 2.24 = 16.98 ,$$

$$S_2 = 1.55 + 1.07 + .74 + .51 = 3.87 \quad .$$

Then substituting S_1 and S_2 into equations (3.5.3), (3.5.4) and (3.5.5), we find that

$$\hat{x} = .228 \ , \quad \hat{\lambda} = .370 \ , \quad \hat{\alpha} = 6.797 \ .$$

Thus we may represent the data in Table 1 by the function

$$\hat{y}_i = 6.797 \, e^{-.370 t_i} \quad .$$

For Case 2, four sums S_q are required and are given by

$$S_q = \sum_{(q-1)n}^{qn-1} \sum_{j=1}^{m} y_{ij} \ , \ q = 1, 2, 3, 4 \quad . \qquad (3.5.6)$$

There are also two functions E_r to be estimated using equation (3.3.23). With $p = 2$, this expression for the $\hat{E}_r$ yields

$$\hat{E}_1 = \frac{\begin{vmatrix} S_1 & S_3 \\ S_2 & S_4 \end{vmatrix}}{\begin{vmatrix} S_1 & S_2 \\ S_2 & S_3 \end{vmatrix}} = \frac{S_1 S_4 - S_2 S_3}{S_1 S_3 - S_2^2} \ , \qquad (3.5.7)$$

$$\hat{E}_2 = \frac{\begin{vmatrix} S_2 & S_3 \\ S_3 & S_4 \end{vmatrix}}{\begin{vmatrix} S_1 & S_2 \\ S_2 & S_3 \end{vmatrix}} = \frac{S_2 \, S_4 - S_3^2}{S_1 \, S_3 - S_2^2} \quad . \tag{3.5.8}$$

Now the exponential estimates $\hat{x}_1$ and $\hat{x}_2$ for Case 2 are roots of the equation

$$x^2 - \hat{E}_1 \, x + \hat{E}_2 = 0 \quad ,$$

which corresponds to equation (3.3.25) for the general Case 3. Therefore, we may take

$$\hat{x}_1 = \frac{1}{2}\left[E_1 + (E_1^2 - 4 \, E_2)^{\frac{1}{2}} \right] \quad , \tag{3.5.9}$$

$$\hat{x}_2 = \frac{1}{2}\left[E_1 - (E_1^2 - 4 \, E_2)^{\frac{1}{2}} \right] \quad .$$

Finally, from (3.3.26) and (3.3.27) we find that

$$\hat{\lambda}_k = - \frac{1}{Kn} \quad \ell n \, \hat{x}_k \quad , \quad k = 1, 2 \quad , \tag{3.5.10}$$

$$\hat{\alpha}_1 = \frac{(1 - \hat{x}_1^{1/n})(S_1 \, \hat{x}_2 - S_2)}{m(1 - \hat{x}_1)(\hat{x}_2 - \hat{x}_1)} \quad , \tag{3.5.11}$$

$$\hat{\alpha}_2 = \frac{(1 - \hat{x}_2^{1/n})(S_2 - S_1 \hat{x}_1)}{m(1 - \hat{x}_2)(\hat{x}_2 - \hat{x}_1)} \quad . \qquad (3.5.12)$$

The estimation for Case 2 may be illustrated using the fictional data in Table 2. Suppose an experimental animal is injected with a test material at time zero. This data purports to represent the concentration of the injected material in the animal at time t_i measured through the cumulative per cent of excretion of that material up to time t_i .

Table 2

CUMULATIVE PER CENT EXCRETION y_i VERSUS TIME t_i

t_i	0	1	2	3	4	5	6	7	8	9	10	11
y_i	0.60	1.82	2.84	3.72	4.40	4.99	5.49	5.86	6.19	6.42	6.65	6.76

Since there are twelve observations and four parameters to be estimated, $n = 3$. As before, K and m are both one. Using equation (3.5.6), we find that

$$S_1 = 0.60 + 1.82 + 2.84 = 5.26 \quad ,$$

$$S_2 = 3.72 + 4.40 + 4.99 = 13.11 \quad ,$$

$$S_3 = 5.49 + 5.86 + 6.19 = 17.54 \quad ,$$

$$S_4 = 6.42 + 6.65 + 6.76 = 19.83 \quad .$$

Then from equations (3.5.7) and (3.5.8) we obtain

$$\hat{E}_1 = 1.5782 \quad , \qquad \hat{E}_2 = 0.5989 \quad .$$

These estimates, substituted into equations (3.5.9), lead to the exponential estimates

$$\hat{x}_1 = 0.9433 \quad , \qquad \hat{x}_2 = 0.6349 \quad .$$

These estimates in turn substituted into equations (3.5.10), (3.5.11), and (3.5.12) give

$$\hat{\lambda}_1 = 0.0195 \quad , \qquad \hat{\lambda}_2 = 0.1514 \quad ,$$

$$\hat{\alpha}_1 = 10.784 \quad , \qquad \hat{\alpha}_2 = -10.167 \quad .$$

Hence the data in Table 2 may be fitted by the equation

$$\hat{y}_i = 10.784\, e^{-0.0195 t_i} - 10.167\, e^{-0.1514 t_i} \quad .$$

3.6 <u>Estimation for Cases 4, 5, and 6</u>

Now that the estimation procedure has been developed for Cases 1 and 2 in particular as well as for Case 3, the results for Case 3 may be utilized to obtain estimators for Case 6. From equation (3.1.8)

$$y_{ij} = \alpha_0 + \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i} + e_{ij} \quad , \quad i = 0, 1, 2, \ldots, (2p+1)n-1 \qquad (3.6.1)$$

under our Case 6 model. Sums S_q may be formed from the observations y_{ij} in a manner similar to that in which they were constructed for Case 3 to give

$$S_q = \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} y_{ij} \ , \ q = 1, 2, \ldots, 2p+1 \ , \qquad (3.6.2)$$

the only difference being in the range of the subscript q . Taking the expectations of the sums S_q , we have

$$\mathcal{E}(S_q) = mn\,\alpha_0 + m \sum_{k=1}^{p} \alpha_k \Lambda_k^{(q-1)n} \frac{(1 - \Lambda_k^n)}{1 - \Lambda_k} \ , \qquad (3.6.3)$$

$$q = 1, 2, \ldots, 2p+1 \ ,$$

where Λ_k is still defined by equation (3.2.4). As before, we equate the sums S_q to their expectations to arrive at the equations

$$S_q = mn\,\hat{\alpha}_0 + \sum_{k=1}^{p} \hat{\alpha}_k \hat{\Lambda}_k^{(q-1)n} \frac{(1 - \hat{\Lambda}_k^n)}{1 - \hat{\Lambda}_k} \ , \qquad (3.6.4)$$

Is there an m missing here? $m \Sigma \hat{\alpha}_k$

$$q = 1, 2, \ldots, 2p+1 \ .$$

At this point the solution for Case 6 diverges from the pattern established for Case 3. We now perform a series of subtractions to obtain the differences

$$S'_q = S_q - S_{q+1} \, , \quad q = 1, 2, \ldots, 2p \quad . \qquad (3.6.5)$$

In forming the S'_q we eliminate $\hat{\alpha}_0$ from our original set of $(2p+1)$ equations, giving the $2p$ new equations

$$S'_q = m \sum_{k=1}^{p} \hat{\alpha}_k \hat{\Lambda}_k^{(q-1)n} \frac{(1 - \hat{\Lambda}_k)^2}{1 - \hat{\Lambda}_k} \, , \quad q = 1, 2, \ldots, 2p \quad . \qquad (3.6.6)$$

Next we define the matrices $\bar{s}$ and $\underline{s}$ of Section 2.5 in terms of the S'_q instead of the S_q, and we call the resulting matrices $\bar{s}'$ and $\underline{s}'$ respectively. Then we let W' be a diagonal matrix with elements

$$w'_{rs} = \frac{m(1 - \Lambda_r^n)^2}{1 - \Lambda_r} \delta_{rs} \quad .$$

Now the set of equations (3.6.6) can be represented by the two matrix equations

$$L \, W' \alpha = \bar{s}' \quad , \qquad (3.6.7)$$

$$L \, W' \, V \alpha = \underline{s}' \quad , \qquad (3.6.8)$$

which correspond to equations (3.3.1) and (3.3.2) for Case 3.

Proceeding as before, we may eliminate the vector α from the equations (3.6.7) and (3.6.8) to obtain the equation

$$L\,V\,L^{-1}\,\bar{s}' = \underline{s}' \quad . \tag{3.6.9}$$

This equation is of the same form in terms of the S'_q as (3.3.4) is in terms of the S_q, and like (3.3.4) it may be solved for estimators of the exponents λ_k. Thus, the solution for estimators of the parameters λ_k is the same for Case 6 in terms of the S'_q as it is for Case 3 in terms of the sums S_q.

From the above discussion, it follows that to estimate the parameters of a Case 6 model it is first necessary to compute the statistics S'_q in accordance with equation (3.6.5). The next step is to substitute the S'_q for the corresponding S_q in the solution already derived for Case 3 to obtain estimators $\hat{\lambda}_k$ of the exponents λ_k. These estimators, when substituted for the λ_k in equation (3.6.7), lead to a set of linear equations in the coefficients α_k, for $k > 0$, which may be easily solved for estimators $\hat{\alpha}_k$ of the α_k. If we again let $\hat{x}_k = \exp.(-\hat{\lambda}_k Kn)$, this set of linear equations which yields the $\hat{\alpha}_k$ may be written as

$$m \sum_{k=1}^{p} \hat{x}_k^{q-1} \frac{(1 - \hat{x}_k)^2}{1 - \hat{x}_k^{1/n}} \hat{\alpha}_k = S'_q \quad , \quad q = 1, 2, \ldots, p \quad . \tag{3.6.10}$$

These equations correspond to the set (3.3.27) which were derived earlier for Case 3. Finally, an estimator $\hat{\alpha}_0$ of α_0 may be found for Case 6 by substitution for the $\hat{x}_k$ and the rest of the $\hat{\alpha}_k$ in the equation

$$\hat{\alpha}_0 = \frac{1}{mn}\left[S_1 - m \sum_{k=1}^{p} \hat{\alpha}_k \frac{(1 - \hat{x}_k)}{1 - \hat{x}_k^{1/n}} \right] . \tag{3.6.11}$$

Now that the new estimation procedure has been presented for the general Case 6 model, estimators for Cases 4 and 5 may be easily determined by setting p equal to one and two respectively in the Case 6 derivation. In this way it can be shown that for Case 4,

$$\hat{x} = \left(\frac{S_2 - S_3}{S_1 - S_2} \right) , \tag{3.6.12}$$

$$\hat{\lambda} = - \frac{1}{Kn} \log \hat{x} = - \frac{1}{Kn} \log \left(\frac{S_2 - S_3}{S_1 - S_2} \right) , \tag{3.6.13}$$

$$\hat{\alpha}_0 = \frac{S_2 - S_1 \hat{\Lambda}^n}{mn(1 - \hat{\Lambda}^n)} = \frac{S_1 S_3 - S_2^2}{mn(S_1 - 2S_2 + S_3)} , \tag{3.6.14}$$

$$\hat{\alpha}_1 = \frac{(S_1 - mn\,\hat{\alpha}_0)(1 - \hat{\Lambda})}{m(1 - \hat{\Lambda}^n)} = \frac{(S_1 - S_2)^3 \left[1 - \left(\frac{S_2 - S_3}{S_1 - S_2} \right)^{\frac{1}{n}} \right]}{m(S_1 - 2S_2 + S_3)^2} . \tag{3.6.15}$$

Also, for Case 5,

$$\hat{E}_1 = \frac{S_1S_4 - S_1S_5 + S_2S_5 - S_2S_3 + S_3^2 - S_3S_4}{S_1S_3 - S_1S_4 + S_2S_3 + S_2S_4 - S_2^2 - S_3^2} , \qquad (3.6.16)$$

$$\hat{E}_2 = \frac{S_3 - S_4 + \hat{E}_1(S_3 - S_2)}{S_2 - S_1} . \qquad (3.6.17)$$

Estimators $\hat{x}_1$, $\hat{x}_2$, $\hat{\lambda}_1$ and $\hat{\lambda}_2$ for Case 5 are obtained from $\hat{E}_1$ and $\hat{E}_2$ in the same way as they were for Case 2 in Section 3.5. Subsequently estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\alpha}_0$ are given by

$$\hat{\alpha}_1 = \frac{\left[(S_2 - S_1)(1 + \hat{x}_2) - S_3 + S_1\right] (1 - \hat{x}_1^{1/n})}{m(1 - \hat{x}_1)^2 (\hat{x}_1 - \hat{x}_2)} , \qquad (3.6.18)$$

$$\hat{\alpha}_2 = \frac{\left[(S_2 - S_1)(1 + \hat{x}_1) - S_3 + S_1\right] (1 - \hat{x}_2^{1/n})}{m(1 - \hat{x}_2)^2 (\hat{x}_2 - \hat{x}_1)} , \qquad (3.6.19)$$

$$\hat{\alpha}_0 = \frac{1}{mn}\left[S_1 - m\,\hat{\alpha}_1 \frac{(1 - \hat{x}_1)}{1 - \hat{x}_1^{1/n}} - m\,\hat{\alpha}_2 \frac{(1 - \hat{x}_2)}{1 - \hat{x}_2^{1/n}} \right] . \qquad (3.6.20)$$

As an illustration of a Case 6 type estimation, let us again fit the data in Table 2, but this time to a Case 4 model. Since there are now only three parameters to be estimated with the twelve observations, $n = 4$. As before, $K = m = 1$. Now from (3.6.2) it follows that

$$S_1 = 0.60 + 1.82 + 2.84 + 3.72 = 8.98 \quad ,$$

$$S_2 = 4.40 + 4.99 + 5.49 + 5.86 = 20.74 \quad ,$$

$$S_3 = 6.19 + 6.42 + 6.65 + 6.76 = 26.02 \quad .$$

Thus (3.6.12) yields

$$\hat{x} = \frac{-5.28}{-11.76} = 0.44898$$

and (3.6.13) gives

$$\hat{\lambda} = -\frac{1}{4} \ln (0.44898) = 0.2002 \quad .$$

Then, substituting into (3.6.14) and (3.6.15), we find that

$$\hat{\alpha}_0 = 7.5806 \quad , \quad \hat{\alpha}_1 = -7.0272 \quad .$$

Hence the data in Table 2 may also be fitted by the equation

$$\hat{y}_i = 7.5806 - 7.0272 \, e^{-0.2002 \, t_i} \quad .$$

3.7 A Comparison with the Prony Method

In the estimation procedure which has been presented in this chapter, sums S_q of observations are substituted in the equation for the expectations $\mathcal{E}(S_q)$ of these sums. Then in the course of the solution of the resulting equations, the symmetric functions $\hat{E}_r$ of the $\hat{x}_k$ may be presented as the ratio (3.3.23) of two determinants. The elements of these determinants are the sums S_q. But if we divide each S_q by mn, we shall divide both the numerator and the denominator of the ratio (3.3.23) by $(mn)^p$, which will leave the equations unchanged except for a substitution of groups means for the corresponding sums S_q. Thus we may obtain the same estimators by using arithmetic means instead of sums, and the estimation procedure may alternatively be thought of as one in which means of observed values are substituted for their expectations.

Another interesting characteristic of the new estimation procedure is its similarity to the Prony method outlined in Chapter II. In fact, a comparison of (3.3.22) and (2.2.2) shows that these two equations are of the same form, but that (2.2.2) involves observations y_i while (3.3.22) can be expressed in terms of the sums S_q or the corresponding arithmetic means. So it might appear that the new method is comprised of the application of the Prony method to group means instead of individual observations. This is not the case however, for such an application of the Prony procedure would consist of taking each group mean to represent the mid-point expectation for that group and then fitting these means by Prony's method. But the expected values of the group means are not equal to the group mid-point expectations, and hence Prony's method applied in this way amounts to fitting an exponential model with parameters which differ from those defined in the new procedure. Thus the new procedure is essentially different from the Prony method. Moreover, the new procedure does not require that the number of parameters in the model be equal to the numbers of observations to be fitted as did Prony's original interpolation method. Neither does it involve questionable least squares calculations as does the estimation version of Prony's method presented by Whittaker and Robinson [30].

IV. THE LIMITING DISTRIBUTIONS OF THE ESTIMATORS

4.1 A Theorem on Limiting Distributions for Large m

Although the estimators produced by the new estimation procedure are comparatively easy to calculate, they are not simple enough to yield easily derived small sample distributions. Reference to analytical investigations of the small sample distributions of some of the estimators will be made in Chapter VI. Also, the results of an extensive empirical study of the distributions of the estimators for Case 1 will be presented there. Meanwhile, in this chapter the limiting distributions of all the estimators derived with the new estimation procedure will be determined. Furthermore, these distributions will be derived as either m, the number of observations made at each observation point t_i, or n, the number of t_i in each of the $2p$ partitions of the t_i, approaches infinity.

We shall first present a theorem from which limiting distributions of the estimators may be determined as $m \to \infty$ with n held fixed. This theorem is applicable to either Case 3 or Case 6. In the derivation for Case 3, $q = 1, 2, \ldots, p$, while for Case 6, $q = 1, 2, \ldots, 2p+1$. Now in Section 3.7 it was mentioned that the estimators obtained by the new procedure may be expressed alternatively as functions of group means instead of the group sums S_q. Denote these means by Y_q, where

$$Y_q = \frac{1}{mn} S_q \ , \tag{4.1.1}$$

and define the expectation of any Y_q to be η_q . From (4.1.1) it follows that

$$\eta_q = \frac{1}{mn} \mathcal{E}(S_q) \quad . \tag{4.1.2}$$

Then equations (3.2.3) and (3.2.5) for the $\mathcal{E}(S_q)$ lead to the relationships

$$\eta_q = \frac{1}{n} \sum_{i=(q-1)n}^{qn-1} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k K i} \tag{4.1.3}$$

$$= \frac{1}{n} \sum_{k=1}^{p} \alpha_k \Lambda_k^{(q-1)n} \frac{(1 - \Lambda_k^n)}{1 - \Lambda_k} \quad , \tag{4.1.4}$$

where $\Lambda_k = \exp.(-\lambda_k K)$. Y_q may be evaluated in a similar way from equations (3.2.1), (3.1.5) and (4.1.3) to give

$$Y_q = \frac{1}{mn} \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} y_{ij}$$

$$= \frac{1}{n} \sum_{i=(q-1)n}^{qn-1} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k K i} + \frac{1}{mn} \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} e_{ij}$$

$$= \eta_q + \frac{1}{m} \sum_{j=1}^{m} \left[\frac{1}{n} \sum_{i=(q-1)n}^{qn-1} e_{ij} \right] \quad .$$

In order to prove the theorem we wish to present, we would like to represent the Y_q as functions of means of identically distributed variates except for constant parameters. Let us define a new error term

$$e_{qj} = \frac{1}{n} \sum_{i=(q-1)n}^{qn-1} e_{ij} \quad , \quad j = 1, 2, \ldots, m \quad . \qquad (4.1.5)$$

Under the assumptions of our model, the errors e_{qj} are identically distributed variates with zero means and common variance σ^2/n . Now if we let

$$\bar{e}_q = \frac{1}{m} \sum_{j=1}^{m} e_{qj} \quad , \qquad (4.1.6)$$

we may write Y_q as

$$Y_q = \eta_q + \bar{e}_q \quad . \qquad (4.1.7)$$

The η_q , as can be seen from equation (4.1.4), are independent of m, and n is being held fixed in order to obtain limiting distributions as $m \longrightarrow \infty$.

Let us define the function $\hat{\theta}$ to be any one of the estimators $\hat{x}_k$, $\hat{\alpha}_k$, or $\hat{\lambda}_k$ obtained by the new estimation procedure for either Case 3 or Case 6. Since $\hat{\theta}$ is a function of the means Y_q which are in turn functions of the $\bar{e}_q$ and the η_q , $\hat{\theta}$ may also be considered as a function of the sample means $\bar{e}_q$ and the population means η_q . As indicated above, the $\bar{e}_q$ are means of identically distributed variates while the Y_q are not, for each Y_q is the mean of observations y_{ij} which under our model do not all have the same mean. Now Hsu [15] proves a theorem which gives the limiting distributions of functions of means such as the $\bar{e}_q$, and his theorem is applicable here. Let us define the point η to be a row vector with the η_q as elements and the point Y to be a row vector with the Y_q as elements. Now in terms of the estimation function $\hat{\theta}$, Hsu's theorem becomes

Theorem 1. If the function $\hat{\theta}(Y)$ of means Y_q possesses continuous second order derivatives of every kind in a neighborhood of the point η , then $\sqrt{mn}\left[\theta(Y) - \theta(\eta)\right]$ is normally distributed in the limit as $m \rightarrow \infty$ with mean zero and variance

$$\sum_q a_q^2 \sigma^2 \tag{4.1.8}$$

as long as $a_q \neq 0$ for some q , where

$$a_q = \frac{\partial}{\partial Y_q} \hat{\theta}(Y) \Big|_{Y=\eta} \quad . \tag{4.1.9}$$

4.2. Limiting Distributions for Large m

Since we have explicit formulas for the estimators for Cases 1, 2, 4, and 5, we could show that the conditions of Theorem 1 are satisfied for these estimators and then we could use Theorem 1 to obtain their limiting distributions as $m \longrightarrow \infty$. However, in this section we shall instead proceed directly to the limiting distributions as $m \longrightarrow \infty$ of the estimators for the more general Cases 3 and 6 even though the estimators for these cases cannot always be represented explicitly by algebraic equations. Although the detailed demonstration in this section will be for Case 3, it will be indicated that the results also hold for Case 6. After limiting distributions for the general Cases 3 and 6 are considered, particular results will be displayed for Cases 1 and 4.

Before considering the $\hat{x}_k$, the $\hat{\lambda}_k$, or the $\hat{\alpha}_k$, it is first necessary that we investigate the behavior of the $\hat{E}_r$ for Y in a neighborhood of η. From equation (3.3.23), it can be seen that each $\hat{E}_r$ may be represented as the ratio of two determinants in the S_q. If we substitute the terms mnY_q for the corresponding S_q in (3.3.23), the mn factors cancel out. Thus we may write $\hat{E}_r$ as

$$\hat{E}_r = \frac{|P_{p-r+1}|}{|P|} \quad , \quad r = 1, 2, \ldots, p \quad , \tag{4.2.1}$$

where $|P|$ and $|P_k|$ are the determinants $|R|$ and $|R_k|$ of equation (3.3.23) with the Y_q replacing the S_q.

