AN ABSTRACT OF THE THESIS OF

Stephen Walter Butler for the Master of Science in
(Name) (Degree)

Statistics
(Major)

Date thesis is presented May 31, 1963

Title Behavior of Estimates of the Mean Life of a Radioactive Source

Abstract approved Redacted for Privacy
(Major professor)

This thesis presents the results of sampling experiments investigating the properties of estimates of the mean life of a radioactive source in the presence of a constant unknown background.

An iterative estimation procedure is used where the estimates are maximum liklihood estimates with asymptotic variances. The estimation procedure was programmed in FORTRAN for the I.B.M. 1620 computer and a large number of examples were run to study the bias, divergence and variance of the estimates. Particular emphasis is placed upon the reduction in the number of divergent cases by improvements in the estimation procedure.

BEHAVIOR OF ESTIMATES OF THE MEAN LIFE OF A RADIOACTIVE SOURCE

by

STEPHEN WALTER BUTLER

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

August 1963

APPROVED:

Redacted for Privacy

Professor of Statistics

In Charge of Major

Redacted for Privacy

Chairman of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 31, 1963

Typed by Ola Gara

ACKNOWLEDGMENT

The author wishes to thank Dr. R. F. Link for his guidance in the completion of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

BEHAVIOR OF ESTIMATES OF THE MEAN LIFE OF A RADIOACTIVE SOURCE

CHAPTER I

INTRODUCTION

In an unpublished paper (6) by Dr. R. F. Link of Oregon State University, an iterative estimation procedure is developed to estimate three parameters associated with the distribution of particles emitted from a radioactive material. The estimates were obtained by the maximum liklihood method and the asymptotic variances are derived. Since the variances are obtained from a limit as $n \to \infty$, it is not known how good the estimates are for small values of n. One consequence of this is that the estimates may be biased for many parameter values. An additional uncertainty associated with this estimation procedure, because it is an iterative procedure, is whether or not successive iterations will always converge.

The objective of this thesis is to provide a bridge between theory and practical application. It is proposed to provide the experimenter with all the information that is needed to apply the estimation procedure with the possible exception of the tables of multipliers for the asymptotic variance of C. (6, p. 8) Of prime importance in providing this bridge are investigations of bias, divergence and variance of the estimates and the effects of parameters upon these properties of the estimates. The estimation procedure is improved to reduce the number of divergent examples. Investigations of bias, divergence, and variance of the estimates and comparisons between the unmodified and the modified estimation procedures are in chapter II.

A computer method of analysis was chosen because the analytic analysis of the estimation equations is extremely lengthy and tedious. An additional reason is that the computer program is needed for the practical application of the estimation procedure. Chapter III contains a discussion of the development of the computer program and modifications made to the estimation procedure.

The presentation would be incomplete without tying together the program and recommendations to show how to use them in practice. The simulated experiments in chapter IV serve this function as well as demonstrating how to repeat the experiment to obtain greater accuracy of the estimates.

An insufficient amount of time was available to obtain as much data on cases with background level as was obtained for the zero background cases. A few cases with different background levels were run to insure that the estimation program worked for these cases and to check the variances with the asymptotic variances.

CHAPTER II

RESULTS

Some of the notations and definitions that will be used throughout the remaining chapters will now be introduced. The average number of particles observed from the background in an interval is A. The number of particles contained in the source at the start of the experiment is B. The mean life of the source is C. The number of intervals of observation is K and the width of the intervals is D. The notation var($\hat{C}$) will be used to denote the mean squared deviations of $\hat{C}$ from true C. The notation var(C) is used to denote the asymptotic variance of C calculated from tables in (6). A parameter with a circumflex (^) above it is an estimate, the true values of a parameter are denoted without the circumflex.

A case is a set of ten examples that were run run for one set of parameter values. The frequencies for each example were calculated with independent sets of random numbers. A total of 37 cases, 370 examples, were run for selected values of the parameters K and with zero background. A total of 10 cases were run with selected values of the parameters K and $\frac{KD}{C}$ with a background level of $\frac{AC}{B}$

= .1,.2. The unmodified procedure for obtaining the estimates, described in chapter I, will be referred to as procedure I. The modifications to the general procedure will be called procedure II and procedure III.

Procedure II is essentially an improvement in the initial estimates with which the iteration procedure is started. Procedure III groups the frequencies and then calculates the estimates by procedure II. Both procedures II and III are discussed more thoroughly in chapter III on program development.

Table 1 gives an overall view of the parameter combinations that were used for most of the tests in this chapter. The numbers in the blocks are the levels of K that were used and each K represents one of the 37 cases. All 37 cases were with $\frac{AC}{B} = 0$ and C = 1. The condensed data for all cases run are in Appendix A.

INVESTIGATIONS OF BIAS IN THE ESTIMATES

As an overall check on bias it was found that the average estimate $\hat{C}$ was below the true value of C in 29 out of 43 cases (67%).

The number of cases where the average $\hat{C}$ was less than C was tabulated in one table for each of the parameters

Table 1. Parameter Combinations for Cases 1 through 37.