Both $|P|$ and $|P_{p-r+1}|$ are continuous everywhere since each of them involves only sums of products of the Y_q. Hence $\hat{E}_r$ is continuous in a neighborhood of η provided that $|P| \neq 0$ at η. Furthermore, derivatives of all orders of $\hat{E}_r$ with respect to the Y_q are all continuous in a neighborhood of η if $|P| \neq 0$ at η, for the k^{th} order derivatives are ratios of continuous sums of products of the Y_q and $|P|^{2^k}$. And since $|P|$ is a continuous function of the Y_q , $|P| \neq 0$ in some neighborhood of η if it is not zero at η . Hence, to demonstrate the continuity of the $\hat{E}_r$ and their derivatives of all orders, or in other words, to demonstrate that the $\hat{E}_r$ are analytic, in a neighborhood of η , it is only necessary to show that

$$|P(\eta)| = |P|\Big|_{Y=\eta} \neq 0 \quad .$$

First observe that since $x_k = \Lambda_k^n$, (4.1.4) may be written as

$$\eta_q = \sum_{k=1}^{p} \alpha_k' x_k^{q-1} \quad ,$$

where

$$\alpha_k' = \frac{1}{n} \alpha_k \frac{(1 - x_k)}{1 - x_k^{1/n}} \quad .$$

Under the assumptions of our model, $\alpha_k' \neq 0$. Now reference to the definitions just given for $|P|$ and $|P(\eta)|$ and also to the definition of $|R|$ given in connection with (3.3.23) reveals that $|P(\eta)|$ has elements

$$\sum_{k=1}^{p} \alpha_k' x_k^{r+s-2} \quad .$$

Thus $|P(\eta)| = |B B^T|$, where B is a p by p matrix with elements $\sqrt{\alpha_s'}\, x_s^{r-1}$. But $|B|$ is in turn given by

$$\left| B \right| = (\alpha_1' \alpha_2' \cdots \alpha_p')^{\frac{1}{2}} \left| L \right| \quad .$$

Now, as was shown in Section 3.3,

$$\left| L \right| = \prod_{i > j} (x_i - x_j) \quad ,$$

and since all the x_k are distinct under the assumptions of our model, $\left| L \right| \neq 0$. So we conclude that

$$\left| P(\eta) \right| = \alpha_1' \alpha_2' \cdots \alpha_p' \left| L \right|^2 \neq 0 \quad . \tag{4.2.2}$$

Therefore, both the $\hat{E}_r$ and their derivatives of all orders with respect to the Y_q are continuous in a neighborhood of η .

As we stated earlier, the estimators $\hat{x}_k$, whose asympotic distributions we are seeking, are roots of a p^{th} degree polynomial equation with coefficients $\hat{E}_r$. Since the roots of such an equation are continuous functions of the coefficients [28, pg. 69], and since it has been shown that the $\hat{E}_r$ are continuous functions of the Y_q in a neighborhood of η , it follows that the $\hat{x}_k$ are continuous functions of the Y_q in such a neighborhood. However, Theorem 1 requires continuity of the second order partial derivatives of the $\hat{x}_k$ in a neighborhood of η , and to prove that these derivatives are continuous we shall refer to an implicit function theorem from the theory of functions of real variables.

This implicit function theorem is given, for instance, by Graves [11]. Applied to the polynomial (3.3.25), where the coefficients $\hat{E}_r$ are analytic functions of the Y_q in a neighborhood of η, it implies not only the existence and continuity of the second order derivatives of the $\hat{x}_k$, but of derivatives of all orders, in a neighborhood of η provided that two conditions are satisfied. If we let

$$g(x, Y) = x^p - \hat{E}_1 x^{p-1} + \hat{E}_2 x^{p-2} - \dots + (-1)^p \hat{E}_p \quad , \qquad (4.2.3)$$

$$g_x(x, Y) = \frac{\partial}{\partial x} g(x, Y) \quad , \qquad (4.2.4)$$

these conditions are

$$g(x, Y)\Big|_{\substack{x=x_k \\ Y=\eta}} = g(x_k , \eta) = 0 \ , \quad k = 1, 2, \dots, p \ ;$$

$$g_x(x, Y)\Big|_{\substack{x=x_k \\ Y=\eta}} = g_x(x_k , \eta) \neq 0 \quad , \quad k = 1, 2, \dots, p \ .$$

Since, as will be indicated later, $\hat{E}_r = E_r$, $r = 1, 2, \dots, p$, at η , and since the E_r are by definition the elementary symmetric functions of the x_k, the first condition always holds. However, the second condition is satisfied if and only if no two of the positive constants x_k are equal. But our model precludes any two x_k from being equal.

Therefore, under the model specified earlier for Case 3, the second order partial derivatives of the $\hat{x}_k$ with respect to the Y_q are continuous in some neighborhood of η .

Before applying Theorem 1 to the estimators for Case 3, we now need only observe that by virtue of the equating of S_q to $\mathcal{E}(S_q)$ in the estimation procedure, when $Y_q = \eta_q$ for all q , the estimators $\hat{x}_k = x_k$, $k = 1, 2, \ldots, p$. Thus, on the basis of Theorem 1, it follows that $\sqrt{mn}(\hat{x}_k - x_k)$ is asymptotically normally distributed with mean zero for large m , where $k = 1, 2, \ldots, p$. Now let us quickly show that both $\sqrt{mn}(\hat{\lambda}_k - \lambda_k)$ and $\sqrt{mn}(\hat{\alpha}_k - \alpha_k)$ are also normally distributed about a mean of zero in the limit as $m \to \infty$.

The estimators $\hat{\lambda}_k$ are related to the $\hat{x}_k$ by the equation

$$\hat{\lambda}_k = -\frac{1}{Kn} \ln \hat{x}_k \quad , \quad k = 1, 2, \ldots, p \quad , \tag{4.2.5}$$

which has second order partial derivatives

$$\frac{\partial^2 \hat{\lambda}_k}{\partial Y_j \, \partial Y_i} = \frac{\frac{\partial}{\partial Y_j}(\hat{x}_k)\frac{\partial}{\partial Y_i}(\hat{x}_k)}{Kn\, \hat{x}_k^2} - \frac{1}{Kn\, \hat{x}_k} \frac{\partial^2 \hat{x}_k}{\partial Y_j \, \partial Y_i} \quad , \tag{4.2.6}$$

$$k = 1, 2, \ldots, p \quad .$$

Since $\hat{x}_k = x_k$ at η and $x_k > 0$ for all k , there exists a neighborhood about η where the $\hat{x}_k > 0$ and have continuous second order partial derivatives. Therefore, from (4.2.5) and (4.2.6) it can be seen that both the $\hat{\lambda}_k$ and their second order partial derivatives with respect to the Y_q must also be continuous in that neighborhood. Also, upon substitution of x_k for $\hat{x}_k$ in (4.2.5), it follows that $\hat{\lambda}_k = \lambda_k$ at η. Hence, we may conclude that $\sqrt{mn}\,(\hat{\lambda}_k - \lambda_k)$ has a limiting normal distribution as $m \to \infty$ with mean zero by virtue of Theorem 1.

Turning our attention now to the estimator $\hat{\alpha}_k$, we first note that by the same argument used to show that $\hat{x}_k = x_k$ at η, it also follows that $\hat{\alpha}_k = \alpha_k$ at η. Thus, to demonstrate the asymptotic normality of $\sqrt{mn}\,(\hat{\alpha}_k - \alpha_k)$ for large m by Theorem 1, it is only necessary to show that the second order partial derivatives of $\hat{\alpha}_k$ with respect to the Y_q are continuous in a neighborhood of η. Substitution of $mn\,Y_q$ for S_q in equation (3.3.27) shows that the $\hat{\alpha}_k$ may be found by solving the set of equations

$$\frac{1}{n} \sum_{k=1}^{p} \hat{x}_k^{q-1} \frac{(1 - \hat{x}_k)}{1 - \hat{x}_k^{1/n}} \hat{\alpha}_k = Y_q \quad , \; q = 1, 2, \ldots, p \; . \tag{4.2.7}$$

Now $\hat{x}_k$ is continuous in a neighborhood of η and at η, $\hat{x}_k = x_k$, where $0 < x_k < 1$. Therefore, there exists a neighborhood of η for which $(1 - \hat{x}_k)/(1 - \hat{x}_k^{1/n})$ as well as $\hat{x}_k^{q-1}$ and Y_q are continuous. It follows from (4.2.7) that in that neighborhood the $\hat{\alpha}_k$ may be represented as ratios of continuous determinants. As in the case of the $\hat{E}_r$, the $\hat{\alpha}_k$ and their second order partial derivatives with respect to the Y_q will be continuous in a neighborhood of η if the determinant in the denominator is not zero at η.

The denominator in the ratio which equals $\hat{\alpha}_k$ may be shown to equal $(c_1, c_2, \ldots, c_p)\,|\,L\,|$, where

$$c_k = \frac{1 - x_k}{n(1 - x_k^{1/n})} \quad , \quad k = 1, 2, \ldots, p \ .$$

Since $x_k \neq 1$ for $\lambda_k > 0$ as specified in the model, each c_k is finite but not equal to zero. Also, we have already seen that under the restrictions of our model, $|L| \neq 0$. Hence each $\hat{\alpha}_k$ has a non-zero denominator at η , and both the $\hat{\alpha}_k$ and their second order partial derivatives with respect to the Y_q are continuous in some neighborhood of η . Therefore, Theorem 1 yields the result that $\sqrt{mn}(\hat{\alpha}_k - \alpha_k)$ has a limiting normal distribution with mean zero as $m \to \infty$.

So far the evaluation of the asymptotic variances of the estimators for Case 3 has not been mentioned. The estimators $\hat{\lambda}_k$ and $\hat{\alpha}_k$ are known functions of the exponential estimators $\hat{x}_k$. Therefore, their asymptotic variances, as well as those of the $\hat{x}_k$, follow directly from equations (4.1.8) and (4.1.9) once the first partial derivatives of the $\hat{x}_k$ with respect to each of the Y_q are known. The evaluation of these partial derivatives is also given by Graves [11] . Let us define column vectors

$$g_Y(x, Y) = \left[\frac{\partial}{\partial Y_1} g(x, Y), \frac{\partial}{\partial Y_2} g(x, Y), \ldots, \frac{\partial}{\partial Y_{2p}} g(x, Y) \right]^T ; \tag{4.2.8}$$

$$X_{kY} = \left[\frac{\partial}{\partial Y_1} \hat{x}_k , \frac{\partial}{\partial Y_2} \hat{x}_k , \ldots, \frac{\partial}{\partial Y_{2p}} \hat{x}_k \right]^T , \quad k = 1, 2, \ldots, p \ , \tag{4.2.9}$$

where $g(x, Y)$ is defined by (4.2.3). Also, let

$$g_Y(x_k, Y) = g_Y(x, Y)\Big|_{x=\hat{x}_k} \quad , \quad k = 1, 2, \ldots, p \quad , \qquad (4.2.10)$$

$$g_x(x_k, Y) = g_x(x, Y)\Big|_{x=\hat{x}_k} \quad , \quad k = 1, 2, \ldots, p \quad , \qquad (4.2.11)$$

where the scalar $g_x(x, Y)$ is defined by (4.2.4). Now the vector X_{kY} of the first partial derivatives of any $\hat{x}_k$ with respect to the Y_q satisfies the equation

$$X_{kY} = - \frac{g_Y(\hat{x}_k, Y)}{g_x(\hat{x}_k, Y)} \quad , \quad k = 1, 2, \ldots, p \quad . \qquad (4.2.12)$$

With the help of (4.2.12), we may evaluate the a_q derivatives of equation (4.1.9) in Theorem 1 for any $\hat{\lambda}_k$. For

$$\frac{\partial \hat{\lambda}_k}{\partial Y_q} = - \frac{\frac{\partial}{\partial Y_q} \hat{x}_k}{Kn \; \hat{x}_k} \quad , \quad q = 1, 2, \ldots, 2p \quad . \qquad (4.2.13)$$

and $\frac{\partial}{\partial Y_q} \hat{x}_k$ for any k and q is given by (4.2.12). Similarly, equation (4.2.7) yields a solution for any $\hat{\alpha}_k$ in terms of the $\hat{x}_k$, and the first partial derivatives of the $\hat{\alpha}_k$ with respect to the Y_q may be written in terms of the $\hat{x}_k$ and their first partial derivatives. Hence, (4.1.9) may also be evaluated for any $\hat{\alpha}_k$ with the help of (4.2.12).

Then once the a_q derivatives are determined for either an $\hat{\alpha}_k$ or a $\hat{\lambda}_k$ estimator, (4.1.8) may be employed to find the asymptotic variance of that estimator for large m .

The limiting distributions just derived for Case 3 may be shown in the same manner to hold for Case 6. We have already seen that Theorem 1 is applicable to Case 6 as well as to Case 3. Therefore, in order to claim the results of this section for Case 6, we need only show that the continuity conditions of Theorem 1 are satisfied for Case 6. And to do this, we need to demonstrate that the determinant $|P(\eta)|$, evaluated by (4.2.2) for Case 3, and the denominators of the $\hat{\alpha}_k$ at $Y = \eta$ are not zero for Case 6. However, from (3.6.5), (3.6.10), (4.1.2) and the discussion leading up to (4.2.2), it can be seen that for Case 6,

$$|P(\eta)| = \prod_{k=1}^{p} \alpha_k' (1 - x_k) \; |L|^2 \quad .$$

Since $x_k \neq 1$ under our model and since the right side of (4.2.2) is not equal to zero under our model, $|P(\eta)| \neq 0$ for Case 6. Similarly, the denominator of $\hat{\alpha}_k$, $k = 1, 2, \ldots, p$, evaluated at η for Case 6 equals $\prod_{k=1}^{p} (1 - x_k)$ times the corresponding denominator for Case 3, and is therefore not equal to zero at η . Finally, from (3.6.11) it can also be seen that $\hat{\alpha}_0$ is continuous at $Y = \eta$. Thus, the limiting distributions as $m \to \infty$ already obtained for Case 3 also apply to Case 6.

In summary of this section thus far, it has been proved that the estimators $\hat{x}_k$, $\hat{\lambda}_k$ and $\hat{\alpha}_k$ obtained by the new estimation procedure for parameters in the general Case 3 and Case 6 models are such that if $\hat{\theta}$ denotes any one of the estimators and θ the corresponding parameter, $\sqrt{mn}(\hat{\theta} - \theta)$ is asymptotically normally distributed with mean zero for n fixed and m large. Also, a method has been given for determining the asymptotic variance of $\sqrt{mn}(\hat{\theta} - \theta)$ by using (4.2.12), (4.1.8) and (4.1.9) in conjunction with the equation which specifies $\hat{\theta}$ in terms of the $\hat{x}_k$. This method is also applicable to Case 6 as well as Case 3. When using it for Case 6, the $\partial\hat{\lambda}_k/\partial Y_q$ are still given by (4.2.13), but the $\partial\hat{\alpha}_0/\partial Y_q$ are obtained from (3.6.11) and the $\partial\hat{\alpha}_k/\partial Y_q$, $k = 1, 2, \ldots, p$, from (3.6.10), where $Y_q = S_q/mn$. Also, the $\partial\hat{x}_k/\partial Y_q$ are still given by (4.2.12), but the vector Y now has $(2p + 1)$ elements instead of $2p$ elements as it had for Case 3.

Before going on to limiting distributions as $n \longrightarrow \infty$, let us look at the limiting distributions as $m \longrightarrow \infty$ for Cases 1 and 4 in particular. Interpreted in terms of Case 1, the conclusions of the last paragraph are that $\sqrt{mn}(\hat{x} - x)$, $\sqrt{mn}(\hat{\lambda} - \lambda)$ and $\sqrt{mn}(\hat{\alpha} - \alpha)$ are all asymptotically normally distributed with zero means for m large. The asymptotic variances for $\sqrt{mn}(\hat{x} - x)$ and $\sqrt{mn}(\hat{\lambda} - \lambda)$ may be determined by differentiating both (3.5.3) and (3.5.4) with respect to Y_1 and Y_2 and then substituting in (4.1.9) and (4.1.8). In this way we find that the asymptotic variances for $\sqrt{mn}(\hat{x} - x)$ and $\sqrt{mn}(\hat{\lambda} - \lambda)$ are

$$\frac{(\eta_1^2 + \eta_2^2)\,\sigma^2}{\eta_1^4} \tag{4.2.14}$$

and

$$\frac{(\eta_1^2 + \eta_2^2)\,\sigma^2}{K^2\, n^2\, \eta_1^2\, \eta_2^2} \tag{4.2.15}$$

respectively. Similarly, (3.5.5) in conjunction with (4.1.8) and (4.1.9) leads to $(a_1^2 + a_2^2)\sigma^2$ for the asymptotic variance of $\sqrt{mn}\,(\hat{\alpha} - \alpha)$, where

$$a_1 = \frac{\eta_1 \left[n(\eta_1 - 2\eta_2)(\eta_1^{1/n} - \eta_2^{1/n}) + (\eta_1 - \eta_2)\,\eta_2^{1/n} \right]}{(\eta_1 - \eta_2)^2\, \eta_1^{1/n}} ;$$

(4.2.16)

$$a_2 = \frac{\eta_1^2 \left[n\,\eta_2\,(\eta_1^{1/n} - \eta_2^{1/n}) - (\eta_1 - \eta_2)\,\eta_2^{1/n} \right]}{(\eta_1 - \eta_2)^2\, \eta_1^{1/n}\, \eta_2} .$$

For Case 4, $\sqrt{mn}\,(\hat{x} - x)$, $\sqrt{mn}\,(\hat{\lambda} - \lambda)$, $\sqrt{mn}\,(\hat{\alpha}_0 - \alpha_0)$ and $\sqrt{mn}\,(\hat{\alpha}_1 - \alpha_1)$ all have limiting normal distributions as $m \to \infty$ with zero means and variances given by $(a_1^2 + a_2^2 + a_3^2)\,\sigma^2$. Differentiation of (3.6.12), (3.6.13) and (3.6.14) and substitution in (4.1.9) shows that for $\sqrt{mn}\,(\hat{x} - x)$,

$$a_1 = -\frac{\eta_2 - \eta_3}{(\eta_1 - \eta_2)^2} , \quad a_2 = \frac{\eta_1 - \eta_3}{(\eta_1 - \eta_2)^2} , \quad a_3 = -\frac{1}{\eta_1 - \eta_2} ;$$

(4.2.17)

for $\sqrt{mn}(\hat{\lambda} - \lambda)$,

$$a_1 = \frac{1}{Kn(\eta_1 - \eta_2)} \ , \quad a_2 = -\frac{\eta_1 - \eta_3}{Kn(\eta_1 - \eta_2)(\eta_2 - \eta_3)} \ , \quad a_3 = \frac{1}{Kn(\eta_2 - \eta_3)} \ ; \tag{4.2.18}$$

and for $\sqrt{mn}(\hat{\alpha}_0 - \alpha_0)$,

$$a_1 = \frac{(\eta_2 - \eta_3)^2}{(\eta_1 - 2\eta_2 + \eta_3)^2} \ ,$$

$$a_2 = -\frac{2(\eta_1 - \eta_2)(\eta_2 - \eta_3)}{(\eta_1 - 2\eta_2 + \eta_3)^2} \ , \tag{4.2.19}$$

$$a_3 = \frac{(\eta_1 - \eta_2)^2}{(\eta_1 - 2\eta_2 + \eta_3)^2} \ .$$

Finally, in terms of the constant $x = (\eta_2 - \eta_3)/(\eta_1 - \eta_2)$, the a_q for $\sqrt{mn}(\hat{\alpha}_1 - \alpha_1)$ are

$$a_1 = \frac{1}{m(1 - x)^3}\left[(1 - x)\, x^{\frac{1}{n}} + n(1 - 3x)(1 - x^{\frac{1}{n}})\right] \ ,$$

$$a_2 = \frac{1}{mx(1 - x)^3}\left[nx\,(1 + 3x)(1 - x^{\frac{1}{n}}) - (1 - x)(1 + x)\, x^{\frac{1}{n}}\right] \ , \tag{4.2.20}$$

$$a_3 = \frac{1}{mx(1-x)^3}\left[(1-x)\,x^{\frac{1}{n}} - 2nx\,(1-x^{\frac{1}{n}})\right] .$$

4.3 Limiting Distributions for Large n

Up to this point limiting distributions have been derived with n constant as $m \longrightarrow \infty$, where n is the number of points at which observations are taken within each interval and m is the number of observations per point. Now let us reverse the situation and hold m constant, so that asymptotic distributions for large n may be found.

From Section 4.1 we have

$$Y_q = \eta_q + \frac{1}{mn} \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} e_{ij} . \qquad (4.3.1)$$

The error term

$$\bar{e}'_q = \frac{1}{n} \sum_{i=(q-1)n}^{qn-1} \left[\frac{1}{m} \sum_{j=1}^{m} e_{ij} \right] \qquad (4.3.2)$$

may be thought of as the mean of the n identically distributed variates

$$\frac{1}{m} \sum_{j=1}^{m} e_{ij} .$$

Also, the term

$$\eta_q = \frac{1}{n} \sum_{k=1}^{p} \alpha_k \Lambda_k^{(q-1)n} \frac{(1-\Lambda_k^n)}{1-\Lambda_k}$$

and is not independent of n . In fact, in Section 5.2 it will be shown that for constant m and constant T $(=Kn)$,

$$\lim_{n \to \infty} \eta_q = \varphi_q \quad , \tag{4.3.3}$$

where the constant φ_q is defined by

$$\varphi_q = \frac{1}{T} \sum_{k=1}^{p} \frac{\alpha_k}{\lambda_k} e^{-\lambda_k (q-1)T} \left(1 - e^{-\lambda_k T}\right) . \tag{4.3.4}$$

It is now apparent that unlike the situation when m was allowed to grow large, when n increases, Y_q may not be represented as in (4.1.7) by a constant plus a mean of identically distributed variates. So now Hsu's theorem may not be applied directly, as it was in Section 4.1, to obtain the desired limiting distributions as $n \to \infty$. However, a modification of Hsu's theorem, used in conjunction with a theorem presented by Cramér, is applicable to the present situation.

From equations (4.3.1) and (4.3.2), we have that

$$Y_q = \eta_q + \bar{e}_q' \quad ,$$

and therefore that

$$\sqrt{n}\,(Y_q - \varphi_q) = \sqrt{n}\,(\eta_q - \varphi_q) + \sqrt{n}\,\bar{e}_q' \quad , \tag{4.3.5}$$

where φ_q is defined by (4.3.4). The Central Limit Theorem [2, pp. 213-218] shows that the error term $\sqrt{n}\,\bar{e}_q'$ in equation (4.3.5) is asymptotically normally distributed for large n about a mean of zero with variance σ^2/m . Moreover, it may also be demonstrated that

$$\lim_{n\to\infty} \sqrt{n}\,(\eta_q - \varphi_q) = 0 \quad , \tag{4.3.6}$$

for reference to (4.1.4) and (4.3.4) shows that

$$\lim_{n\to\infty} \sqrt{n}\,(\eta_q - \varphi_q) = \sum_{k=1}^{p} B_k \lim_{n\to\infty} \left[\frac{T\,\lambda_k - n(1 - e^{-\lambda_k T/n})}{\sqrt{n}\;T\,\lambda_k(1 - e^{-\lambda_k T/n})}\right] , \tag{4.3.7}$$

where the constant

$$B_k = \alpha_k\, e^{-\lambda_k T(q-1)} \;(1 - e^{-\lambda_k T}) \quad .$$

Then several applications of L'Hospital's rule yield the result that

$$\lim_{n\to\infty} \left[\frac{T\,\lambda_k - n(1 - e^{-\lambda_k T/n})}{\sqrt{n}\;T\,\lambda_k(1 - e^{-\lambda_k T/n})}\right] = \lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0 \quad ,$$

thus reducing (4.3.7) to (4.3.6). Now, applying a theorem given by Cramér [2, pg. 254] to the equation (4.3.5), we conclude that $\sqrt{n}\,(Y_q - \varphi_q)$ has the same limiting distribution as $\sqrt{n}\;\bar{e}_q'$. That is, for n large $\sqrt{n}\,(Y_q - \varphi_q)$ is asymptotically normally distributed with mean zero and variance σ^2/m .

Now let us refer again to the theorem of Hsu's [15] used in Section 4.1. In deriving the limiting distributions of functions of sample means, Hsu utilizes only one property of normalized means such as the $\sqrt{n}\,\bar{e}'_q$: namely, their limiting distributions. Therefore, since $\sqrt{n}\,(Y_q - \varphi_q)$ has the same limiting distribution as $\sqrt{n}\,\bar{e}'_q$, Hsu's theorem may just as well be proved in terms of the $(Y_q - \varphi_q)$ instead of the means $\bar{e}'_q$. Such a proof in the context of this paper would lead to

<u>Theorem 2</u>. If the function $\hat{\theta}(Y)$ of means Y_q possesses continuous second order derivatives of every kind in a neighborhood of the point φ , then $\sqrt{mn}\left[\hat{\theta}(Y) - \hat{\theta}(\varphi)\right]$ is normally distributed in the limit as $n \rightarrow \infty$ with mean zero and variance

$$\sum_q b_q^2\, \sigma^2 \tag{4.3.8}$$

as long as $b_q \neq 0$ for some q , where

$$b_q = \frac{\partial}{\partial Y_q}\hat{\theta}(Y)\Bigg|_{Y=\varphi} \quad . \tag{4.3.9}$$

In Theorem 2, φ is the point with the φ_q as coordinates, while Y and $\hat{\theta}$ are the same as defined in Section 4.1. Theorem 2, like Theorem 1, holds for both Case 3 and Case 6. Moreover, the derivatives $\frac{\partial}{\partial Y_q}\hat{\theta}(Y)$, as stated in the last section, may be evaluated with the help of (4.2.12).

In Section 4.2 we expressed the estimators obtained by the new procedure in terms of the sample means Y_q , and then we went on to show that the estimators themselves as well as their second order partial derivatives with respect to the Y_q are continuous in a neighborhood of η . A study of Section 4.2 reveals that to demonstrate similar continuity conditions for a neighborhood of φ for both Case 3 and Case 6, we need only show that the following three properties hold for Case 3 under the assumptions of our model:

(1) $\varphi_r \neq \varphi_s \ , \quad r \neq s \quad ;$

(2) $\left|P(\varphi)\right| = \left|P\right|\Big|_{Y=\varphi} \neq 0 \quad ;$

(3) $x_k \Big|_{Y=\varphi} = x_k \quad .$

Since our model specifies that $\alpha_k \neq 0$ and $\lambda_k > 0$, $k = 1, 2, \ldots, p$, and that $\lambda_r \neq \lambda_s$, $r \neq s$, we deduce from equation (4.3.4) that $\varphi_r \neq \varphi_s$ for $r \neq s$. Furthermore, when $Y = \varphi$,

$$\varphi_q = \sum_{k=1}^{p} u_k \, x_k^{q-1} \quad ,$$

where

$$u_k = \frac{\alpha_k}{\lambda_k T} \left(1 - e^{-\lambda_k T} \right) \quad , \quad k = 1, 2, \ldots, p \quad .$$

Note that $u_k \neq 0$ in accordance with our model for every k . Now at φ , $\left|P\right|$, defined in Section 4.2 in connection with equation (4.2.1), has elements

$$\sum_{k=1}^{p} u_k \, x_k^{r+s-2} \quad .$$

Hence, by comparing these elements with those given for $|P(\eta)|$ in Section 4.2, we conclude by analogy with (4.2.2) that

$$|P(\varphi)| = u_1 u_2 \ldots u_p \; |L|^2 \quad . \qquad (4.3.10)$$

Since it has been shown that $|L| \neq 0$ under the assumptions of our model, it follows that $|P(\varphi)| \neq 0$.

We have shown that the first and second properties necessary to prove continuity of the estimators and their second order derivatives in a neighborhood of φ are satisfied. In order to show that the last one holds, namely that at φ , $\hat{x}_k = x_k$, $k = 1, 2, \ldots, p$, we recall from (4.3.3) that $\lim_{n\to\infty} \eta_q = \varphi_q$. Consequently, since $\hat{x}_k$ is continuous in a neighborhood of η , $\lim_{n\to\infty} \hat{x}(\eta) = \hat{x}(\varphi)$. But $\hat{x}(\eta) = x$ for all n , and therefore $\hat{x}(\varphi) = x$. Hence, the estimators $\hat{x}_k$, $\hat{\lambda}_k$ and $\hat{\alpha}_k$ and their second order partial derivatives are continuous in a neighborhood of φ .

Now we wish to complete the demonstration that $\sqrt{mn}\,(\hat{\theta} - \theta)$ is normally distributed in the limit as $n \to \infty$ with zero mean and variance given by (4.3.8), where $\hat{\theta}$ may denote any of the estimators $\hat{x}_k$, $\hat{\lambda}_k$ or $\hat{\alpha}_k$ and θ denotes the corresponding population parameter. To do this it is necessary to show that $\lim_{n\to\infty} \hat{\alpha}_k = \alpha_k$ at φ . Then Theorem 2 may be applied to give all the desired limiting distributions, since it has already been shown that $\hat{x}_k = x_k$ at φ , and consequently that $\hat{\lambda}_k = \lambda_k$ at that point.

Substituting the φ_q for the corresponding Y_q in (3.3.27), we can see that the $\hat{\alpha}_k$ may be determined at φ by solving the following linear equations for the $\hat{\alpha}_k$:

$$\frac{1}{n} \sum_{k=1}^{p} \hat{x}_k^{q-1} \frac{(1 - \hat{x}_k)}{1 - \hat{x}_k^{1/n}} \hat{\alpha}_k = \frac{1}{T} \sum_{k=1}^{p} \frac{\alpha_k}{\lambda_k} x_k^{q-1} (1 - x_k) \quad , \qquad (4.3.11)$$

$$q = 1, 2, \ldots, p \quad .$$

But $\hat{x}_k = x_k$ at φ , so substitution of x_k for $\hat{x}_k$ in (4.3.11) and multiplication of both sides of that equation by $\lambda_k T$ yields

$$\sum_{k=1}^{p} x_k^{q-1} (1 - x_k) \left[\frac{T\lambda_k \hat{\alpha}_k}{n(1 - x_k^{1/n})} - \alpha_k \right] = 0 \quad , \qquad (4.3.12)$$

$$q = 1, 2, \ldots, p \quad .$$

Keeping in mind that we are interested in limiting distributions for the $\hat{\alpha}_k$ as $n \rightarrow \infty$, let us evaluate

$$\lim_{n \rightarrow \infty} \frac{T\lambda_k}{n(1 - x_k^{1/n})} = \lim_{n \rightarrow \infty} \frac{\lambda_k T/n}{1 - e^{-\lambda_k T/n}} \qquad (4.3.13)$$

where $\lambda_k T$ is constant. This limit is equal to

$$\lim_{t \rightarrow 0} \frac{t}{1 - e^{-t}}$$

which by L'Hospital's rule in turn equals $\lim_{t \rightarrow 0} e^t = 1$. Thus, as $n \rightarrow \infty$, (4.3.12) becomes

$$\sum_{k=1}^{p} x_k^{q-1}(1 - x_k)(\hat{\alpha}_k - \alpha_k) = 0 \;, \quad q = 1, 2, \ldots, p \;, \tag{4.3.14}$$

which has as its solution, $\hat{\alpha}_k = \alpha_k$. Hence, in the limit as $n \rightarrow \infty$, $\hat{\alpha}_k = \alpha_k$ at φ .

Now, on the basis of the results obtained in this section and the proofs given in Section 4.2, we conclude that as $n \rightarrow \infty$, the distributions of $\sqrt{mn}(\hat{x}_k - x_k)$, $\sqrt{mn}(\hat{\lambda}_k - \lambda_k)$ and $\sqrt{mn}(\hat{\alpha}_k - \alpha_k)$ are asymptotically normal with zero means and variances calculated from (4.3.8) in Theorem 2 with the help of (4.3.9) and (4.2.12). Explicit formulas for the constants b_q given by (4.3.9) may be determined for Cases 1 and 4 by substituting φ_q for η_q in the formulas for the corresponding a_q given in Section 4.2.

4.4 An Additional Limiting Distribution for Large n

In this section we shall derive another limiting distribution as $n \rightarrow \infty$ which will be utilized in Chapter VI and which will help summarize the results of this chapter. Let us consider the limiting distribution as $n \rightarrow \infty$ of

$$\frac{\sqrt{mn}\left[\hat{\theta}(Y) - \hat{\theta}(\eta)\right]}{\sqrt{\sum_q a_q \sigma^2}} = \left\{\frac{\sqrt{mn}\left[\hat{\theta}(Y) - \hat{\theta}(\varphi)\right] - \sqrt{mn}\left[\hat{\theta}(\varphi) - \hat{\theta}(\eta)\right]}{\sqrt{\sum_q b_q^2 \sigma^2}}\right\} \sqrt{\frac{\sum_q b_q^2 \sigma^2}{\sum_q a_q^2 \sigma^2}} \;. \tag{4.4.1}$$

As before, $q = 1, 2, \ldots, p$ for Case 3 and $q = 1, 2, \ldots, 2p+1$ for Case 6. From (4.1.9), (4.3.3) and (4.3.9) we deduce that

$$\lim_{n \to \infty} \sqrt{\frac{\sum_q b_q \sigma^2}{\sum_q a_q^2 \sigma^2}} = 1 . \tag{4.4.2}$$

Moreover, since it has been shown in Sections 4.2 and 4.3 that at either η or φ, $\hat{x}_k = x_k$ and $\hat{\lambda}_k = \lambda_k$, $k = 1, 2, \ldots, p$,

$$\lim_{n \to \infty} \sqrt{mn} \left[\hat{\theta}(\varphi) - \hat{\theta}(\eta) \right] = 0 \tag{4.4.3}$$

for $\hat{\theta} = \hat{x}_k$ or $\hat{\lambda}_k$.

In order to complete this proof, we need to demonstrate that (4.4.3) also holds for $\hat{\theta} = \hat{\alpha}_k$, $k = 1, 2, \ldots, p$. Now if the set of equations (3.3.27) is solved for any given $\hat{\alpha}_k$ at both η and φ, the two solutions will be ratios of determinants with identical denominators. In our consideration of the continuity of the $\hat{\alpha}_k$ at η and φ we have already seen that this denominator is not zero at either η or φ. By virtue of (4.3.13), it can also be seen that this denominator approaches a constant as $n \to \infty$. The determinants in the numerators of the two solutions will also be the same except for the k^{th} columns. In the solution for $\hat{\alpha}_k$ at η, this k^{th} column will be the column vector η^T, while in the solution at φ it will be φ^T. Now, like the denominators, the cofactor of the elements in the k^{th} columns will be identical for the two solutions and will approach constant limits as $n \longrightarrow \infty$. Therefore, expanding the numerator determinants about their k^{th} columns, we find that

$$\hat{\alpha}_k(\varphi) - \hat{\alpha}_k(\eta) = \sum_{q=1}^{p} (\varphi_q - \eta_q)\, U_q \ , \quad k = 1, 2, \ldots, p \ ,$$

where $\lim_{n \to \infty} U_q$ is a finite constant for every q . Now when m is held constant,

$$\lim_{n \to \infty} \sqrt{mn}\left[\hat{\alpha}_k(\varphi) - \hat{\alpha}_k(\eta)\right] = \sqrt{m} \sum_{q=1}^{p} \left[\lim_{n \to \infty} \sqrt{n}\,(\varphi_q - \eta_q)\right]\left[\lim_{n \to \infty} U_q\right],$$

$$k = 1, 2, \ldots, p \ .$$

But from (4.3.6), $\lim_{n \to \infty} \sqrt{n}\,(\varphi_q - \eta_q) = 0$ for all q . Therefore, (4.4.3) is also satisfied when $\hat{\theta} = \hat{\alpha}_k$, $k = 1, 2, \ldots, p$.

Now from (4.4.1), (4.4.2), (4.4.3) and a theorem given by Cramér [2, pg. 254] , we deduce that the left side of (4.4.1) and

$$\frac{\sqrt{mn}\left[\hat{\theta}(Y) - \hat{\theta}(\varphi)\right]}{\sqrt{\sum_q b_q \sigma^2}} \tag{4.4.4}$$

have the same limiting distribution as $n \longrightarrow \infty$. But from Theorem 2 it follows that (4.4.4) has an asymptotic standard normal distribution from n large. Therefore, since $\hat{\theta}(\eta) = \theta$, we have that

$$\frac{\sqrt{mn}\,[\hat{\theta} - \theta]}{\sqrt{\sum_q a_q \sigma^2}} \tag{4.4.5}$$

has a limiting standard normal distribution as $n \longrightarrow \infty$. But (4.4.5) is also a standard normal variate in the limit as $m \longrightarrow \infty$, a result that follows from Theorem 1. Thus, the results of this chapter may be summarized by saying that the distribution of (4.4.5) approaches the standard normal distribution as either m or $n \rightarrow \infty$ and by noting that $a_q \longrightarrow b_q$ as $n \longrightarrow \infty$, where a_q is defined by (4.1.9) and b_q by (4.3.9).

V. THE PROPERTIES OF THE ESTIMATORS

5.1 Sufficiency

This chapter will be concerned with the statistical properties of the estimators derived with the new estimation procedure. In this section we shall consider whether or not the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$ are sufficient. To do this it is necessary to examine the joint density function of the observations y_{ij} . However, our model does not specify the distribution of the y_{ij} but only requires that the corresponding errors, e_{ij} , be identically distributed with mean zero and common variance σ^2 . Hence, in order to study the sufficiency of our estimators, we shall first make the additional assumption that the errors e_{ij} , and consequently the observations y_{ij} , are normally distributed.

Now, for Case 3, each y_{ij} has the density function

$$f(y_{ij}\, ;\, \alpha_k, \lambda_k, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp. \left[- \frac{1}{2\sigma^2} (y_{ij} - \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i})^2 \right] . \tag{5.1.1}$$

Thus the joint density function of the y_{ij} is

$$\prod_{i=0}^{2pn-1} \prod_{j=1}^{m} f(y_{ij}\, ;\, \alpha_k, \lambda_k, \sigma^2) = \frac{1}{(2\pi\, \sigma^2)^{pmn}} \exp. \left[- \frac{1}{2\sigma^2} \sum_{i=0}^{2pn-1} \sum_{j=1}^{m} (y_{ij} - \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i})^2 \right] . \tag{5.1.2}$$

If the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$ obtained by the new procedure are sufficient, then the density function (5.1.2) must necessarily be factorable into two functions, one of which involves only the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$ and the parameters of (5.1.2) while the other is independent of the parameters α_k and λ_k (see [2, pp. 488-489]) . Thus, after expanding the exponent of (5.1.2), we can see that the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$ are sufficient only if the sum

$$\sum_{i=0}^{2pn-1} \sum_{j=1}^{m} (y_{ij} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i}) = \sum_{i=0}^{2pn-1} \left[(\sum_{j=1}^{m} y_{ij}) \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i} \right] \tag{5.1.3}$$

can be expressed without explicitly involving a product of the observations y_{ij} and the parameters α_k and λ_k . It can be shown that this is possible only if $n = 1$. Then

$$\sum_{j=1}^{m} y_{ij} = S_i = m \sum_{k=1}^{p} \hat{\alpha}_k e^{-\hat{\lambda}_k Kn(i-1)} \frac{(1 - e^{-\hat{\lambda}_k Kn})}{1 - e^{-\hat{\lambda}_k K}} , \tag{5.1.4}$$

and therefore in (5.1.3) the sum $\sum_{j=1}^{m} y_{ij}$ may be replaced by a function of the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$. That is, when the errors e_{ij} in our model are normally distributed, the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$ are sufficient only if the number of points t_i at which observations are taken is equal to the number of parameters in the model. This result may be shown to hold for Case 6 as well as Case 3.

The estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$ of the new estimation procedure have been found to be sufficient in only one instance other than the one already mentioned under the assumption of normality. This instance occurs when both m and $n = 1$, as in the method originally presented by Prony. In this situation, as can be seen from equation (5.1.4), each y_{ij} is itself a function of the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$, and therefore the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$ are sufficient regardless of the distribution of the errors e_{ij}.

5.2 Consistency

In order to prove consistency for the estimators derived with the new estimation procedure as $m \to \infty$ while n is held fixed, we need only utilize results already obtained. Means Y_q have been defined as have their expectations η_q, and these means converge in probability as $m \to \infty$ to the corresponding η_q. Furthermore, continuity of each of the estimators $\hat{x}$, $\hat{\alpha}_k$ and $\hat{\lambda}_k$ in a neighborhood of η was demonstrated in Section 4.2. It was also shown in that section that at η, $\hat{x}_k = x_k$, $\hat{\alpha}_k = \alpha_k$ and $\hat{\lambda}_k = \lambda_k$. Hence, on the basis of Slutsky's theorem [27], we conclude that the estimators $\hat{x}$, $\hat{\alpha}_k$ and $\hat{\lambda}_k$ converge in probability to x, α_k and λ_k respectively as $m \to \infty$. But an estimator $\hat{\theta}$ is a consistent estimate of θ if it converges to θ in probability (see [18, pg. 3]). Therefore, $\hat{x}$, $\hat{\alpha}_k$ and $\hat{\lambda}_k$ are consistent estimators of x, α_k and λ_k respectively for large m.

In order to use Slutsky's theorem to prove that the estimators are consistent as $n \to \infty$ with m fixed, we need to show that each Y_q converges in probability to some constant as $n \to \infty$. Such a probability limit can be found even though, as shown in Section 4.3, the

Y_q cannot be regarded as means of identically distributed variates.

From equations (3.2.1) and (3.1.5),

$$S_q = m \sum_{i=(q-1)n}^{qn-1} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k Ki} + \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} e_{ij} .$$

Although we are letting n grow large, we wish to keep the domain for the t_i constant in length, where $t_i = Ki$. That is, when the number of points at which observations are made is increased, the intervals between points are shortened so that the length of the interval for which the y_{ij} sum to S_q is constant. This constant, as defined earlier, is $T(=Kn)$. Now we may rewrite S_q as a sequence in n without involving the variable K as follows:

$$S_q = m \sum_{i=(q-1)n}^{qn-1} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k \frac{Ti}{n}} + \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} e_{ij} .$$

Thus, we may express $Y_q = S_q/mn$ as

$$Y_q = \frac{1}{T} \sum_{i=(q-1)n}^{qn-1} \alpha_k e^{-\lambda_k \frac{Ti}{n}} \frac{T}{n} + \frac{1}{mn} \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} e_{ij} . \tag{5.2.1}$$

Now from the definition of a definite integral [8] , it can be seen that

$$\lim_{n\to\infty} \frac{1}{T} \sum_{i=(q-1)n}^{qn-1} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k \frac{Ti}{n}} \frac{T}{n} = \frac{1}{T} \int_{(q-1)T}^{qT} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t}\, dt$$

$$= \frac{1}{T} \sum_{k=1}^{p} \frac{\alpha_k}{\lambda_k} e^{-\lambda_k (q-1)T} (1 - e^{-\lambda_k T}) \quad . \tag{5.2.2}$$

Also, since in the model the e_{ij} are independent, identically distributed variates with mean zero and variance σ^2, the mean

$$\frac{1}{mn} \sum_{i=(q-1)n}^{qn-1} \sum_{j=1}^{m} e_{ij}$$

converges in probability to zero as $n \to \infty$. Therefore, taking the probability limit of both sides of equation (5.2.1) as $n \to \infty$, we can see that Y_q converges in probability to the constant given in equation (5.2.2). Thus, referring to the definition of φ_q given by formula (4.3.4), we have that Y_q converges in probability to φ_q as $n \to \infty$.

In Section 4.3 it was proved that at the point φ , $\hat{x}_k$ and $\hat{\lambda}_k$ are continuous and equal to x_k and λ_k respectively. Thus Slutsky's theorem [27] leads to the result that $\hat{x}_k$ and $\hat{\lambda}_k$ are consistent estimators of x_k and λ_k respectively when m is held fixed as $n \to \infty$. It was also demonstrated in Section 4.3 that $\hat{\alpha}_k$ is continuous at φ , but instead of showing that $\hat{\alpha}_k = \alpha_k$ at φ , it was only shown that at φ , $\lim_{n\to\infty} \hat{\alpha}_k = \alpha_k$. In other words, the estimator $\hat{\alpha}$ is a continuous function of n for this case, for the solution of equation (4.3.11) involved a factor

$$\frac{T \lambda_k}{n(1 - x_k^{1/n})} \quad . \tag{5.2.3}$$

It was shown in Section 4.3 that the expression in (5.2.3) $\longrightarrow 1$ as $n \longrightarrow \infty$. Therefore, if we apply Slutsky's theorem to equation (4.3.11) apart from the factor (5.2.3), and then if we utilize the theorem that the limit of a quotient is equal to the quotient of the corresponding limits, where the limit in the denominator is not zero, we still obtain the desired result. Namely, as $n \longrightarrow \infty$, $\hat{\alpha}_k$ converges in probability to α_k, and hence $\hat{\alpha}_k$ is a consistent estimator of α_k for large n.

It is interesting to note that φ_q is functionally independent of m. However, as mentioned in Section 4.1, η_q is a function of n. In fact, since

$$\lim_{n \to \infty} \frac{T \lambda_k}{n(1 - x_k^{1/n})} = 1 \quad ,$$

from (4.1.4) we deduce that

$$\lim_{n \to \infty} \eta_q = \frac{1}{T} \sum_{k=1}^{p} \frac{\alpha_k}{\lambda_k} (1 - x_k) x_k^{q-1} \quad .$$

Therefore, equation (4.3.4) reveals that

$$\lim_{n \to \infty} \eta_q = \varphi_q \quad .$$

Hence, as both m and $n \rightarrow \infty$, the Y_q converge in probability to the corresponding constants φ_q. It is immaterial whether the limiting process is carried out with respect to m or with respect to n first. The conclusion then is that the same consistency properties which hold for m fixed when $n \rightarrow \infty$ still hold when both m and $n \rightarrow \infty$, namely, that $\hat{x}_k$, $\hat{\alpha}_k$ and $\hat{\lambda}_k$ are consistent estimators of x_k, α_k and λ_k respectively.

5.3 Bias

Although the estimators obtained with the new estimation procedure are consistent, and therefore unbiased in the limit, they are not unbiased for small samples. However, it does not seem feasible to determine analytically the extent of the bias in general for small samples. Instead, in this section an approximation to the bias will be given only for the estimator $\hat{x}$ for Cases 1 and 4. Later the extent of the bias will be investigated empirically in Chapter VI, which gives the results of an extensive sampling survey for Case 1.

Let us first consider the exponential estimator $\hat{x}$ for Case 1, where $\hat{x} = Y_2/Y_1$. If we expand $\hat{x}$ in a Taylor series about the point $\eta = (\eta_1, \eta_2)$, we find that

$$\hat{x} = \frac{\eta_2}{\eta_1} + \sum_{r=1}^{\infty} \frac{(-1)^r \eta_2}{\eta_1^{r+1}} (Y_1 - \eta_1)^r + \sum_{r=1}^{\infty} \frac{(-1)^{r-1} (Y_1 - \eta_1)^{r-1} (Y_2 - \eta_2)}{r\, \eta_1^r} . \tag{5.3.1}$$

This series converges only when $|Y_1 - \eta_1| < \eta_1$ since Y_2/Y_1 has a singularity when $Y_1 = 0$. Now, recalling from Chapter III that η_2/η_1 equals the constant x , we may approximate the difference $(\hat{x} - x)$ by

$$\hat{x} - x \simeq \sum_{r=1}^{N} \frac{(-1)^r \eta_2}{\eta_1^{r+1}} (Y_1 - \eta_1)^r + \sum_{r=1}^{N} \frac{(-1)^{r-1} (Y_1 - \eta_1)^{r-1} (Y_2 - \eta_2)}{r \eta_1^r} , \tag{5.3.2}$$

where only a finite number N of the terms in (5.3.1) are used in the approximation. Taking the expected values of both sides of (5.3.2), we can approximate the bias, $\mathcal{E}(\hat{x} - x)$, of the estimator $\hat{x}$ relative to the constant exponential x as follows:

$$\mathcal{E}(\hat{x} - x) \simeq \sum_{r=2}^{N} \frac{(-1)^r \eta_2}{\eta_1^{r+1}} \mu_r (Y_1) , \tag{5.3.3}$$

where $\mu_r(Y_1)$ is the r^{th} order central moment of Y_1 . Now if in (5.3.3) we replace the expectations η_1 and η_2 by the sample means Y_1 and Y_2 , we obtain the bias approximation

$$\mathcal{E}(\hat{x} - x) \simeq \sum_{r=2}^{N} \frac{(-1)^r Y_2}{Y_1^{r+1}} \mu_r (Y_1) \Bigg|_{\eta=Y} . \tag{5.3.4}$$

Let us evaluate (5.3.4) under the assumption that the errors e_{ij} in our Case 1 model are normally distributed with mean zero and common variance σ^2 . Actually, since the expansion (5.3.1) converges only when Y_1 lies in a circle in the positive quadrant, it appears as if we should further restrict the errors in such a way that Y_1 will always be positive. But Fieller [5] has shown that when $\eta_1 > 0$ and large relative to the standard error of Y_1, such a curtailed normal distribution for the errors e_{ij} differs very little from the usual normal distribution. Now for Case 1, η_1 and η_2 are both positive provided that the coefficient $\alpha > 0$, and $|\eta_1| > |\eta_2|$ even though the corresponding means, Y_1 and Y_2, have the same variance. Therefore, η_1 would be expected to be large in absolute value relative to the standard error of Y_1. Hence, when the Case 1 model is fitted to positive data, there is no need to further restrict the errors e_{ij} once they are assumed to be normally distributed.

When the errors e_{ij} are assumed to be normally distributed,

$$\mu_r(Y_1) = 0 \qquad \text{for } r \text{ odd},$$

$$= (r-1)(r-3)\ldots(1)\left(\frac{\sigma^2}{mn}\right)^{\frac{r}{2}} \quad \text{for } r \text{ even.}$$

Thus substitution for $\mu_r(Y_1)$ in (5.3.4) gives

$$\mathcal{E}(x - x) \simeq \sum_{v=1}^{M} \frac{\eta_2}{\eta_1^{2v+1}} (2v-1)(2v-3)\ldots(1)\left(\frac{\sigma^2}{mn}\right)^{v}\Bigg|_{\eta=Y} \qquad (5.3.5)$$

$$\simeq \hat{x} \sum_{v=1}^{M} (2v-1)(2v-3)\ldots(1)\left(\frac{\sigma^2}{mn\,Y_1^2}\right)^{v},$$

where M is the largest integer in $\frac{N}{2}$. But the right side of the approximation (5.3.5) is always positive. Therefore, when the errors e_{ij} are not only identically distributed as specified by our model but are also normally distributed, the expected bias of $\hat{x}$ for Case 1, which is approximated by (5.3.5), is always positive. This bias decreases as m , n or Y_1 becomes larger or as $\hat{x}$ decreases.

A development similar to that already presented for Case 1 may be used to obtain an approximation to the bias of $\hat{x}$ for Case 4. Corresponding to equation (5.3.1), for Case 4, when $\left| Y_1 - \eta_1 - Y_2 - \eta_2 \right| < \left| \eta_1 - \eta_2 \right| > 0$, we have the expansion

$$\hat{x} = x \left(1 + \frac{Y_2 - \eta_2 - Y_3 + \eta_3}{Y_1 - \eta_1 - Y_2 + \eta_2}\right) \sum_{r=0}^{\infty} (-1)^r \left(\frac{Y_1 - \eta_1 - Y_2 + \eta_2}{\eta_1 - \eta_2}\right)^r . \tag{5.3.6}$$

From this expansion we obtain an approximation, corresponding to that given by (5.3.4) for Case 1, to the bias in $\hat{x}$ for Case 4. This approximation is

$$\mathcal{E}(\hat{x} - x) \simeq \hat{x} \sum_{r=1}^{N} \left. \frac{(-1)^r \sum_{k=0}^{r} \frac{r!}{k!(r-k)!} \mu_k(Y_1)\mu_{r-k}(Y_2)}{(Y_1 - Y_2)^r} \right|_{Y=\eta}$$

$$+ \hat{x} \sum_{r=1}^{N} \left. \frac{(-1)^r \sum_{k=0}^{r} \frac{r!}{k!(r-k)!} \mu_k(Y_1)\mu_{r-k+1}(Y_2)}{(Y_2 - Y_3)(Y_1 - Y_2)^r} \right|_{Y=\eta} . \tag{5.3.7}$$

If we set $N = 6$ and if we assume that the errors e_{ij} in our Case 4 model are normally distributed, (5.3.7) becomes

$$\mathcal{E}(\hat{x} - x) \simeq \frac{2\hat{x}\sigma^2}{mn(Y_1 - Y_2)^2}\left[1 - \frac{3\sigma^2}{mn(Y_2 - Y_3)(Y_1 - Y_2)} + \frac{6\sigma^2}{mn(Y_1 - Y_2)^2}\right.$$

$$\left. - \frac{30\sigma^4}{m^2n^2(Y_2 - Y_3)(Y_1 - Y_2)^3} + \frac{60\sigma^4}{m^2n^2(Y_1 - Y_2)^4}\right] . \qquad (5.3.8)$$

As before for Case 1, the approximation (5.3.8) for Case 4 is valid only if the expectation of the denominator of $\hat{x}$, in this case $(\eta_1 - \eta_2)$, is positive and large relative to the standard error of that denominator. Thus (5.3.8) should be used as an approximation to the bias of $\hat{x}$ relative to x for Case 4 only if $(Y_1 - Y_2)$ is positive and large relative to its standard error even though the errors e_{ij} are assumed to be normally distributed.

5.4 Efficiency

The estimators yielded by the new estimation procedure are not in general efficient, and no measures of their small sample efficiencies are available. However, since maximum likelihood estimators are asymptotically efficient, the asymptotic efficiency of an estimator from the new method can be determined by taking the ratio of the asymptotic variance of the corresponding maximum likelihood estimator to that of the estimator in question.

The asymptotic variances of the maximum likelihood estimators of the parameters in our model can be found by inverting a matrix of products of first partial derivatives of $\mathcal{E}(y_{ij})$ with respect to those parameters (see [9]) . If we let θ_k , k = 1, 2, ..., 2p , represent our Case 3 parameters when the errors e_{ij} are assumed to be normally distributed, then this matrix has elements

$$c_{rs} = \sum_{i=0}^{2pn-1} \sum_{j=1}^{m} \frac{\partial\, \mathcal{E}(y_{ij})}{\partial\, \theta_r} \frac{\partial\, \mathcal{E}(y_{ij})}{\partial\, \theta_s} .$$

For example, for Case 1 with $\theta_1 = \alpha$ and $\theta_2 = \lambda$,

$$\frac{\partial\, \mathcal{E}(y_{ij})}{\partial\, \alpha} = e^{-\lambda t_i} , \quad \frac{\partial\, \mathcal{E}(y_{ij})}{\partial\, \lambda} = -\alpha\, t_i\, e^{-\lambda t_i} ,$$

and the matrix with which we are concerned has elements

$$c_{11} = m \sum_{i=0}^{2pn-1} e^{-2\lambda t_i} ,$$

$$c_{12} = c_{21} = -m\,\alpha \sum_{i=0}^{2pn-1} t_i\, e^{-2\lambda t_i} ,$$

$$c_{22} = m\,\alpha^2 \sum_{i=0}^{2pn-1} t_i^2\, e^{-2\lambda t_i} .$$

If we denote the corresponding elements of the inverse of this matrix by c^{rs} , r , $s = 1, 2$, then the asymptotic variances of the maximum likelihood estimators of α and λ are given by $c^{11}\sigma^2$ and $c^{22}\sigma^2$ respectively.

Now when $n = 1$ and the errors e_{ij} in our model are normally distributed, the estimators obtained by the new procedure, which have already been shown to be sufficient in this instance, are also asymptotically efficient. A glance at (4.2.15) and (4.2.16) and the definition of η_q given by (4.1.4) shows that as m increases, the asymptotic variances of both $\hat{\alpha}$ and $\hat{\lambda}$ for Case 1 decrease proportionally. Also, it can be seen that if everything but α is held fixed, the asymptotic variance of $\hat{\lambda}$ is inversely proportional to α^2 while the variance of $\hat{\alpha}$ is not affected by changes in α . Moreover, if K , the distance between the observation points t_i , is allowed to vary while the product λK as well as m , n and α remain constant, the asymptotic variance of $\hat{\lambda}$ varies inversely with K^2 while the asymptotic variance of $\hat{\alpha}$ is again unaffected. But the asymptotic variances presented above for the maximum likelihood estimators of α and λ for Case 1 can be shown to be influenced in the same way by changes in m , in α , or in K when λK is held constant. Therefore, in order to obtain an idea of the asymptotic efficiency of the estimators yielded by the new method for Case 1, we need only consider the relative effects on the asymptotic variances of the maximum likelihood estimators and those from the new procedure of allowing n to be greater than one and of varying λ without changing K or n . Such a comparison is made in Tables 3 and 4.

The first rows of Tables 3 and 4 give the asymptotic variances divided by σ^2 of the maximum likelihood estimators of α and λ while the second rows contain the corresponding values for the estimators from the new procedure computed through direct substitution in (4.2.15)

Table 3

Case 1 Asymptotic Variances and Efficiencies for Different Values of n

$\alpha = m = 1, \quad T = Kn = 2, \quad \lambda = \ln 2$

	$n = 1$		$n = 2$		$n = 4$		$n = 8$	
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$
M.L. Variances/σ^2	1.000	4.250	0.952	1.974	0.771	1.149	0.519	0.650
N.P. Variances/σ^2	1.000	4.250	1.605	3.778	1.697	2.593	1.251	1.530
Efficiency	1.000	1.000	0.593	0.522	0.455	0.443	0.415	0.425

Table 4

Case 1 Asymptotic Variances and Efficiences for Different Values of λ

$\alpha = m = 1, \quad T = Kn = 2, \quad n = 2$

	$\lambda = \frac{1}{4} \ln 2$		$\lambda = \frac{1}{2} \ln 2$		$\lambda = \ln 2$		$\lambda = 2 \ln 2$	
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$
M.L. Variances/σ^2	0.794	0.345	0.866	0.619	0.952	1.974	0.996	13.229
N.P. Variances/σ^2	0.952	0.443	1.118	0.858	1.605	3.778	4.826	82.240
Efficiency	0.834	0.780	0.775	0.721	0.593	0.522	0.206	0.161

and (4.2.16). The third rows list asymptotic efficiencies of estimators yielded by the new method. Note that these asymptotic efficiencies decrease as either n or λ becomes larger.

Similar results to those already cited for Case 1 have been obtained for Case 4. Again the asymptotic variances of both the maximum likelihood estimators and those from the new procedure are inversely proportional to m while in both cases the asymptotic variances of the estimators of λ are also inversely proportional to α_1^2 and to K^2 when λK is held fixed. Both sets of asymptotic variances for estimators of α_0 and α_1 are unaffected by changes in α_1 and in no instance does the value of α_0 or the sign of α_1 enter into the calculation of asymptotic variances. Tables 5 and 6 for Case 4 correspond to Tables 3 and 4 for Case 1 and indicate the effect of changes in n and λ on the asymptotic efficiencies of the estimators derived with the new procedure. The variances for the new procedure were calculated by substitution in (4.2.18), (4.2.19), and (4.2.20).

Table 5

Case 4 Asymptotic Variances and Efficiencies for Different Values of n

$\pm\alpha_1 = m = 1, \quad T = Kn, \quad \lambda = \ell n\, 2$

	n = 2			n = 4		
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\lambda}$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\lambda}$
M.L. Variances/σ^2	1.109	1.708	5.905	0.494	0.952	3.194
N.P. Variances/σ^2	1.981	1.844	16.593	0.991	1.529	11.387
Efficiency	0.560	0.927	0.356	0.499	0.623	0.280

Table 6

Case 4 Asymptotic Variances and Efficiencies for Different Values of λ

$\pm\alpha_1 = m = 1, \quad T = Kn = 2, \quad n = 2$

	$\lambda = \frac{1}{2}\ \ell n\ 2$			$\lambda = \ell n\ 2$		
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\lambda}$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\lambda}$
M.L. Variances/σ^2	6.564	5.727	3.445	1.109	1.708	5.905
N.P. Variances/σ^2	16.500	12.239	9.608	1.981	1.844	16.593
Efficiency	0.398	0.468	0.359	0.560	0.927	0.356

The results presented in this section are limited in scope and apply only when the errors e_{ij} are normally distributed. However, Tables 3-6 do show that the estimators produced by the new procedure are certainly not in general efficient and that asymptotically they are quite inefficient under the conditions of this section.

5.5 Optimum Construction of the Sums S_q

In Section 3.2 it was indicated that there are several ways of forming the sums S_q from the observations y_{ij} for which essentially the same method for estimating the parameters of our model may be followed. The procedure given there for calculating the S_q was said to be an optimum procedure in some respects. In this section it will be shown that it is indeed a better method than certain alternative methods.

As an alternative construction for the S_q , let us take

$$S_q^* = \sum_{j=1}^{m} \sum_{v=0}^{n-1} y_{2pv+q-1,j} \; ; \; q = 1, 2, \ldots, 2p \quad . \qquad (5.5.1)$$

Note that the subscript i of the model presented in Section 3.1, which is represented in equation (5.5.1) by the subscript (2pv + q-1) , still ranges from zero to (2pn-1) . In order to form the sums S_q^* , we divide the domain of the t_i into n equal intervals instead of 2p intervals as before. Then we let S_1^* be the sum of the observations made at the first observation points in all of the intervals, S_2^* be the sum of the observations from the second observation points of all the intervals, and so on. In the remainder of this section we shall continue to denote entities connected with the alternative construction of the sums S_q with an asterisk as a superscript. So, corresponding to equation (3.2.3), we have

$$\mathcal{E}(S_q^*) = m \sum_{v=0}^{n-1} \sum_{k=1}^{p} \alpha_k \Lambda_k^{2pv-1+q} \quad , \quad q = 1, 2, \ldots, 2p \, , \qquad (5.5.2)$$

which sums to give

$$\mathcal{E}(S_q^*) = m \sum_{k=1}^{p} \alpha_k \Lambda_k^{q-1} \frac{(1 - \Lambda_k^{2pn})}{1 - \Lambda_k^{2p}} \quad , \quad q = 1, 2, \ldots, 2p \, . \qquad (5.5.3)$$

As before, $\Lambda_k = \exp.(-\lambda_k K)$. Following the same procedure used in Section 3.3, we set (5.5.3) equal to S_q^* , and then we solve for estimators

α_k^* and λ_k^* of the parameters α_k and λ_k.

In order to facilitate the solutions for α_k^* and λ_k^*, it is again expedient to resort to matrix algebra. We shall once more make use of the matrix α defined in Section 3.3, and we shall form column matrices $\bar{s}^*$ and $\underline{s}^*$ by substituting S_q^* for S_q in the matrices $\bar{s}$ and $\underline{s}$ previously defined. It is also necessary to modify the definitions of the elements of the p by p matrices L, W and V used in Section 3.3 as follows:

$$\ell_{rs}^* = \Lambda_s^{r-1} \quad ,$$

$$w_{rs}^* = \frac{m(1 - \Lambda_r^{2pm})}{1 - \Lambda_r^{2p}} \delta_{rs} \quad ,$$

$$v_{rs}^* = \Lambda_r^p \delta_{rs} \quad .$$

Corresponding to equations (3.3.1) and (3.3.2) for the S_q, the equations for the S_q^* may now be represented by the two matrix equations

$$L^* W^* \alpha = \bar{s}^* \quad , \tag{5.5.4}$$

$$L^* W^* V^* \alpha = \underline{s}^* \quad . \tag{5.5.5}$$

Solving (5.5.4) for α and substituting the result in (5.5.5), we arrive at the equation

$$L^* V^* L^{*-1} \bar{s}^* = \underline{s}^* \quad . \tag{5.5.6}$$

Because of the definitions given above, the vectors $\bar{s}^*$ and $\underline{s}^*$ in the equation (5.5.6) are of the same form in terms of the S_q^* as are the matrices $\bar{s}$ and $\underline{s}$ of Section 3.3 in terms of the S_q . Furthermore, V^* and L^* are of the same form in the Λ_k that the matrices V and L of Section 3.3 are in terms of the x_k . Thus, from the analogy between (5.5.6) and (3.3.4), we see that the solution given in Section 3.3 for the x_k is the correct solution for the Λ_k in terms of the newly defined S_q^* . That is, the solution obtained for Λ_k^n in terms of the S_q defined in Section 3.3 is now the solution for Λ_k in terms of the S_q^* . Hence, it follows that solutions for $\hat{\alpha}_k^*$ and $\hat{\lambda}_k^*$ may easily be obtained in much the same way that $\hat{\alpha}_k$ and $\hat{\lambda}_k$ were derived in Section 3.3.

Not only can the S_q^* be used to obtain estimators in a way similar to that developed for the S_q , but some of the properties of $\hat{\alpha}_k$ and $\hat{\lambda}_k$ can also be shown to hold for the estimators $\hat{\alpha}_k^*$ and $\hat{\lambda}_k^*$. In particular, the estimators $\hat{\alpha}_k^*$ and $\hat{\lambda}_k^*$ are consistent for large m when n is held fixed. However, consistency for m fixed and n large no longer obtains. To demonstrate this, let us attempt to parallel the consistency proof given in Section 5.2 as $n \longrightarrow \infty$, but with the variable S_q^* instead of S_q .

From (5.5.1) it follows that

$$S_q^* = m \sum_{v=0}^{n-1} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k K(2pv+q-1)} + \sum_{v=0}^{n-1} \sum_{j=1}^{m} e_{ij} \quad . \tag{5.5.7}$$

Carrying out the same sort of manipulations that were used in Section 5.2, we find that (5.5.7) yields

$$Y_q^* = \frac{1}{2pT} \sum_{v=0}^{n-1} \sum_{k=1}^{p} \alpha_k e^{-\lambda_k \frac{T}{n}(2pv+q-1)} 2p \frac{T}{n} + \frac{1}{mn} \sum_{v=0}^{n-1} \sum_{j=1}^{m} e_{ij} .$$

The last term, which is the mean of the independent, identically distributed errors e_{ij} , is again zero in the limit as $n \to \infty$. So, referring once more to the definition of a definite integral [8] , we see that

$$\lim_{n \to \infty} Y_q^* = \frac{1}{2pT} \sum_{k=1}^{p} \int_0^{2pT} \alpha_k e^{-\lambda_k t} dt$$

$$= \frac{1}{2pT} \sum_{k=1}^{p} \frac{\alpha_k}{\lambda_k} (1 - e^{-\lambda_k 2pT}) . \qquad (5.5.8)$$

But the right side of (5.5.8) is independent of q , and hence in the limit as $n \to \infty$, all the Y_q^* are equal. That is, the constant limits φ_q^* corresponding to the φ_q defined by equation (4.3.4) are all equal to the constant given by the right side of equation (5.5.8). Therefore, the determinant $|P(\varphi^*)|$, corresponding to the determinant $|P(\varphi)|$ evaluated by equation (4.3.10), is singular and equal to zero.

Since $|P(\varphi^*)| = 0$ when the new estimation procedure is developed in terms of the S_q^* , it follows from equation (3.3.23) that the estimators $E_r^* \to \infty$ if $n \to \infty$. Hence the exponential estimators $\hat{x}_k^*$ also $\to \infty$ if $n \to \infty$ and are then neither admissible nor consistent. Therefore, the estimators $\hat{\alpha}_k^*$ and $\hat{\lambda}_k^*$, which are computed from the $\hat{x}_k^*$, are not defined for large n either, and hence they are not consistent estimators of the parameters α_k and λ_k . So it is evident that increasing the number of observation points t_i would not be likely to improve the accuracy of the estimators obtained with the alternative method presented in this section for forming the sums S_q . Therefore, the construction presented for the sums S_q in Chapter III is better than that given in this section for the S_q^* .

So far we have considered only one alternative formation for the sums S_q which leads to summable geometric series for the expectations of the S_q , and which is therefore amenable to an estimation procedure similar to that developed in Chapter III. There are many other alternative constructions which involve both the approach used to obtain the S_q and that used to arrive at the S_q^* . For instance, the domain of the observation points t_i might be divided in half, with the observations from the first half being used to form S_1 , S_2 , ..., S_p by one of these methods and with the remainder of the observations being used to form S_{p+1} , S_{p+2}, ..., S_{2p} by the other method. All such constructions would make at least two of the constants φ_r and φ_s equal for $r \neq s$, and hence, like the alternative method already considered, they would result in estimators which would not be consistent for large n . The other likely alternative constructions are such that the S_q would not all be sums of the same number of y_{ij} . This would complicate the solutions for the estimators considerably and would not be advantageous except perhaps in special cases. Hence, we conclude that the construction for the sums S_q presented in Chapter III is an optimum construction, at least by comparison with the alternative constructions considered here.

5.6 Conditions for Obtaining Admissible Estimates

So far in this chapter it has been shown that the estimators yielded by the new estimation procedure are consistent, but not in general sufficient, unbiased or asymptotically efficient. It has also been demonstrated that the construction given in Chapter III for the sums S_q is better than several alternative constructions which would lead to the same sort of estimation procedure as that presented in this paper. Now we shall study conditions for the existence of admissible solutions for the estimators of the new estimation procedure.

In the model specified in Chapter III, $\lambda_k > 0$, $k = 1, 2, \ldots, p$ and $\lambda_r \neq \lambda_s$ for $r \neq s$, making $0 < x_k < 1$, $k = 1, 2, \ldots, p$. In addition, the x_k are real and distinct. So in order for the estimators $\hat{x}_k$ to be admissible, we shall require that $0 < \hat{x}_k < 1$, $k = 1, 2, \ldots, p$, and that the $\hat{x}_k$ be distinct and real. Now the $\hat{x}_k$ are the p roots of the polynomial

$$\hat{E}_0 x^p - \hat{E}_1 x^{p-1} + \hat{E}_2 x^{p-2} - \ldots + (-1)^{p-1} \hat{E}_{p-1} x + (-1)^p \hat{E}_p = 0 \,, \tag{5.6.1}$$

where $\hat{E}_0 \equiv 1$. Since the $\hat{E}_r$ are the elementary symmetric functions of the $\hat{x}_k$, in order for $0 < \hat{x}_k < 1$, $k = 1, 2, \ldots, p$, it is necessary that

$$0 < \hat{E}_r < \frac{p!}{r!\,(p-r)!} \,, \quad r = 1, 2, \ldots, p \,,$$

and that

$$\hat{E}_{p-r} < \hat{E}_r \; ; \; r = 0, 1, 2, \ldots, \; \left[(p-1)/2\right] \quad ,$$

where $\left[(p-1)/2\right]$ denotes the largest integer in $(p-1)/2$. Also, it is necessary that $\hat{E}_p < \hat{E}_1$. Furthermore, given $\hat{E}_r > 0$, it follows from Newton's rule of signs that a necessary condition for all the $\hat{x}_k$ to be real is that each of the quantities

$$\left(\frac{\hat{E}_1}{{}_pC_1}\right)^2 - \frac{\hat{E}_0}{{}_pC_0}\,\frac{\hat{E}_2}{{}_pC_2} \; ; \left(\frac{\hat{E}_2}{{}_pC_2}\right)^2 - \frac{\hat{E}_1}{{}_pC_1}\,\frac{\hat{E}_3}{{}_pC_3} \; ; \ldots$$

$$\ldots \left(\frac{\hat{E}_{p-1}}{{}_pC_{p-1}}\right)^2 - \frac{\hat{E}_{p-2}}{{}_pC_{p-2}}\,\frac{\hat{E}_p}{{}_pC_p}$$

be positive, where

$${}_pC_k = \frac{p!}{k!\,(p-k)!} \quad .$$

Note that these conditions on the $\hat{E}_r$ are only necessary conditions, and that they are not sufficient to guarantee an admissible solution for the $\hat{x}_k$. Additional conditions which are both necessary and sufficient for the roots of (5.6.1), whether they are real or complex, to be less than one in absolute value are given by Samuelson [25] .

Instead of testing the $\hat{E}_r$ against all of the conditions given above, it is usually more expedient merely to compute the $\hat{E}_r$ and note whether or not they are all positive. Then if one or more of the $\hat{E}_r$ is negative, the set of estimators $\hat{x}_k$ is not admissible. But if the $\hat{E}_r$ are all positive, Sturm's theorem [28, pp. 103-107] may be used to determine the number of real roots of the polynomial (5.6.1) which lie between zero and one. Sturm's theorem states that

> "there exists a set of real polynomials $f(x)$, $f'(x)$, $f_2(x)$, $f_3(x)$, ..., $f_m(x)$ whose degrees are in descending order, such that, if $b > a$, the number of distinct real roots of $f(x) = 0$ between $x = a$ and $x = b$ is equal to the excess of the number of changes of sign in the sequence $f, f', f_2, \ldots, f_m$ when $x = a$ over the number of changes in sign when $x = b$."

$f'(x)$ denotes the first derivative of $f(x)$. Now let q_1 be the quotient and $(-f_2)$ the remainder in the division of f by f'. Then $f_2(x)$ is given by

$$f_2 = q_1 f' - f \quad .$$

The other functions of Sturm's theorem may be similarly defined as follows:

$$f_3 = q_2 f_2 - f' \quad ,$$

$$f_4 = q_3 f_3 - f_2 \quad ,$$

$$\vdots$$

$$f_m = q_{m-1} f_{m-1} - f_{m-2} \quad .$$

The new estimation procedure leads to an admissible set of estimators $\hat{x}_k$ if and only if the application of Sturm's theorem shows that there are p real roots $\hat{x}_k$ between zero and one, and provided the $\hat{x}_k$ are distinct. Fulfillment of this latter condition can usually be demonstrated only by solving for the roots $\hat{x}_k$ of the polynomial (5.6.1).

Sturm's theorem gives a satisfactory way of testing for an admissible set of estimates $\hat{x}_k$ once the coefficients $\hat{E}_r$ have been calculated. Also, the conditions given above for the $\hat{E}_r$ which are necessary for an admissible solution may be helpful in weeding out inadmissible solutions, but again they cannot be applied unless the $\hat{E}_r$ have been calculated. Since the calculation of the $\hat{E}_r$ is rather arduous for $p > 3$, it would be desirable to obtain conditions for an admissible solution for the $\hat{x}_k$ which could be imposed upon the sums S_q. However, no such conditions which can be readily applied have been found except when $p = 1$ or 2.

For Case 1, where $p = 1$ and where $\hat{x}$ corresponds to the $\hat{x}_k$ in the above discussion, $\hat{x} = S_2/S_1$. It is immediately evident that $\hat{x}$ is admissible, that is, that $\hat{x}$ is real and lies between zero and one, whenever S_1 and S_2 are not zero and are of the same sign with $|S_1| > |S_2|$. For the modified exponential function, Case 4,

$$\hat{x} = \frac{S_2 - S_3}{S_1 - S_2} \quad .$$

Now $\hat{x}$ is an admissible estimator whenever the sequence S_1, S_2, S_3 is strictly monotone, either increasing or decreasing, with $|S_1 - S_2| > |S_2 - S_3|$.

One of the conditions necessary for an admissible solution for the $\hat{x}_k$ for Case 2 is that $\hat{E}_1$ and $\hat{E}_2$ be positive. From equations

(3.5.7) it can be seen that this condition requires that the expressions $S_1 S_4 - S_2 S_3$, $S_1 S_3 - S_2^2$ and $S_2 S_4 - S_3^2$ all be of the same sign. Factoring $S_2 S_4$ out of the first expression, $S_2 S_3$ out of the second, and $S_3 S_4$ out of the third, it follows that these expressions will be of the same sign when all the S_q are positive if and only if the sequence S_1/S_2 , S_2/S_3 , S_3/S_4 is strictly monotone. Since the S_q usually are all positive in practice, this is a convenient necessary condition which often may be easily used to eliminate an inadmissible solution. The other necessary conditions given previously for the $\hat{E}_r$ may also be expressed in terms of the S_q for Case 2, but they are sufficiently complicated so that it is as easy to carry out the actual solution as it is to make the tests.

The discussion given above concerning conditions which must be satisfied if an admissible solution is to result does not give any clear indication of whether or not the new estimation procedure will lead to admissible estimates in most practical problems. In the sampling survey which will be reported in Chapter VI, some idea of how often admissible solutions may be expected to result will be gained. Also, situations which lead to inadmissible estimates will be more clearly depicted and in Chapter VII it will be shown that for some such situations the new estimation procedure may be used to fit a different model than that specified in this paper. One more point should be brought up here. So far in this section we have only discussed admissible solutions for the $\hat{x}_k$. But when $\hat{x}_k$ is admissible, $\hat{\lambda}_k$ will be real and positive. Hence $\hat{\lambda}_k$ will have the same range as that specified for λ_k , and so when $\hat{x}_k$ is admissible, $\hat{\lambda}_k$ will be admissible too. The same is true of the $\hat{\alpha}_k$ except in rare instances when the $\hat{x}_k$ are admissible but one of the $\hat{\alpha}_k$ is zero.

VI. SMALL SAMPLE STUDIES

6.1 Distributions and Confidence Limits

In this chapter we shall obtain confidence limits for estimators from the new procedure and then we shall present the data from an empirical sampling study. However, it is first of interest to note that although no work has been done on the exact distributions of the estimators $\hat{\alpha}_k$ and $\hat{\lambda}_k$ for either Case 3 or Case 6, the small sample distributions of the $\hat{x}_k$ have been considered. For either Case 1 or Case 4, when the errors e_{ij} are normally distributed, $\hat{x}$ is the ratio of two normally distributed variates, and its distribution has been studied by Fieller [5] and Merrill [22]. For the general cases, Case 3 and Case 6, the $\hat{x}_k$ are roots of the polynomial (3.3.25) with the $\hat{E}_r$, which are real and continuous, as coefficients. The distributions of such roots have been investigated by Hamblen [12] and Girshick [10]. Although these papers are of mathematical interest, the distributions derived are too complex to yield distributions or confidence limits for either the $\hat{\alpha}_k$ or the $\hat{\lambda}_k$.

Exact confidence limits are available, however, for λ for Cases 1 and 4, provided that the e_{ij} are normally distributed. Fieller [6] shows that

> "if y and z are estimates of ζ and η subject to random errors normally distributed about zero mean, and if v_{yy}, v_{yz}, v_{zz} are joint estimates, based on f degrees of freedom and independent of y and z, of the variances and covariance of the error distribution, then the fiducial [confidence] range for $\beta = \eta/\zeta$ consists of these values for which

$$(z^2 - t^2 v_{zz}) - 2\beta(yz - t^2 v_{zy}) + \beta^2 (y^2 - t^2 v_{yy}) \leq 0$$

where t is the appropriate level of the Student distribution for f degrees of freedom."

An estimate s^2 for σ^2 may be obtained as indicated in the next paragraph. Then for Case 1, with $y = S_1$, $z = S_2$, $\zeta = \mathcal{E}(S_1)$, $\eta = \mathcal{E}(S_2)$, $v_{yz} = 0$, $v_{yy} = mns^2 = v_{zz}$, $\beta = \mathcal{E}(S_1)/\mathcal{E}(S_2) = x$, it follows from Fieller's theorem that a confidence interval for x consists of those values of x for which

$$(S_2^2 - mn\, s^2\, t_\alpha^2) - 2x\, S_1 S_2 + x^2 (S_1^2 - mn\, s^2\, t_\alpha^2) \leq 0\,, \qquad (6.1.1)$$

where t_α is the α-level critical value of the Student t-statistic with the same number of degrees of freedom as the estimate s^2. The inequality (6.1.1) is equivalent to the confidence interval

$$\frac{S_1 S_2 - s\left[mn\,(S_1^2 + S_2^2 - mn\, s^2\, t_\alpha^2) \right]^{1/2} t_\alpha}{S_1^2 - mn\, s^2\, t_\alpha^2} \leq x \leq$$

$$\frac{S_1 S_2 + s\left[mn\,(S_1^2 + S_2^2 - mn\, s^2\, t_\alpha^2 \right]^{1/2} t_\alpha}{S_1^2 - mn\, s^2\, t_\alpha^2}, \qquad (6.1.2)$$

which is more convenient for calculation. Since $\lambda = \frac{1}{T} \ln x$, where $T = Kn$, and λ is therefore a monotone function of x, for λ, corresponding to (6.1.2), there exists the α-level confidence interval

$$-\frac{1}{T}\ln\left\{S_1 S_2 + s\left[mn \quad (S_1^2 + S_2^2 - mn\, s^2\, t^2_{\alpha})\right]^{1/2} \quad t_{\alpha}\right\}$$

$$\leq \quad \lambda - \frac{1}{T}\ln\,(S_1^2 - mn\, s^2\, t_{\alpha}^2\,)$$

$$\leq \quad -\frac{1}{T}\ln\left\{S_1\, S_2 - s\left[mn\, (S_1^2 + S_2^2 - mn\, s^2\, t_{\alpha}^2\right]^{1/2} \quad t_{\alpha}\right\}. \tag{6.1.3}$$

We mentioned above that an estimate s^2 of σ^2 is available. If m , the number of observations made at each point t_i , is greater than one, an estimate s_i^2 of σ^2 may be formed for each t_i by computing

$$s_i^2 = \frac{1}{m-1} \sum_{j=1}^{m} (y_{ij} - \bar{y}_i) \tag{6.1.4}$$

where

$$\bar{y}_i = \frac{1}{m} \sum_{j=1}^{m} y_{ij} \ . \tag{6.1.5}$$

Since under our model the e_{ij} are assumed to be homogeneous, the s_i^2 may then be pooled to form

$$s^2 = \frac{1}{2pn} \sum_{i=0}^{2pn-1} s_i^2 \tag{6.1.6}$$

with $2pn(m-1)$ degrees of freedom for Case 3 or

$$s^2 = \frac{1}{(2p+1)n} \sum_{i=0}^{(2p+1)n-1} s_i^2 \qquad (6.1.7)$$

with $(2p+1)n\,(m-1)$ degrees of freedom for Case 6. As in the case of a linear regression, σ^2 may also be estimated using the mean square deviation of the y_{ij} from regression. The usual practice for a non-linear regression is to assign the same number of degrees of freedom to this estimate as it would have in the linear case, namely, $2p(n-1)$ for Case 3 and $(2p+1)(n-1)$ for Case 6. If, when $m > 1$, the expected mean square deviation from regression is not greater than the expected error mean square estimated by (6.1.6) or (6.1.7), the two mean squares may be pooled in the calculation of an estimate s^2 for σ^2. When $m = 1$, the mean square deviation from regression is the only estimate available for σ^2.

Now that the estimation of σ^2 has been discussed, let us also apply Fieller's theorem to Case 4 when the e_{ij} are assumed to be normally distributed. In this case we let $y = S_1 - S_2$, $z = S_2 - S_3$, $\zeta = \mathcal{E}(S_1) - \mathcal{E}(S_2)$, $\eta = \mathcal{E}(S_2) - \mathcal{E}(S_3)$, $v_{yz} = -mn\, s^2$, $v_{yy} = v_{zz} = 2mn\, s^2$, $\beta = [\mathcal{E}(S_2) - \mathcal{E}(S_3)] / [\mathcal{E}(S_1) - \mathcal{E}(S_2)] = x$. So a confidence interval for x consists of those values of x for which

$$\left[(S_2 - S_3)^2 - 2\,mn\,s^2 t_\alpha^2\right] - 2x\left[(S_1 - S_2)(S_2 - S_3) + mn\,s^2 t_\alpha^2\right]$$

$$+ x^2\left[(S_1 - S_2)^2 - 2\,mn\,s^2 t_\alpha^2\right] \leq 0. \qquad (6.1.8)$$

Like (6.1.1) for Case 1, for Case 4 the inequality (6.1.8) leads to a confidence interval for λ . If we let

$$L_r = -\frac{1}{T}\ln\left\{\left[(S_1 - S_2)(S_2 - S_3) + mn\, s^2\, t_\alpha^2\right] + (-1)^r\, s\,(mn)^{1/2}\right.$$
$$\left.\left[2(S_1 - S_2)(S_2 - S_3) + 2(S_1 - S_2)^2 + 2(S_2 - S_3)^2 - 3mn\, t_\alpha^2\, s^2\right]^{\frac{1}{2}} t_\alpha\right\}$$
$$+ \frac{1}{T}\ln\left[(S_1 - S_2)^2 - 2\, mn\, s^2\, t_\alpha^2\right], \; r = 0, 1 , \qquad (6.1.9)$$

this α-level confidence interval is

$$L_0 \leq \lambda \leq L_1 \; . \qquad (6.1.10)$$

In addition to the special cases already considered, approximate confidence limits for any parameter α_k or λ_k estimated by the new procedure may be derived from the results of Chapter IV. In Section 4.4 we found that (4.4.5) has asymptotically a standard normal distribution. Thereford, if we again let $\hat{\theta}$ represent any estimator derived with the new procedure and let θ be the corresponding parameter, it can be shown that the distribution of

$$\frac{\sqrt{mn}\,(\hat{\theta} - \theta)}{s\sqrt{\sum_q A_q^2}} \qquad (6.1.11)$$

approximates the Student t-distribution with the number of degrees of freedom assigned to s^2 , where now

$$a_q \Big|_{\eta = Y} = A_q = \frac{\partial \hat{\theta}}{\partial Y_q} . \qquad (6.1.12)$$

Hence, for an approximate α-level confidence interval about θ , we take

$$\hat{\theta} - \left(\frac{1}{mn} \sum_q A_q^2\right)^{\frac{1}{2}} s\, t_\alpha \leq \theta \leq \hat{\theta} + \left(\frac{1}{mn} \sum_q A_q^2\right)^{\frac{1}{2}} s\, t_\alpha . \qquad (6.1.13)$$

The computation of confidence limits using (6.1.10) and (6.1.13) will be illustrated in Chapter VII.

6.2 An Empirical Study for Case 1

In order to learn more about the small sample characteristics of the estimators developed in this paper, we investigated the properties of these estimators empirically for Case 1. The computations were done on the Oak Ridge National Laboratory's automatic digital computer, the Oracle. A more extensive study was originally planned, but it has been possible to consider only this special case during the time this paper has been in preparation. Nevertheless, the results reported here will help in our evaluation of the new procedure.

For Case 1, observations y_{ij} were generated in accordance with the model presented in Section 3.1 with the additional specification that the errors e_{ij} be normally distributed. Each e_{ij} was computed by first generating sixteen random variates from a rectangular distribution with zero mean and then taking the mean of these variates. In the calculation of the errors, σ was taken to be forty per cent of the mean expected value of y_{ij}. Then m was allowed to take on the values 1, 2, 4, 8, 16 and 32, since, for instance, doubling m may also be interpreted as halving σ^2.

The computations were carried out with λ taking on four different values, namely, $\frac{1}{4}\ln 2$, $\frac{1}{2}\ln 2$, $\frac{3}{4}\ln 2$ and $\ln 2$. Also, n was set equal to 2, 4, 6, 8 and 16. Each of the samples was generated with $T = Kn = 2$ and with $\alpha = 1$. So altogether, 120 sets of parameters were used in the calculations. The choice of λ values, as we shall see later, makes it possible to investigate, for instance, the number of half lives which should be observed in order to accurately estimate the rate of decay of a radioactive substance. Also, this empirical study may be extended to any non-zero value of α, for, under the conditions of our study, changing α by a given factor would not affect $\hat{\lambda}$ or its variance, but it would multiply $\hat{\alpha}$ by that factor and the variance of $\hat{\alpha}$ by the square of that factor. For each set of parameters the calculations were continued until 1024 samples were generated which led to admissible estimates. Meanwhile, the number of inadmissible solutions obtained was recorded. The proportions of inadmissible solutions, for all sets of parameters for which such solutions occurred, are given later in Table 18. The distributions of the estimates $\hat{\alpha}$ and $\hat{\lambda}$ were also recorded, as were the sample means and variances of the estimates. The sample means and variances of $\hat{\lambda}$ computed for Case 1 are displayed in Tables 7-10 while those for $\hat{\alpha}$ are given in Tables 11-14.

Table 7

SAMPLE MEANS AND VARIANCES OF $\hat{\lambda}$

Case 1: $\lambda = \frac{1}{4} \ell n\ 2 = .17329$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	$\bar{\hat{\lambda}}$	.26860	.21347	.18819	.18402	.17993
	$V(\hat{\lambda})$	.04027	.01885	.01208	.01014	.00565
2	$\bar{\hat{\lambda}}$	.21418	.18453	.17776	.17695	.17309
	$V(\hat{\lambda})$	.01876	.00932	.00691	.00528	.00272
4	$\bar{\hat{\lambda}}$	.18414	.17194	.17534	.17460	.17306
	$V(\hat{\lambda})$	.00895	.00532	.00334	.00280	.00153
8	$\bar{\hat{\lambda}}$	.17771	.17423	.17481	.17384	.17369
	$V(\hat{\lambda})$	.00549	.00280	.00176	.00139	.00066
16	$\bar{\hat{\lambda}}$	.17445	.17296	.17269	.17369	.17345
	$V(\hat{\lambda})$	.00287	.00135	.00094	.00071	.00034
32	$\bar{\hat{\lambda}}$	.17314	.17466	.17298	.17222	.17312
	$V(\hat{\lambda})$	.00128	.00063	.00050	.00034	.00017

Table 8

SAMPLE MEANS AND VARIANCES OF $\hat{\lambda}$

Case 1: $\lambda = \frac{1}{2}\ln 2 = .34657$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	$\bar{\hat{\lambda}}$	.40387	.36855	.35917	.36313	.34925
	$v(\hat{\lambda})$	.07633	.03609	.02026	.01776	.00747
2	$\bar{\hat{\lambda}}$	.38508	.35447	.35114	.35237	.34763
	$v(\hat{\lambda})$	.04206	.01569	.00945	.00751	.00366
4	$\bar{\hat{\lambda}}$	.35260	.34791	.34705	.34915	.34697
	$v(\hat{\lambda})$	.01563	.00681	.00474	.00360	.00187
8	$\bar{\hat{\lambda}}$	.35180	.34817	.34717	.34918	.34758
	$v(\hat{\lambda})$	.00728	.00368	.00252	.00182	.00089
16	$\bar{\hat{\lambda}}$	.34559	.34580	.34556	.34659	.34751
	$v(\hat{\lambda})$	.00362	.00177	.00117	.00087	.00046
32	$\bar{\hat{\lambda}}$	.34610	.34645	.34806	.34667	.34605
	$v(\hat{\lambda})$	.00168	.00090	.00057	.00045	.00023

Table 9

SAMPLE MEANS AND VARIANCES OF $\hat{\lambda}$

Case 1: $\lambda = \frac{3}{4}\ell n\ 2 = .51986$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	$\bar{\hat{\lambda}}$	.56358	.56597	.53871	.54126	.53195
	$V(\hat{\lambda})$	.10925	.06588	.03537	.02661	.01138
2	$\bar{\hat{\lambda}}$	.56561	.54429	.54461	.52758	.52270
	$V(\hat{\lambda})$	.07119	.02363	.01646	.01165	.00549
4	$\bar{\hat{\lambda}}$	.53827	.52756	.52690	.52374	.52199
	$V(\hat{\lambda})$	.02642	.01182	.00689	.00546	.00249
8	$\bar{\hat{\lambda}}$	.53017	.52346	.52332	.52205	.52037
	$V(\hat{\lambda})$	.01122	.00509	.00370	.00262	.00136
16	$\bar{\hat{\lambda}}$	.52119	.52121	.52000	.52016	.52174
	$V(\hat{\lambda})$	.00505	.00271	.00172	.00128	.00068
32	$\bar{\hat{\lambda}}$	.52188	.52004	.52049	.52039	.52077
	$V(\hat{\lambda})$	.00274	.00129	.00082	.00062	.00033

Table 10

SAMPLE MEANS AND VARIANCES OF $\hat{\lambda}$

Case 1: $\lambda = \ln 2 = .69315$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	$\bar{\hat{\lambda}}$	.74090	.75742	.74609	.72464	.70333
	$V(\hat{\lambda})$	.17459	.12672	.07491	.05131	.02077
2	$\bar{\hat{\lambda}}$	.75802	.72593	.71080	.71305	.70481
	$V(\hat{\lambda})$	.11015	.04908	.03134	.02196	.00911
4	$\bar{\hat{\lambda}}$	.73395	.70627	.69991	.69999	.69679
	$V(\hat{\lambda})$	.05105	.01929	.01311	.00921	.00410
8	$\bar{\hat{\lambda}}$	.71267	.69422	.69904	.70162	.69429
	$V(\hat{\lambda})$	.01945	.00856	.00620	.00433	.00209
16	$\bar{\hat{\lambda}}$	.69912	.69628	.69559	.69602	.69553
	$V(\hat{\lambda})$	.00878	.00431	.00280	.00203	.00097
32	$\bar{\hat{\lambda}}$	.69944	.69782	.69502	.69372	.69394
	$V(\hat{\lambda})$	.00449	.00219	.00145	.00101	.00054

Table 11

SAMPLE MEANS AND VARIANCES OF $\hat{\alpha}$

Case 1: $\lambda = \frac{1}{4} \ell n\ 2 = .17329$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	$\bar{\hat{\alpha}}$	1.10565	1.06526	1.02912	1.01586	1.01211
	$V(\hat{\alpha})$	.07412	.05069	.03722	.03112	.01830
2	$\bar{\hat{\alpha}}$	1.04753	1.02247	1.00948	1.00764	1.00104
	$V(\hat{\alpha})$	.03606	.02610	.02124	.01606	.00867
4	$\bar{\hat{\alpha}}$	1.01895	1.00219	1.00553	1.00660	1.00229
	$V(\hat{\alpha})$	.02146	.01444	.00955	.00871	.00499
8	$\bar{\hat{\alpha}}$	1.00557	1.00072	1.00300	.99913	1.00214
	$V(\hat{\alpha})$	.01124	.00789	.00516	.00413	.00215
16	$\bar{\hat{\alpha}}$	1.00021	1.00151	.99986	1.00060	1.00039
	$V(\hat{\alpha})$	.00610	.00368	.00269	.00211	.00112
32	$\bar{\hat{\alpha}}$	.99984	1.00100	1.00039	.99907	.99947
	$V(\hat{\alpha})$	.00277	.00174	.00143	.00103	.00057

Table 12

SAMPLE MEANS AND VARIANCES OF $\hat{\alpha}$

Case 1: $\lambda = \frac{1}{2} \ell n\ 2 = .34657$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	$\bar{\hat{\alpha}}$	1.04065	1.02595	1.01566	1.02488	1.00466
	$V(\hat{\alpha})$	.06302	.05296	.03621	.03431	.01486
2	$\bar{\hat{\alpha}}$	1.01674	1.00838	1.00822	1.00969	1.00233
	$V(\hat{\alpha})$	.03876	.02585	.01778	.01459	.00791
4	$\bar{\hat{\alpha}}$	1.00062	1.00137	1.00026	1.00210	1.00029
	$V(\hat{\alpha})$	.01959	.01172	.00834	.00730	.00378
8	$\bar{\hat{\alpha}}$	1.00285	1.00016	1.00064	1.00264	1.00217
	$V(\hat{\alpha})$	.00949	.00629	.00458	.00367	.00183
16	$\bar{\hat{\alpha}}$	.99806	.99846	.99873	1.00115	1.00163
	$V(\hat{\alpha})$	.00441	.00294	.00223	.00176	.00100
32	$\bar{\hat{\alpha}}$	.99755	1.00101	1.00155	1.00086	.99879
	$V(\hat{\alpha})$	.00231	.00157	.00113	.00089	.00050

Table 13

SAMPLE MEANS AND VARIANCES OF $\hat{\alpha}$

Case 1: $\lambda = \frac{3}{4} \ln 2 = .51986$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	$\bar{\hat{\alpha}}$	1.02012	1.03710	1.01548	1.01806	1.01606
	$V(\hat{\alpha})$	.05896	.05486	.03745	.03047	.01469
2	$\bar{\hat{\alpha}}$	1.01639	1.01927	1.02384	1.00997	1.00620
	$V(\hat{\alpha})$	.03207	.02408	.01885	.01394	.00698
4	$\bar{\hat{\alpha}}$	1.00481	1.00344	1.00344	1.00396	1.00247
	$V(\hat{\alpha})$	.01562	.01162	.00811	.00655	.00351
8	$\bar{\hat{\alpha}}$	1.00647	1.00179	1.00424	1.00223	1.00108
	$V(\hat{\alpha})$	.00804	.00525	.00402	.00319	.00180
16	$\bar{\hat{\alpha}}$	1.00029	1.00107	.99995	1.00049	1.00231
	$V(\hat{\alpha})$	.00368	.00277	.00209	.00156	.00097
32	$\bar{\hat{\alpha}}$	1.00202	.99972	1.00021	1.00234	1.00063
	$V(\hat{\alpha})$	.00193	.00151	.00093	.00084	.00045

Table 14

SAMPLE MEANS AND VARIANCES OF $\hat{\alpha}$

Case 1: $\lambda = \ell n\ 2 = .69315$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	$\bar{\hat{\alpha}}$	1.00805	1.03636	1.04081	1.02354	1.01021
	$V(\hat{\alpha})$	..05597	.06306	.04506	.03784	.01844
2	$\bar{\hat{\alpha}}$	1.01943	1.02348	1.01741	1.01605	1.01020
	$V(\hat{\alpha})$	.02987	.02721	.02147	.01619	.00839
4	$\bar{\hat{\alpha}}$	1.01249	1.00832	1.00566	1.00610	1.00461
	$V(\hat{\alpha})$	.01485	.01193	.00911	.00757	.00384
8	$\bar{\hat{\alpha}}$	1.00397	1.00100	1.00256	1.00666	1.00099
	$V(\hat{\alpha})$	.00739	.00552	.00454	.00358	.00198
16	$\bar{\hat{\alpha}}$	1.00105	1.00334	1.00129	1.00198	1.00228
	$V(\hat{\alpha})$	.00370	.00272	.00212	.00168	.00091
32	$\bar{\hat{\alpha}}$	1.00263	1.00184	1.00215	1.00146	1.00083
	$V(\hat{\alpha})$	.00181	.00148	.00108	.00083	.00049

Let us examine the variances in Tables 7-14 to determine the accuracy of the estimation relative to changes in m and n for constant λ, α and $T = Kn$. It is important in this analysis that we keep in mind the dual role of m of either determining the magnitude of σ or of specifying the number of observations taken for each t_i. For example, data recorded for $m = 4$, instead of being interpreted as having occurred with $\sigma = .4/2pn \sum_{i=0}^{2pn-1} \mathcal{E}(y_{ij})$ and $m = 4$, could be thought of as having arisen with $\sigma = .2/2pn \sum_{i=0}^{2pn-1} \mathcal{E}(y_{ij})$ and $m = 1$. Thus the variances given for $m = 1$, though quite large, are not alarming since in this instance σ is also large. Note that the sample variances of both $\hat{\lambda}$ and $\hat{\alpha}$ are approximately halved each time m is doubled. That is, the sample variances of both $\hat{\lambda}$ and $\hat{\alpha}$ are inversely proportional to m. The same appears to be true with respect to n in Tables 7-10 for all the variances given for $\hat{\lambda}$ and in Tables 11-14 for the variances of $\hat{\alpha}$ as n progresses from 8 to 16. Moreover, increases in n up to $n = 8$ also decrease the sample variances of $\hat{\alpha}$ somewhat, but not proportionally.

Tables 7-14 also indicate the effect of changes in m and n on the bias of the estimates. The averages $\bar{\hat{\lambda}}$ given in Tables 7-10 are predominantly positively biased, as would be expected on the basis of Section 4.3. In fact, only fifteen of the 120 averages reported in Tables 7-10 are negatively biased. Furthermore, in each table the bias is greatest for small m and n and it tends to decrease with increasing m and n. The same trend is noticeable in the averages $\bar{\hat{\alpha}}$ recorded in Tables 11-14, where only twelve of the 120 averages are negatively biased. This is not surprising, for a positive bias in $\hat{\lambda}$ makes exp. $(-\hat{\lambda}\, t_i)$ negatively biased and can be compensated for by a positive bias in the corresponding $\hat{\alpha}$.

Now that we have investigated the effects of changes in m and n on the estimates for Case 1, let us analyze Tables 7-14 with respect to changes in λ. Comparisons among these tables show that the positive bias of $\hat{\lambda}$ becomes more pronounced as λ increases. The same is true to a lesser extent for $\hat{\alpha}$, even though the sample variances for $\hat{\alpha}$ tend to decrease as λ increases, at least until $\lambda = \frac{3}{4} \ln 2$, where in some instances the downward trend is reversed. On the other hand, increasing λ under the conditions of our study increases the sample variance of $\hat{\lambda}$ in every instance. However, of more interest than changes in the actual bias and sample variance of $\hat{\lambda}$ with increasing λ are the effects on both the bias and the standard deviation of $\hat{\lambda}$ relative to λ. Table 15, which has been computed from Table 8, indicates the magnitude of these statistics for our study. By constructing similar tables from Tables 7, 9 and 10 it can be shown that relative bias of $\hat{\lambda}$ is reduced for small m and n as λ increases. But as m and n become large, the relative bias of $\hat{\lambda}$ decreases more rapidly for small λ than for large λ. Furthermore, the standard deviation of $\hat{\lambda}$ relative to λ decreases as λ becomes larger.

If λ is increased, $\mathcal{E}(y_{ij}) = \alpha$ exp. $(-\lambda\, t_i)$ becomes smaller, and the e_{ij} as computed in our sampling study also become smaller. Thus we might expect the variation in $\hat{\lambda}$ not only to be less relative to λ for large λ than for small λ, but to also be less in absolute

Table 15

STANDARD DEVIATION AND BIAS OF $\hat{\lambda}$ RELATIVE TO λ

Case 1: $\lambda = \frac{1}{2}\ln 2 = .34657$

m		n = 2	n = 4	n = 6	n = 8	n = 16
1	bias/λ	.165	.063	.036	.048	.008
	s.d./λ	.636	.300	.169	.148	.062
2	bias/λ	.111	.023	.013	.017	.003
	s.d./λ	.350	.131	.079	.063	.030
4	bias/λ	.017	.004	.001	.007	.001
	s.d./λ	.130	.057	.039	.030	.016
8	bias/λ	.015	.005	.002	.008	.003
	s.d./λ	.061	.031	.021	.015	.007
16	bias/λ	-.003	-.003	-.003	.000	.003
	s.d./λ	.030	.015	.010	.007	.004
32	bias/λ	-.001	-.000	.004	.000	-.002
	s.d./λ	.014	.007	.004	.004	.002

value. However, an explanation of the actual behavior of $\hat{\lambda}$ as λ increases may be found by considering the relationship of λ and $T = Kn$. The calculations summarized in Tables 7-14 were all done with t_i ranging from 0 to 4, that is, with $T = 2$. Since $\mathcal{E}(y_{ij}) \longrightarrow 0$ more rapidly with increasing i for a large λ than for a smaller λ, an increase in λ when T is kept constant tends to make more observations nearly zero and of little use in the estimation of λ, which is essentially a rate of decline. Hence, the poorer estimation observed for the larger values of λ in Tables 7-10 may be caused by a failure to reduce T as λ is increased. To investigate this possibility further, let us study the effect of changes in T on $\hat{\lambda}$ and $\hat{\alpha}$ when both λ and n are held fixed.

Suppose in progressing from Table 11 to Table 14 we regard the increase in λ instead as an increase in $T = Kn$ without changing either the product λK or $\mathcal{E}(y_{ij})$. The y_{ij} will not be changed by this interpretation because of the way in which the e_{ij} are generated in our empirical study. Therefore, neither will $\bar{\hat{\alpha}}$ nor the variance of $\hat{\alpha}$ be changed for any given pair of m and n values, for the calculation of $\hat{\alpha}$ from the y_{ij} involves neither λ, K nor T. Thus Tables 11-14 are not changed by the new interpretation. The same is not true of Tables 7-10, however, for $\hat{\lambda}$ is inversely proportional to K, making the variance of $\hat{\lambda}$ inversely proportional to K^2. Thus from Tables 7-10 we could construct new tables for different values of either K or T with λ and n held constant. A few entries which would appear in such tables are given in Table 16.

Table 16

SAMPLE MEANS AND VARIANCES OF $\hat{\lambda}$ FOR VARYING T

Case 1: $\lambda = \frac{1}{4} \ln 2 = .17329$

m	n		T = 2	T = 4	T = 6	T = 8
2	8	$\bar{\hat{\lambda}}$	.17695	.17619	.17586	.17826
		$V(\hat{\lambda})$	.00528	.00188	.00129	.00137
4	16	$\bar{\hat{\lambda}}$	.17306	.17349	.17400	.17420
		$V(\hat{\lambda})$	.00153	.00047	.00028	.00026
8	4	$\bar{\hat{\lambda}}$	.17423	.17409	.17449	.17356
		$V(\hat{\lambda})$	.00280	.00092	.00057	.00054
16	2	$\bar{\hat{\lambda}}$	.17445	.17280	.17373	.17478
		$V(\hat{\lambda})$	.00287	.00091	.00056	.00055
32	6	$\bar{\hat{\lambda}}$	.17298	.17403	.17350	.17376
		$V(\hat{\lambda})$	.00050	.00014	.00009	.00009

The behavior of $\bar{\hat{\lambda}}$ and $V(\hat{\lambda})$ in Table 16 as T increases is similar to that of $\bar{\hat{\alpha}}$ and $V(\hat{\alpha})$ in Tables 11-14 as either T or λ increases. From these tables we conclude that if m and n are held constant, an increase in T tends to reduce the sample variance of both $\hat{\lambda}$ and $\hat{\alpha}$ up to a certain point, after which it appears that at least the variance of $\hat{\alpha}$ increases and that the bias of both $\hat{\lambda}$ and $\hat{\alpha}$ increases. But in actual experimentation, an increase in the range of the t_i is usually accomplished by increasing n without keeping T constant. As we have seen, increasing n with T constant tends to decrease the sampling variances, but it is subject to diminishing returns as n becomes larger. And when making T larger is accompanied by increasing n, we would still eventually expect poorer estimation with the new procedure. Hence the results in Table 7-16 indicate that, for instance, continuing to observe half lives of a decaying radioactive substance will yield better estimates for a Case 1 model at first, but only until the observations level off near zero.

In addition to studying the effects of changes in m, n, λ and T on $\hat{\lambda}$, $\hat{\alpha}$ and their variances, let us compare some of the small sample variances reported in this section with the corresponding asymptotic variances given in Section 5.4. Table 17 presents several pairs of variances, and in the calculation of the asymptotic variance in each pair σ^2 was computed in the manner prescribed for our empirical study. Note the close agreement of the small sample variances with the respective asymptotic variances.

Table 17

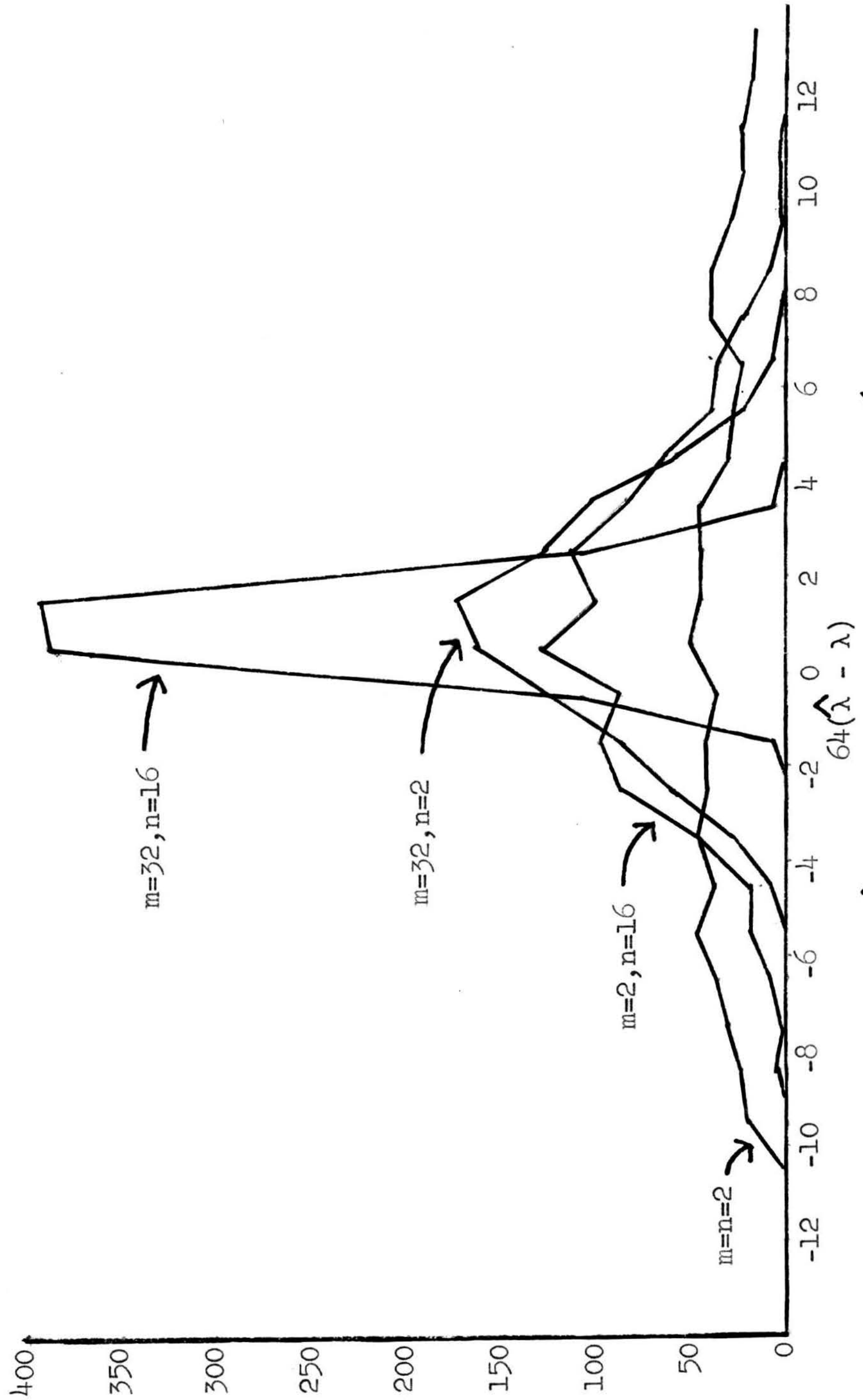
A COMPARISON OF ASYMPTOTIC AND SMALL SAMPLE VARIANCES

$m = 1, T = 2, \lambda = \ln 2$

$\sigma = k' \text{ mean } \mathcal{E}(y_{ij})$

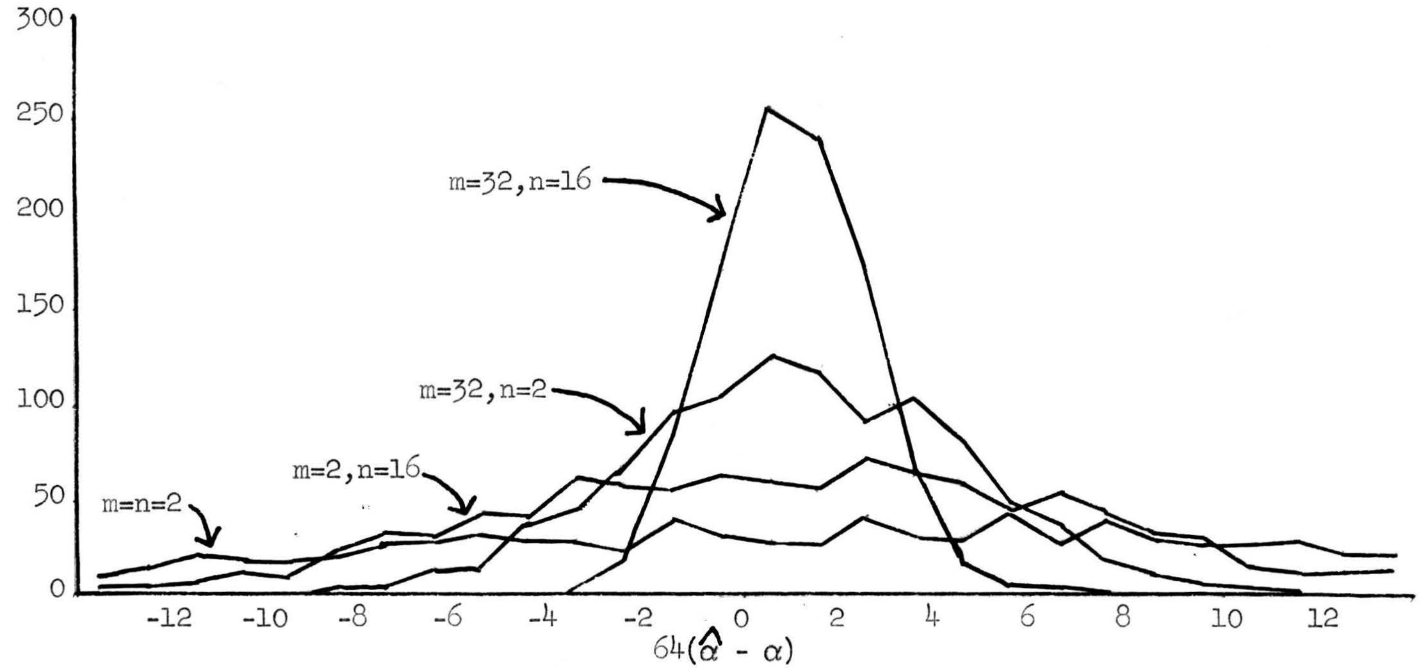
k'		n = 2 $V(\hat{\alpha})$	$V(\hat{\lambda})$	n = 4 $V(\hat{\alpha})$	$V(\hat{\lambda})$	n = 8 $V(\hat{\alpha})$	$V(\hat{\lambda})$
.2	Asymptotic	.01411	.03320	.01087	.01660	.00679	.00830
	Small Sample	.01485	.05105	.01193	.01929	.00757	.00921
.14142	Asymptotic	.00706	.01660	.00544	.00830	.00340	.00415
	Small Sample	.00739	.01945	.00552	.00856	.00358	.00433
.1	Asymptotic	.00353	.00830	.00272	.00415	.00170	.00208
	Small Sample	.00370	.00878	.00272	.00431	.00168	.00203

The sample means and variances of Tables 7-17 have helped describe the small sample distributions of $\hat{\lambda}$ and $\hat{\alpha}$ for Case 1. To further depict these distributions, the sampling distributions of $\hat{\lambda}$ about λ and of $\hat{\alpha}$ about $\alpha = 1$ were recorded for each set of parameters. For both estimates intervals of 1/64 were used with sixteen intervals on each side of the parameter in question. It would not be feasible to present each of these distributions here, but some of them are given in Figures 1 and 2 to illustrate the effect of increasing m and n on these distributions. These figures indicate the approach of these distributions to normality as either m or n grows large, a result demonstrated analytically in Chapter IV.



DISTRIBUTION OF $(\hat{\lambda} - \lambda)$ FOR CASE 1 WITH $\lambda = 1/4 \ln 2 = 0.17329$

Figure 1



DISTRIBUTION OF $(\hat{\alpha} - \alpha)$ FOR CASE 1 WITH $\alpha = 1$

Figure 2

To complete our discussion of the empirical sampling study for Case 1, we need only look at the proportions of inadmissible solutions obtained during the calculations. For Case 1, an inadmissible solution occurs when $Y_2 \geq Y_1$, as mentioned in Section 5.6. Also, as stated before, the inadmissible solutions reported here occurred while 1024 admissible solutions were being computed for each set of parameters. Table 18 includes all the inadmissible solutions yielded by the empirical work for Case 1. The results given there are encouraging because in only one instance is there an entry for $\sigma < 0.14142$ times the mean expectation of y_{ij}. However, the table indicates that under our Case 1 model when the e_{ij} are normally distributed, the new estimation procedure will produce inadmissible solutions with a fairly high frequency when λ is small relative to T.

The empirical sampling study for Case 1 discussed in this section has reflected favorably upon the new estimation procedure. Yet from this study we cannot infer that the new method behaves as well for more general cases of our model. For instance, as the number of terms in the model increases more necessary conditions must be satisfied in order for a solution to be admissible, so we would expect a higher frequency of inadmissible solutions. However, this sampling study does indicate that the new procedure is adequate when its model is applicable.

Table 18

PROPORTIONS OF INADMISSIBLE SOLUTIONS FOR CASE 1

m	n	$\lambda = \frac{1}{4}\ln 2$	$\lambda = \frac{1}{2}\ln 2$	$\lambda = \frac{3}{4}\ln 2$	$\lambda = \ln 2$
1	2	.192	.038	.011	.002
	4	.116	.015	0	0
	6	.064	.002	0	0
	8	.057	.001	0	0
	16	.003	0	0	0
2	2	.110	.015	0	0
	4	.038	0	0	0
	6	.019	0	0	0
	8	.008	0	0	0
4	2	.051	0	0	0
	4	.007	0	0	0
	6	.003	0	0	0
8	2	.006	0	0	0

VII. EXTENSIONS AND ILLUSTRATIONS

7.1 Extensions of the Model

The model in Section 3.1 was formulated so that it could be realistically applied to many problems involving exponential fitting, and yet it was restricted sufficiently to make the development of the new estimation procedure relatively simple. However, there are several useful extensions of the model which require only minor alterations of the estimation procedure. Some of these are indicated in Section 3.1, where several assumptions of the model are declared unnecessary as far as the estimation itself is concerned, but either are necessary in order for certain properties of the estimators to hold or else are necessary to make the model conform to the experimental situations to which it is most often applied. In this section some additional extensions of the model will be proposed.

The first extension results from removing the requirement that all the λ_k be real. As pointed out by Willers [31] , there are some situations in which complex exponents are meaningful. Then the model fitted can conveniently be represented in terms of sine and cosine terms as well as exponentials, thus giving a new model to which the new estimation procedure applies. Another trivial modification of the model consists of using a positive number other than e as the base of all the exponentials and logarithms in this paper.

As stated in the introduction, the estimation can also be carried out when the e_{ij} are not homogeneous. And when σ^2 varies only from group to group, the limiting distributions of Chapter IV are still valid if a slight change is made. In this instance, the expression (4.1.8) for the asymptotic variance of an estimator for n large becomes

$$V(\hat{\theta}) = \sum_q a_q^2 \sigma_q^2 \quad , \tag{7.1.1}$$

where σ_q^2 is the variance of the observations in the q^{th} group, and (4.3.8) is similarly affected. When estimates s_q^2 are substituted for the σ_q^2 in (7.1.1) and when all of the s_q^2 have the same number of degrees of freedom, we may assign that number of degrees of freedom to $V(\hat{\theta})$. But the developments of the next paragraph will make it possible for the s_q^2 to have different numbers of degrees of freedom in accordance with an extension of our model. In this case, we shall be conservative and assign to $V(\hat{\theta})$ the smallest of these degrees of freedom.

The new procedure can be further extended to a model in which an unequal number of observations are made at some of the points t_i. In fact, it is only necessary that the number of observations be the same for all of the t_i within any given group. Suppose in the procedure as developed in Chapter III we let j range from 1 to m_q, where the subscript q denotes one of the $2p$ groups as before. Then, if we replace each S_q by S_q/m_q and let $m = 1$, the development in Chapter III applies to this formulation. Although this extension increases the applicability of the new method, its use also requires more care in the planning of an experiment. Previously, if an experimenter wanted to use the new procedure without introducing any approximations, he had to take observations at a prescribed number of points. Yet, for instance, if he took observations at twelve points, he could attempt to fit either 2 , 3 , 4 , 6 or 12 exponential terms with these observations. But if he utilized the extension of the model presented in this paragraph, he would be restricted further in choosing the number of terms to be fitted.

The final modifications which we will suggest here concern the constant α_0 in the Case 6 model. Instead of remaining constant throughout an experiment, α_0 could for instance be a function of t_i . In most experimental situations this would mean that α_0 varies with time and would lead to the model

$$y_{ij} = \alpha_{0i} + \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i} + e_{ij} .$$

Unless α_{0i} were a periodic function with period n , this model would complicate the new procedure considerably. Yet a solution appears to be feasible in some instances.

Instead of varying with i , α_0 could reflect, say, block effects when an experiment is carried out with several animals. That is, α_0 could be a function of j , making

$$y_{ij} = \alpha_{0j} + \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t_i} + e_{ij} . .$$

This formulation would not change the estimation appreciably since the constant term $\sum_{j=1}^{m} \alpha_{0j}$ would be included in every sum S_q. Another possible extension along this line would be the inclusion of block, time and interaction effects in the model. This would of course unduly complicate the estimation with the new procedure except perhaps in special cases.

7.2 Approximations to the Model

In order to utilize the new procedure it is often necessary to make approximations. Sometimes the t_i are not spaced in accordance with the model or else they cannot be divided evenly into $2p$ groups. In some such situations a few interpolations or extrapolations will supply the missing data and make it possible to apply the new method. This is done in the examples in Sections 7.4 and 7.5. Also, an experimenter sometimes takes his data at unequal intervals in such a way that when the t_i are divided into $2p$ groups of length T, n varies from group to group. In this instance a sum S_q may be formed as usual for each group. Then a solution may be carried out as if m_q instead of n changed between groups without altering the product $m_q n$ for any group. This latter approximation is rather crude, but interpolations and extrapolations such as those suggested at first often do not weaken the estimation if they are few in number relative to the number of observations.

A useful approximation is also available when m varies not only from group to group, but for t_i within the same group. In Section 7.1 we saw that, when m changes only from group to group, S_q/m_q may be substituted for S_q in the solution in Chapter III with $m = 1$. Similarly, as an approximation we may average the observations y_{ij} for each point t_i when m varies within a group. Then these averages may be substituted for the y_{ij} in the new procedure with $m = 1$. In this instance, the variance σ^2 in the basic model is a variance of means and varies from one observation point to another, and this should be taken into account in the computation of s^2.

7.3 An Illustration for Case 1

In this section the new estimation procedure will be used to fit a Case 1 model to the data from an experiment conducted by Paul Urso in the Biology Division of the Oak Ridge National Laboratory. Mr. Urso made nucleated bone marrow cell counts on mice both before and after X-irradiation of 900 roentgens. These counts are reported in Table 19, with those made before irradiation recorded for zero days after irradiation.

After plotting the averages given in the last row of Table 19, the experimenter suggested that the data be fitted to a single exponential, that is, to a Case 1 model. Since this entails the estimation of only two parameters by the new procedure, we shall partition the data of Table 19 into two groups, with the counts for days 0 and 1 in the first group and those for days 2 and 3 in the second, and hence $n = 2$. Also, since the interval between successive series of counts is one day in each instance, $K = 1$. The number of mice for which counts were made varies within each of the groups, so we shall follow the recommendation made in Section 7.2 of replacing the observed counts y_{ij} for any i, that is, for any day, by the average y_{ij} for that day and by letting $m = 1$ in the estimation equations. The average daily counts are given in the last row of Table 19. From these averages, using

Table 19

BONE MARROW CELL COUNTS OF X-IRRADIATED MICE

	Days after X-Irradiation			
	0	1	2	3
	11,137,500	3,062,000	437,000	96,250
	9,418,750	3,075,000	766,666	112,500
	10,287,500	5,050,000	1,087,500	237,500
	12,487,500	3,312,500	368,750	75,000
Bone	11,700,000	2,775,000	1,206,250	150,000
Marrow	10,023,750	1,058,750	500,000	90,000
Counts	12,062,500	2,000,000	85,000	100,000
	10,437,500	3,475,000	416,666	118,750
		2,675,000	450,000	162,500
			737,500	
			281,250	
			756,250	
Averages	10,944,375.0	2,942,583.3	591,111.0	126,944.0

(3.5.1) and (3.5.2), we compute

$$S_1 = 10,944,375.0 + 2,942,583.3 = 13.8870 \times 10^6 \quad ,$$

$$S_2 = 591,111.0 + 126,944.0 = 0.7181 \times 10^6 \quad .$$

Then (3.5.3), (3.5.4) and (3.5.5) yield

$$\hat{x} = \frac{0.7181}{13.8870} = 0.05171 \quad ,$$

$$\hat{\lambda} = -\frac{1}{2}\ln 0.05171 = 1.4811 \quad ,$$

$$\hat{\alpha} = \frac{(1 - 0.22739)\,(13.8870)^2 \times 10^{12}}{13.1689 \times 10^6} = 11.3143 \times 10^6 \; .$$

Note that in the calculation of $\hat{\alpha}$ for this example

$$\hat{x}^{\frac{1}{n}} = \exp.\left[-\left(-\frac{1}{2}\ln \hat{x}\right)\right] = \exp.\ (-\hat{\lambda})$$

and can be found merely by looking up $\hat{\lambda}$ in tables of the negative exponential function.

Now we may represent the data from Mr. Urso's experiment by the estimation equation

$$\hat{y}_{ij} = (11.3143 \times 10^6)\, e^{-1.4811\, t_i} \quad .$$

From this equation the following predicted values may be computed:

$$\hat{y}_0 = 11.3143 \times 10^6 \quad ,$$

$$\hat{y}_1 = 2.5728 \times 10^6 \quad ,$$

$$\hat{y}_2 = 0.5850 \times 10^6 \quad ,$$

$$\hat{y}_3 = 0.1330 \times 10^6 \quad .$$

It is a general feature of the new procedure that sums $\hat{S}_q$ calculated from the y_{ij} equal the S_q computed earlier from the experimental data. So, as a check on our computations, we compute

$$\hat{S}_1 = (11.3143 + 2.5728) \times 10^6 = 13.8871 \times 10^6 \quad ,$$

$$\hat{S}_2 = (0.5850 + 0.1330) \times 10^6 = 0.7180 \times 10^6 \quad ,$$

and note that $\hat{S}_q = S_q$, $q = 1, 2,$ within the limits of rounding errors. Then, as a measure of goodness of fit, we calculate

$$\sum_{i=0}^{3} (\hat{y}_i - \bar{y}_i)^2 = \left[(.3699)^2 + (.3698)^2 + (.0061)^2 + (.0061)^2\right] \times 10^{12}$$

$$= 0.2737 \times 10^{12} \quad . \qquad (7.3.1)$$

This sum of squared deviations of the means from the regression will be compared later with similar sums computed for other methods of estimation.

Next we should like to estimate the variances of $\hat{\lambda}$ and $\hat{\alpha}$, but in order to do this we must first compute an estimate s^2 of the variance of the average count per day. Using (6.1.4), we compute sample variances of counts within each day to obtain

$$1157.599 \times 10^9 \;,\quad 1178.785 \times 10^9 \;,\quad 108.0151 \times 10^9 \;,\quad 2.5032 \times 10^9$$

for days 0, 1, 2 and 3 respectively. The corresponding sample variances for the averages are

$$s_0^2 = 144.700 \times 10^9 \;,$$

$$s_1^2 = 130.976 \times 10^9 \;,$$

$$s_2^2 = 9.001 \times 10^9 \;,$$

$$s_3^2 = .278 \times 10^9 \;,$$

with 7, 8, 11 and 8 degrees of freedom respectively. All of the latter sample variances may be pooled to form

$$s^2 = 63.587 \times 10^9 \qquad (7.3.2)$$

with 34 degrees of freedom. However, s_2^2 and especially s_3^2 are quite a bit smaller than s_0^2 and s_1^2. Therefore, it would seem reasonable to assume that σ^2 is homogeneous only for y_{ij} in the same group and to compute asymptotic variances for $\hat{\alpha}$ and $\hat{\lambda}$ as indicated by (7.1.1). Further support for this approach is furnished by the deviations from regression used in the calculation of (7.3.1). For an estimate of σ^2 from the first group we pool s_0^2 and s_1^2 to obtain

$$s^2(1) = 137.381 \times 10^9 \tag{7.3.3}$$

with fifteen degrees of freedom while for the second group from s_2^2 and s_3^2 we calculate

$$s^2(2) = 5.328 \times 10^9 \tag{7.3.4}$$

with nineteen degrees of freedom. Variance estimates for $\hat{\lambda}$ and $\hat{\alpha}$ will be computed using both the estimate (7.3.2) and the estimates (7.3.3) and (7.3.4).

Now let us estimate the asymptotic variance of $\hat{\lambda}$ by equation (4.2.15). For η_q, $q = 1, 2$, we take the mean $Y_q = S_q/mn$. Hence in this example we let

$$\eta_1 = \frac{1}{2}(13.8870 \times 10^6) = 6.9435 \times 10^6 ,$$

$$\eta_2 = \frac{1}{2}(0.7180 \times 10^6) = 0.3590 \times 10^6 .$$

Substituting these estimates of the η_q along with the estimate (7.3.2) of σ^2 in (4.2.15) and dividing by mn , since we want the variance of $\hat{\lambda}$ instead of $\sqrt{mn}\,\hat{\lambda}$, we obtain

$$V(\hat{\lambda}) = 0.06183 \quad .$$

Similarly, if we use (7.3.3) and (7.3.4) to estimate σ_q^2 , $q = 1, 2$, substitution in (7.1.1) yields

$$V(\hat{\lambda}) = 0.00552 \quad .$$

Taking the square roots of these variances, we compute the standard deviations

$$\text{s.d.}(\hat{\lambda}) = 0.2487$$

when s^2 is calculated from all the observations and

$$\text{s.d.}(\hat{\lambda}) = 0.0743$$

when σ^2 is estimated separately for each group.

To compute asymptotic variances for $\hat{\alpha}$ we first evaluate the corresponding a_q , defined by (4.1.9), by substitution in (4.2.16). For this example,

$$a_1 = 1.7804 \quad , \quad a_2 = 2.9192 \quad .$$

Then, using the overall estimate (7.3.2) of σ^2 , we have

$$V(\hat{\alpha}) = \frac{1}{mn}\,(a_1^2 + a_2^2)s^2 = 371.717 \times 10^9 \quad .$$

On the other hand, the estimates (7.3.3) and (7.3.4) lead to

$$V(\hat{\alpha}) = \frac{1}{mn}\left[a_1^2 s^2(1) + a_2^2 s^2(2)\right] = 240.437 \times 10^9 .$$

The corresponding estimates of the standard deviation of $\hat{\alpha}$ are 0.6097×10^6 and 0.4903×10^6 respectively.

The standard deviations we have just computed are those that would be calculated by substitution for $s(\frac{1}{mn} \sum_q A_q^2)^{1/2}$ in (6.1.12). Hence we may quickly apply (6.1.12) to obtain approximate confidence intervals for λ and α. Using the five per cent level of Student's t-statistic, which for 34 degrees of freedom is 2.032, and using standard deviations calculated with s^2 computed from all the data, we compute the 95 per cent confidence intervals

$$0.9757 \leq \lambda \leq 1.9865 ,$$

$$10.0754 \times 10^6 \leq \alpha \leq 12.5532 \times 10^6 .$$

The corresponding confidence intervals computed using group estimates of σ^2 and the five per cent level of Student's t-statistic with 15 degrees of freedom, as recommended in Section 7.1, are

$$1.3228 \leq \lambda \leq 1.6394 ,$$

$$10.2695 \times 10^6 \leq \alpha \leq 12.3591 \times 10^6 .$$

These sets of confidence intervals illustrate the need for making realistic assumptions about σ^2, that is, for not assuming that σ^2 is homogeneous throughout an experiment when in fact it varies from group to group.

The data in Table 19 afford us an opportunity to compare the new procedure with the other methods of fitting mentioned in Chapter II. An application of Prony's method, as given in [30], to the averages displayed in Table 19 leads to the estimates 1.3314 and 9.3622×10^6 for $\hat{\lambda}$ and $\hat{\alpha}$ respectively. These estimates in turn yield

$$\sum_{i=0}^{3} (\hat{y}_i - \bar{y}_i)^2 = 2.7299 \times 10^{12} \quad .$$

The "peeling off" procedure, which for Case 1 reduces to fitting the logarithms of the observations to a straight line by least squares, gives

$$\hat{\lambda} = 1.5291 \quad , \quad \hat{\alpha} = 11.3500 \times 10^6 \quad ,$$

$$\sum_{i=0}^{3} (\hat{y}_i - \bar{y}_i)^2 = 0.4009 \times 10^{12} \quad .$$

Finally, the iterative Deming procedure leads to least squares estimates

$$\hat{\lambda} = 1.3530 \quad , \quad \hat{\alpha} = 10.9609 \times 10^6 \quad ,$$

with

$$\sum_{i=0}^{3} (\hat{y}_i - \bar{y}_i)^2 = 0.0361 \times 10^{12} \quad .$$

These would also be the maximum likelihood estimates under assumptions of normality. Note that the sum of squared deviations of the means from the regression given earlier by (7.3.1) for the new procedure is, next to that for the Deming least squares method, the smallest of those computed in this section.

7.4 An Illustration for Case 4

The new estimation procedure has also been used to analyze the data from some physics experiments at the Oak Ridge National Laboratory. In one of these experiments Dr. Marvin Slater placed cylinders of paraffin between a neutron source and a polyethylene-ethylene proportional counter and then he recorded the amount of radiation transmitted to the counter through paraffin cylinders of different lengths. The counts he made are reported in Table 20.

Table 20

Neutron Counts for Different Lengths t_i of a Paraffin Cylinder

t_i	0	2	4	8	12	16
Counts	67.9	36.3	17.2	8.2	3.5	2.8

In this experiment there are two ways in which radiation can be transmitted to the counter. One is directly through the paraffin and would be expected to be exponentially related to the length of the paraffin cylinder. Scattered radiation reflected from the walls and other surroundings would also reach the counter and would be expected to be constant. Therefore, a Case 4 regression was suggested for the data in Table 20. But this data does not satisfy all the requirements of our Case 4 model, for all the increments in the lengths of the paraffin cylinder are not equal. Two approximate solutions were tried however. In one the count for $t_i = 2$ was discarded and a value for $t_i = 20$ of 2.6 was extrapolated from a plot of the logarithms of the data so that the number of t_i would be an integral multiple of three, the number of parameters. Thus, in this solution, $K = 4$ and $n = 2$. Since interpolation is more apt to be accurate than is extrapolation, interpolated counts for $t_i = 6$, 10 and 14 were used in another solution with $K = 2$ and $n = 3$. The first of these alternative estimations resulted in the smaller sum of squares of deviations from regression of the six original observations, and therefore we shall present that estimation here.

The first step in the estimation, with $m = p = 1$, $n = 2$ and $K = 4$, is to compute

$$S_1 = 67.9 + 17.2 = 85.1 ,$$

$$S_2 = 8.2 + 3.5 = 11.7 ,$$

$$S_3 = 2.8 + 2.6 = 5.4$$

from (3.6.2). Then (3.6.12), (3.6.13), (3.6.14) and (3.6.15) yield

$$\hat{x} = \frac{(11.7 - 5.4)}{(85.1 - 11.7)} = 0.085831 ,$$

$$\hat{\lambda} = -\frac{1}{8} \ln\ 0.085831 = 0.30692\ ,$$

$$\hat{\alpha}_0 = \frac{(85.1)(5.4) - (11.7)^2}{2\left[85.1 - 2(11.7) + 5.4\right]} = 2.404,$$

$$\hat{\alpha}_1 = \frac{(85.1 - 11.7)^3 \left[1 - (0.085831)^{1/2}\right]}{\left[85.1 - 2(11.7) + 5.4\right]^2} = 62.099.$$

Hence the regression equation is

$$\hat{y}_i = 2.404 + 62.099\ e^{-0.30692\ t_i}$$

and

$$\hat{y}_0 = 64.50,\ \hat{y}_2 = 36.02,\ \hat{y}_4 = 20.60,\ \hat{y}_8 = 7.73,$$

$$\hat{y}_{12} = 3.97,\ \hat{y}_{16} = 2.86,\ \hat{y}_{20} = 2.54.$$

In order to check our calculations, we use the y_i to compute $\hat{S}_1 = 85.10$, $\hat{S}_2 = 11.70$ and $\hat{S}_3 = 5.40$ which agree exactly with the S_q computed earlier. Also, we find that the sum of squares of deviations from the regression of the observed y_i , including y_2 , is 23.6438.

Next let us estimate σ^2 , the variance of the counts in Table 20 under our Case 4 model. Since only one observation was taken for each t_i, we must use the mean square deviation of the observations y_i from the corresponding regression values, y_i , to estimate σ^2 . This mean square, calculated using both y_2 and y_{20}, equals 7.8825 and has three degrees of freedom since there are three parameters to be estimated with six original observations.

Variance estimates for $\hat{\lambda}$, $\hat{\alpha}_0$ and $\hat{\alpha}_1$, may now be computed from the asymptotic variance formulas given in Section 4.2. For $\hat{\lambda}$, substitution of $S_q/mn = S_q/2$ for η_q , $q = 1, 2, 3$, gives

$$a_1 = 0.00341, \quad a_2 = -.04309, \quad a_3 = .03968.$$

Then substituting these values and $s^2 = 7.8825$ in equation (4.1.8), which gives the asymptotic variance of $\sqrt{mn}\,(\hat{\lambda} - \lambda)$, and dividing by $mn = 2$, we compute

$$V(\hat{\lambda}) = 0.013566.$$

Similarly, (4.2.19) and (4.2.20) in conjunction with (4.1.8) yield

$$V(\hat{\alpha}_0) = 5.80988, \quad V(\hat{\alpha}_1) = 29.80689.$$

From these variances we calculated the standard deviations

$$s.d.(\hat{\lambda}) = 0.11647, \quad s.d.(\hat{\alpha}_0) = 2.41037, \quad s.d.(\hat{\lambda}_1) = 5.45957.$$

As in the example in Section 7.3, these standard deviations may be substituted in (6.1.13) to obtain approximate confidence limits. The 95 per cent confidence limits computed in this way are

$$-0.06369 \leq \lambda \leq 0.67753$$
$$-5.266 \leq \alpha_0 \leq 10.074$$
$$44.727 \leq \alpha_1 \leq 79.471.$$

These limits are too wide to be of any use whatsoever. Moreover, computation of 95 per cent confidence limits for λ using (6.1.9) and (6.1.10) gives

$$0.01622 \leq \lambda \leq \infty .$$

Inspection of equation (6.1.9) shows that the upper confidence limit for λ computed from that equation will usually be ∞ when s^2 is large. In fact, all the extremely wide confidence limits calculated in this section result from an inordinately large estimate of σ^2.

To complete this illustration, let us again present estimates calculated by some other estimation methods. To apply Prony's method to a Case 4 model we first form $y_i' = y_i - y_{i+1}$, $i = 0, 1, 2, ---, 3n-2$. The y_i' would be expected to follow a Case 1 model and may be fitted by the Prony method as outlined in Chapter II. Then the Case 4 estimates for the y_i may be computed from the Case 1 results obtained for the y_i'. This extension of the Prony method to Case 4 is similar to that given for the new procedure in Section 3.6. The estimates yielded by the Prony method for this example are

$$\hat{\lambda} = 0.41824, \quad \hat{\alpha}_0 = -5.682, \quad \hat{\alpha}_1 = 104.006.$$

In this example $\alpha_0 \geq 0$, and therefore we shall take $\hat{\alpha}_0 = 0$ as the Prony estimate of α_0. To apply the "peeling off" procedure to this illustration we must first assign a value to $\hat{\alpha}_0$. If we let $\hat{\alpha}_0 = 2.404$, the estimate calculated by the new procedure, the estimates of λ and α_1 yielded by the "peeling off" method are

$$\hat{\lambda} = 0.32099, \quad \hat{\alpha}_1 = 62.579.$$

Finally, the third iteration with the Deming method after inserting the new procedure estimates as initial estimates gives corrections -0.00021, -0.076 and 0.01135 which, when added to the products of the second iteration, yield estimates 0.35258, 2.836 and 65.261 respectively for λ, α_0 and α_1. In both this example and in the one presented in Section 7.3 the Prony method apparently gives as good an estimate of λ as the new procedure does, but the new procedure results in more reasonable estimates of the α_k. The sums of squares of the deviations of the y_i from the $\hat{y}_i$ are 16.9220 and 7.7837 for the "peeling off" and Deming procedures respectively. Such a sum of squares was not computed for the Prony method because of its negative estimate of α_0.

7.5 An Illustration for Case 2

As a final illustration of the new estimation procedure we shall apply it to the data in Table 21. The logarithms of frequencies given there describe the distribution of background pulses generated in a proportional counter by neutron interaction with walls and gas plus pulses due to circuit noise. The experiment was conducted by Dr. M. L. Randolph at the Oak Ridge National Laboratory. No counts were made for pulse heights of 14, 26, and 28 and those displayed in Table 21 for these pulse heights were obtained by interpolation. A plot of the data suggests a Case 2 model and therefore that is the model we shall attempt to fit. This example will not be studied as completely as were those in Sections 7.3 and 7.4, but only enough calculations will be carried out to illustrate the general approach given in Chapter IV for the computation of asymptotic variances.

There are sixteen evenly spaced t_i in Table 21 with an interval of two between successive t_i and there are four parameters to be estimated, so $K = 2$ and $n = 4$. Also, only one logarithm is recorded for each t_i, so $m = 1$. It can be shown that substitution of these values along with the data in Table 21 in the estimation equations derived for Case 2 in Section 3.5 yields

$$S_1 = 18.600 \ , \ S_2 = 1.227 \ , \ S_3 = 0.158 \ , \ S_4 = 0.091 \ ;$$

$$\hat{E}_1 = 1.0457 \ , \ \hat{E}_2 = 0.0607 \ ; \ \hat{x}_1 = 0.9840 \ , \ \hat{x}_2 = 0.0617 \ ;$$

$$\hat{\lambda}_1 = 0.00202 \ , \ \hat{\lambda}_2 = 0.3482 \ ; \ \hat{\alpha}_1 = 0.0217 \ , \ \hat{\alpha}_2 = 9.8977 \ .$$

Table 21

Logarithms y_i of Frequencies of Pulse Heights t_i

Generated in a Proportional Counter

t_i	y_i
0	10.430
2	4.703
4	2.327
6	1.140
8	0.615
10	0.325
12	0.170
14	0.117
16	0.050
18	0.040
20	0.046
22	0.022
24	0.036
26	0.021
28	0.018
30	0.016

Let us use these results to estimate the asymptotic variances of $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\alpha}_1$ and $\hat{\alpha}_2$ apart from the estimation of σ^2.

In the general procedure for computing asymptotic variances it is first necessary to evaluate the $\partial\hat{x}_k/\partial Y_q$ given by (4.2.12) in conjunction with (4.2.3), (4.2.4) and (4.2.8) through (4.2.11). From these equations it follows that

$$\frac{\partial}{\partial Y_q}\hat{x}_k = -\frac{\frac{\partial}{\partial Y_q} g(\hat{x}_k , Y)}{g_x(\hat{x}_k , Y)} \quad , \tag{7.5.1}$$

where for Case 2

$$g(\hat{x}_k , Y) = \hat{x}_k^2 - \hat{E}_1 \hat{x}_k + \hat{E}_2 \quad ,$$

$$\frac{\partial}{\partial Y_q} g(\hat{x}_k , Y) = -\hat{x}_k \frac{\partial\hat{E}_1}{\partial Y_q} + \frac{\partial\hat{E}_2}{\partial Y_q} \quad , \tag{7.5.2}$$

$$g_x(\hat{x}_k , Y) = 2\hat{x}_k - \hat{E}_1 \quad . \tag{7.5.3}$$

In all of these equations and in the rest of this section, $k = 1, 2$ and $q = 1, 2, 3, 4$ for Case 2. Using (7.5.3) and the estimates already computed for this example, we find that

$$g_x(\hat{x}_1 , Y) = 0.9223 \ , \quad g_x(\hat{x}_2 , Y) = -0.9223 \quad .$$

But to calculate the $\frac{\partial}{\partial Y_q} g(\hat{x}_k , Y)$, which are necessary for the evaluation of (7.5.1), we must first compute the $\partial\hat{E}_r/\partial Y_q$.

From (3.5.7) and (3.5.8) with $S_q = mn\, Y_q$,

$$\hat{E}_1 = \frac{Y_1Y_4 - Y_2Y_3}{Y_1Y_3 - Y_2^2} = \frac{J_2}{J_1} ,$$

$$\hat{E}_2 = \frac{Y_2Y_4 - Y_3^2}{Y_1Y_3 - Y_2^2} = \frac{J_3}{J_1} ,$$

where we define

$$J_1 = Y_1Y_3 - Y_2^2 \ , \quad J_2 = Y_1Y_4 - Y_2Y_3 \ , \quad J_3 = Y_2Y_4 - Y_3^2 \ .$$

Then differentiation of $\hat{E}_1$ and $\hat{E}_2$ yields

$$\frac{\partial \hat{E}_1}{\partial Y_1} = \frac{J_1Y_4 - J_2Y_3}{J_1^2} \ , \quad \frac{\partial \hat{E}_2}{\partial Y_1} = \frac{- J_3Y_3}{J_1^2} ,$$

$$\frac{\partial \hat{E}_1}{\partial Y_2} = \frac{- J_1Y_3 + 2\, J_2Y_2}{J_1^2} \ , \quad \frac{\partial \hat{E}_2}{\partial Y_2} = \frac{J_1Y_4 + 2\, J_3Y_2}{J_1^2} ,$$

$$\frac{\partial \hat{E}_1}{\partial Y_3} = \frac{- J_1Y_2 - J_2Y_1}{J_1^2} \ , \quad \frac{\partial \hat{E}_2}{\partial Y_3} = \frac{- 2\, J_1Y_3 - J_3Y_1}{J_1^2} ,$$

$$\frac{\partial \hat{E}_1}{\partial Y_4} = \frac{Y_1}{J_1} \ , \qquad \frac{\partial \hat{E}_2}{\partial Y_4} = \frac{Y_2}{J_1} \ .$$

For our example $Y_q = S_q/4$. Therefore,

$$Y_1 = 4.6500 \ , \ Y_2 = 0.3068 \ , \ Y_3 = 0.0395 \ , \ Y_4 = 0.0228 \ ,$$

and

$$J_1 = 0.08955 \ , \ J_2 = 0.09390 \ , \ J_3 = 0.005435 \ .$$

From these Y_q and J_r values we compute

$$\frac{\partial \hat{E}_1}{\partial Y_1} = -0.2079 \ , \ \frac{\partial \hat{E}_2}{\partial Y_1} = -0.02677 \ ,$$

$$\frac{\partial \hat{E}_1}{\partial Y_2} = 6.7444 \ , \ \frac{\partial \hat{E}_2}{\partial Y_2} = 0.6705 \ ,$$

$$\frac{\partial \hat{E}_1}{\partial Y_3} = -3.9701 \ , \ \frac{\partial \hat{E}_2}{\partial Y_3} = -4.0334 \ ,$$

$$\frac{\partial \hat{E}_1}{\partial Y_4} = 51.9263 \ , \ \frac{\partial \hat{E}_2}{\partial Y_4} = 3.4260 \ .$$

Now we have everything at hand to compute the $\partial g(\hat{x}_k , Y)/\partial Y_q$ as given by (7.5.2). Then substitution of these quantities and the $g_x(\hat{x}_k , Y)$ previously computed into (7.5.1) for our example enables us to compute

$$\frac{\partial \hat{x}_1}{\partial Y_1} = -0.1928 \ , \ \frac{\partial \hat{x}_2}{\partial Y_1} = -0.01513 \ ,$$

$$\frac{\partial \hat{x}_1}{\partial Y_2} = 6.4686 \ , \quad \frac{\partial \hat{x}_2}{\partial Y_2} = 0.2759 \ ,$$

$$\frac{\partial \hat{x}_1}{\partial Y_3} = 0.1375 \ , \quad \frac{\partial \hat{x}_2}{\partial Y_3} = -4.1077 \ ,$$

$$\frac{\partial \hat{x}_1}{\partial Y_4} = 51.6855 \ , \quad \frac{\partial \hat{x}_2}{\partial Y_4} = 0.2420 \ .$$

For the $\hat{\lambda}_k$, the a_q of Theorem 1 as defined by (4.1.9) may now be calculated by substitution in equation (4.2.13). In this way for $\hat{\lambda}_1$ we compute

$$a_1 = 0.02449 \ , \ a_2 = -0.8217 \ , \ a_3 = 0.01747 \ , \ a_4 = -6.5657 \ ,$$

and for $\hat{\lambda}_2$ we calculate

$$a_1 = 0.03066 \ , \ a_2 = -0.5591 \ , \ a_3 = 8.3246 \ , \ a_4 = -0.4904 \ .$$

Then substituting in (4.1.8) and dividing by $mn = 4$ since we want the variance of the $\hat{\lambda}_k$ instead of the $\sqrt{mn}(\hat{\lambda}_k - \lambda_k)$, we compute

$$V(\hat{\lambda}_1) = 10.9461 \ s^2 \ , \quad V(\hat{\lambda}_2) = 17.4633 \ s^2 \ .$$

Now let us replace each S_q in (3.5.11) and (3.5.12) by the corresponding $mn \ Y_q$, thus representing the $\hat{\alpha}_k$ as

$$\hat{\alpha}_1 = \frac{n(1 - \hat{x}_1^{\frac{1}{n}})(Y_1\hat{x}_2 - Y_2)}{(1 - \hat{x}_1)(\hat{x}_2 - \hat{x}_1)} \quad ,$$

$$\hat{\alpha}_2 = \frac{n(1 - \hat{x}_2^{\frac{1}{n}})(Y_2 - Y_1\hat{x}_1)}{(1 - \hat{x}_2)(\hat{x}_2 - \hat{x}_1)} \quad .$$

Then by differentiating the expression for the $\hat{\alpha}_1$ and by referring to (4.1.9), we find that for $\hat{\alpha}_1$,

$$a_q = \frac{\partial\hat{\alpha}_1}{\partial Y_q} = n\,(1 - \hat{x}_1)^{-2}\,(\hat{x}_2 - \hat{x}_1)^{-2} \Big\{ (1 - \hat{x}_1)(\hat{x}_2 - \hat{x}_1) \Big[(1 - \hat{x}_1^{\frac{1}{n}})(Y_1 \frac{\partial\hat{x}_2}{\partial Y_q} + \hat{x}_2 \frac{\partial Y_1}{\partial Y_q} - \frac{\partial Y_2}{\partial Y_q}) - (Y_1\hat{x}_2 - Y_2)(\frac{1}{n}\hat{x}_1^{\frac{1}{n} - 1} \frac{\partial\hat{x}_1}{\partial Y_q}) \Big]$$

$$- (1 - \hat{x}_1^{\frac{1}{n}})(Y_1\hat{x}_2 - Y_2)\Big[(1 - \hat{x}_1) \quad (\frac{\partial\hat{x}_2}{\partial Y_q} - \frac{\partial\hat{x}_1}{\partial Y_q}) \quad - (\hat{x}_2 - \hat{x}_1) \frac{\partial\hat{x}_1}{\partial Y_q} \Big] \Big\} \quad . \qquad (7.5.4)$$

Similarly, for $\hat{\alpha}_2$,

$$a_q = \frac{\partial\hat{\alpha}_2}{\partial Y_q} = n(1 - \hat{x}_2)^{-2}(\hat{x}_2 - \hat{x}_1)^{-2} \Big\{ (1 - \hat{x}_2)(\hat{x}_2 - \hat{x}_1)\Big[(1 - \hat{x}_2^{\frac{1}{n}})(\frac{\partial Y_2}{\partial Y_q} - Y_1 \frac{\partial\hat{x}_1}{\partial Y_q}$$

$$- \hat{x}_1 \frac{\partial Y_1}{\partial Y_q}) - (Y_2 - Y_1\hat{x}_1)(\frac{1}{n}\hat{x}_2^{\frac{1}{n} - 1} \frac{\partial\hat{x}_2}{\partial Y_q})\Big]$$

$$- (1 - \hat{x}_2^{\frac{1}{n}})(Y_2 - Y_1\hat{x}_1)\left[(1 - \hat{x}_2)(\frac{\partial\hat{x}_2}{\partial Y_q} - \frac{\partial\hat{x}_1}{\partial Y_q}) - (\hat{x}_2 - \hat{x}_1)\frac{\partial\hat{x}_2}{\partial Y_q}\right]\Bigg\} \quad .$$

All of the quantities needed to evaluate these derivatives have already been calculated, so from (7.5.3) and (7.5.5) we compute

$$a_1 = 0.0168 \ , \ a_2 = -0.5576 \ , \ a_3 = 20.6095 \ , \ a_4 = -3.2847$$

for $\hat{\alpha}_1$ and

$$a_1 = 2.5540 \ , \ a_2 = -7.1179 \ , \ a_3 = 76.2966 \ , \ a_4 = -1.9009$$

for $\hat{\alpha}_2$. Then substitution in (4.1.8) and division by $mn = 4$ as before for the $\hat{\lambda}_k$ yields

$$V(\hat{\alpha}_1) = 108.9630 \ s^2 \quad , \quad V(\hat{\alpha}_2) = 1470.4930 \ s^2 \quad .$$

An estimate s^2 of σ^2 may be found in the same manner as it was in Section 7.4, and it can be shown to equal 0.03997. Then the following standard deviations may be calculated:

$$\text{s.d.}(\hat{\lambda}_1) = 0.6614 \quad , \quad \text{s.d.}(\hat{\lambda}_2) = 0.8355 \quad ,$$

$$\text{s.d.}(\hat{\alpha}_1) = 2.0869 \quad , \quad \text{s.d}(\hat{\alpha}_2) = 7.6665 \quad .$$

VIII. SUMMARY AND CONCLUSIONS

A new estimation procedure has been developed in this paper for a model specifying a linear combination of exponentials with data taken at evenly spaced points, and for that model with a constant term added. Besides the derivation of the estimation equations for the new procedure, the distributions and statistical properties of the resultant estimators were studied. It was found that as the number of observation points or as the number of observations taken at each such point becomes large, the estimators are consistent and that their distributions approach normal distributions. It was also found that the estimators are biased and generally inefficient. Then the new procedure was shown to be optimum relative to certain similar procedures and conditions necessary for admissible solutions were investigated. Confidence intervals were developed and several examples plus a sampling survey were presented. Extensions of the model for the new procedure as well as some approximations to be used in the application of the procedure were also suggested.

In the efficiency study it was shown that the estimators yielded by the new procedure are inefficient relative to the corresponding maximum likelihood estimators, which may be obtained by iterative methods. The convergence of these iterative methods is often slow. But the new procedure is not an iterative procedure, and its estimates are much easier to compute than are the maximum likelihood estimates. So we conclude that the new method, when its model is realistic, is advantageous relative to the method of maximum likelihood if adequate computing facilities are not available to calculate maximum likelihood estimates. Also, if it is more practical to take a large number of observations, and thus obtain small error estimates with the new method even though it is inefficient, than it is to take fewer observations and compute maximum likelihood estimates, the new procedure is again recommended. In any case, the new procedure provides a quick and easy way of computing initial estimates for iterative maximum likelihood calculations.

The limited empirical comparisons which have been made between the new procedure and other non-iterative, easily applied procedures do not provide an adequate basis for judging these methods relative to each other. However, the new procedure appears to be simpler computationally than the Prony and "peeling off" methods. Also, if the variance of the observations can be accurately estimated, variance estimates for the estimates from the new procedure can be calculated and useful confidence limits for the corresponding parameters can be constructed. No such measures of error are in general available for estimates from the Prony and "peeling off" procedures. Furthermore, unlike those for the "peeling off" method, the new procedure calculations do not require any judgment decisions. So if the model for the new procedure is appropriate, this procedure is in several respects optimum relative to the other non-iterative methods discussed in this paper.

IX. ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to for his patient guidance and to for his helpful criticism.

The author is also grateful to , who typed most of the thesis, to , who helped , and to , who typed an earlier draft of the thesis.

The research for this thesis was carried out at the Oak Ridge National Laboratory and was made possible by a fellowship sponsored by the Oak Ridge Institute of Nuclear Studies.

X. BIBLIOGRAPHY

1. Aitken, A. C., Determinants and Matrices, Edinburgh and London, Oliver and Boyd, 1949.

2. Cramér, H., Mathematical Methods of Statistics, Princeton, Princeton University Press, 1946.

3. Deming, W. E., Statistical Adjustment of Data, New York, John Wiley and Sons, Inc., 1946, 128-141.

4. Feurzeig, W. and Tyler, S. A., "A Note of Exponential Fitting of Empirical Curves," Argonne National Laboratory Quarterly Report ANL-4401, Lemont, Illinois, 1950, 14-29.

5. Fieller, E. C., "The Distribution of the Index in a Normal Bivariate Population," Biometrika, 24 (1932), 428-440.

6. Fieller, E. C., "A Fundamental Formula in the Statistics of Biological Assay and Some Applications," Quarterly Journal of Pharmacy and Pharmacology, 17 (1944), 117-123.

7. Fisher, R. A., "The Mathematical Foundations of Theoretical Statistics," Philosophical Transactions A, Royal Society of London, 222 (1922), 309-368.

8. Franklin, P., Methods of Advanced Calculus, New York and London, McGraw-Hill Book Company, Inc., 1944, 144-145.

9. Garwood, F., "The Application of Maximum Likelihood to Dosage-Mortality Curves," Biometrika, 32 (1941), 46-58.

10. Girshick, M. A., "Note on the Distribution of Roots of a Polynomial with Random Complex Coefficients," The Annals of Mathematical Statistics, 13 (1942), 235-238.

11. Graves, L. M., The Theory of Functions of Real Variables, New York and London, McGraw-Hill Book Company, Inc., 1946, 138-140.

12. Hamblen, J. W., Distributions of Roots of Algebraic Equations with Variable Coefficients, Ph.D. dissertation, Purdue University, 1955.

13. Householder, A. S., "On Prony's Method of Fitting Exponential Decay Curves and Multiple-Hit Survival Curves," Oak Ridge National Laboratory Report ORNL-455, Oak Ridge, Tennessee, 1949.

14. Householder, A. S., Principles of Numerical Analysis, New York and London, McGraw-Hill Book Company, Inc., 1953, 226-229.

15. Hsu, P. L., "The Limiting Distribution of Functions of Sample Means and Applications to Testing Hypotheses," Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, (August, 1945 and January, 1946), Berkeley and Los Angeles, University of California Press, 1949, 359-364.

16. Jordon, C., Calculus of Finite Differences, New York, Chelsea Publishing Company, 1947, 543-549.

17. Keeping, E. S., "A Significance Test for Exponential Regression," The Annals of Mathematical Statistics, 22 (1951), 180-198.

18. Kendall, M. G., The Advanced Theory of Statistics, Vol. II, London Charles Griffin and Company, Limited, 1946.

19. Koopman, B. O., "On Distributions Admitting a Sufficient Statistic," Transactions of the American Mathematics Society, 39 (1936), 399-409.

20. Koshal, R. S., "Application of the Method of Maximum Likelihood to the Improvement of Curves Fitted by the Method of Moments," Journal of the Royal Statistical Society, 96 (1933), 303-313.

21. Koshal, R. S., "Maximum Likelihood and Minimal χ^2 in Relation to Frequency Curves," Annals of Eugenics, 9 (1939), 209-231.

22. Merrill, A. S., "Frequency Distribution of an Index when Both the Components Follow the Normal Law," Biometrika, 20A (1928), 53-63.

23. Prony, A. L., "Essai Expérimental et Analytique sur les Lois de la Dilatabilité des Fluids Élastiques, et sur celles de la Force Expansive de la Vapeur de l' Eau et de la Vapeur de l' Alkol, à Différentes Températures," Journal de l' École Polytechnique, Cahier 2 (an IV), 24-35.

24. Running, T. R., Empirical Formulas, New York, John Wilay and Sons, Inc., 1917, 58-61.

25. Samuelson, P. A., "Conditions that the Roots of a Polynomial Be Less than Unity in Absolute Value," Annals of Mathematical Statistics, 12 (1941), 360-364.

26. Smith, R. E. and Morales, M. F., "On the Theory of Blood Tissue Exchanges. II. Applications," Bulletin of Mathematical Biophysics, 6(1944), 133-135.

27. Sverdrup, E., "The Limit Distribution of a Continuous Function of Random Variables," Skandinavisk Aktuarietidskrift, 35 (1952), 1-10.

28. Turnbull, H. W., Theory of Functions, Edinburgh and London, Oliver and Boyd, 1947.

29. Wald, A., "The Fitting of Straight Lines if Both Variables are Subject to Error," Annals of Mathematical Statistics, 11 (1940), 284-300.

30. Whittaker, E. T. and Robinson, G., The Calculus of Observations, London and Glasgow, Blackie and Son, Limited, 1944, 369-371.

31. Willers, F. A., Practical Analysis, New York, Dover Publications, Inc., 1948, 355-363.

The vita has been removed from the scanned document

Abstract of

A NEW ESTIMATION PROCEDURE FOR LINEAR COMBINATIONS OF EXPONENTIALS

by

Richard Garth Cornell

Many experimental problems in the natural sciences result in data which can best be represented by linear combinations of exponentials of the form

$$f(t) = \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t} \quad .$$

Among such problems are those dealing with growth, decay, ion concentration, and survival and mortality. Also, in general, the solution to any problem which may be represented by linear differential equations with constant coefficients is a linear combination of exponentials. In most problems like those which have been mentioned, the parameters α_k and λ_k have biological or physical significance. Therefore, in fitting the function $f(t)$ to the data it is not only necessary that the function approximate the data closely, but it is also necessary that the parameters α_k and λ_k be accurately estimated. Furthermore, a measure of the accuracy of the estimation of the parameters is required.

A new estimation procedure for linear combinations of exponentials is developed in this paper. Unlike the iterative maximum likelihood and least squares methods for estimating the parameters for such a model, the new procedure is noniterative and can be easily applied. Also, in contrast

to other non-iterative methods, error estimates are available for the parameter estimates yielded by the new procedure.

In the model for the new procedure the points t_i at which observations are taken are assumed to be equally spaced and the number of such points is specified to be an integral multiple of the number of parameters to be estimated. Moreover, each observation is specified to have expectation f (t_i), where f is the function mentioned earlier. The coefficients α_k are assumed to be non-zero and the exponents λ_k are assumed to be distinct and positive. Then in the derivation of new procedure, the observations are reduced to as many sums as there are parameters to be estimated. Each of these sums is equated to its expected value and the resultant equations are solved for estimators of the parameters.

The estimators from the new procedure are shown to be asymptotically normally distributed as either the number of points at which observations are taken or the number of observations made at each such point approaches infinity. The asymptotic variances obtained are used to form approximate confidence limits for the α_k and λ_k. The statistical properties of the estimators are also studied. It is found that they are consistent, but not in general unbiased or efficient. Asymptotic efficiencies are calculated for a few sets of parameter values and a bias approximation is obtained for two special cases. The new method is also shown to be optimum relative to certain similar methods and necessary conditions for the new procedure to lead to admissible estimates are studied.

In the last portion of the thesis a sampling study is reported for observations generated with a model containing only one exponential term and with errors which are normally distributed. The small sample biases and variances for the estimates computed from these observations are given and the effects of changes in the parameters in the model are investigated. Then some actual experimental data are fitted using both the new procedure and some alternative methods. The final chapter in the body of the thesis contains a critical evaluation of the new procedure relative to other estimation methods.