$\frac{KD}{C}$ \ B	100	500	1,000	10,000
1	5 8	4 5	3 5 8	3 5
2	5 5 8	3 4 5 8	3 4 5 8	3 5
3	5	5	5	5
4		3 5 8	3 4 5 8	5
8			3 4 8	

The numbers in the blocks are the levels of K that were used.

B, K, and KD/C. The object of the tables was to test the null hypothesis, that the parameter has no effect on bias, for each of the three parameters. A χ^2 test at the .05 level of significance was used for each case. The null hypotheses were accepted for the parameters B and K and these tables were not included. The computed χ^2 for table 2 is 9.00 which is significant and KD/C is therefore assumed to have an effect upon bias. From observing the changes in the proportion of cases where $\hat{C} < C$ in table 2 it appears desirable to test for a change in the effect of KD/C. The null hypothesis, that there is no difference in the proportion of biased cases for KD/C = 3, 4 and 8, was accepted at the .05 level of significance. The interpretation of these tests is that the estimates are biased when $KD/C < 3$ but they did not exhibit a significant amount of bias when $KD/C \geq 3$.

From table 3 it can be seen that the estimates did not have a significant amount of bias when the asymptotic variance was less than .05.

Table 2. The Effect of KD/C on Bias.

	KD/C 1	2	3	4	8	3,4,8 Total	1,2,3, 4,8 Total
Number of Cases where $\hat{C} < C$	8	12	2	5	2	9	29
Total Number of Cases	8	16	4	12	3	19	43
Proportion of Cases $\hat{C} < C$	1.00	.75	.50	.42	.66	.47	.67

Table 3. The Change in Bias with the Asymptotic Variance.

Asymptotic Variance of C	Proportion of Cases Where $\hat{C} < C$
0 to .05	.48
.05 to 1.0	.73
over 1.0	1.00

INVESTIGATIONS OF DIVERGENCE OF THE ESTIMATES

In collecting the data several criteria were used to determine the convergence or divergence of the estimates. The best indicator of convergence was found to be a decreasing step size, where step size is defined as the absolute value of the difference between $\hat{C}$ in the ith iteration and $\hat{C}$ in the i-1 iteration. An example was judged divergent if the step sizes increased in three successive iterations. A limit had to be placed on the number of iterations that would be allowed so that computer time would not be wasted. This limit was set at 12 iterations.

A large number of divergent examples was encountered in cases where KD/C or B or both were small. Divergence was considered to be the most undesirable property of the estimation procedure because, as shown in the experiments in chapter IV, even a very poor estimate would enable the parameters to be adjusted and the experiment repeated for greater accuracy. The greatest amount of time and effort was, therefore, devoted to reducing the number of divergent examples by modifications to the estimation procedure. The modifications are discussed in chapter III.

The number of divergent examples in tables 4, 5, 6 and 7 are the results from procedure II. In order to eliminate interaction effects, each level of B includes data only for cases with the possible parameter combinations of K = 5 and KD/C = 1,2,3. A test at the .05 significance level will be made of the null hypothesis, B has no effect on divergence. The computed χ^2 is 7.55 which is significant and the null hypothesis is rejected.

Each block of table 5 includes all the cases for the possible parameter combinations of K = 5 and all four levels of B. A test at the .05 level of significance will be made of the null hypothesis, KD/C has no effect upon divergence. The computed χ^2 value for this test is 46.21 which is significant and the hypothesis is rejected.

Table 4. The Effect of B on Divergence

	B				
	100	500	1,000	10,000	Total
Number Divergent	14	10	6	6	36
Total Number	30	40	30	30	130
Proportion Divergent	.47	.25	.20	.20	.277

Table 5. The Effect of KD/C on Divergence.

	KD/C			
	1	2	3	Total
Number Divergent	27	6	3	36
Total Number	40	50	40	130
Proportion Divergent	.675	.120	.075	.277

A test of the effect of K on divergence was not significant at the .10 level.

In the discussion of the theory in (6) the only recommendation on the interval width is that it should not be too large. With respect to divergence of the estimates there apparently is a minimum interval width. When the interval width is reduced the number of counts per interval decreases. As the interval size is reduced a point will be reached where the number of counts per interval become so small that the random variation in number of counts between intervals would render the estimation procedure useless.

Table 6 shows the proportion divergent examples above and below a D/C value of .50 when B = 1,000. D/C

Table 6. Number of Divergent Examples Above and Below D/C = .50

	D/C Below .50	D/C Above .50
Number Divergent	22	12
Total Number	60	150
Proportion Divergent	.367	.080

is the number of mean lives of the source that is contained in the interval D. The table includes only cases where B = 1,000. The recommended value D/C = .50 may be decreased slightly when B $>$ 1,000, but it should be increased when B $<$ 1,000.

Table 7. The Change in Fraction Divergent with the Asymptotic Variance of C.

Asymptotic Variance of C	Fraction Divergent	Total Number of Examples
0 to .05	.06	160
.05 to .10	.23	40
.10 to .30	.33	80
over .30	.46	90

INVESTIGATIONS OF THE VARIANCE OF $\hat{C}$

Direct comparisons between the observed variances and the asymptotic variances were not possible because of the influence of the divergent examples. Almost without exception the observed variances were smaller than the asymptotic variances when there were two or more divergent examples in a case. This fact coupled with the results of table 7 lead one to believe that the examples that were divergent would have had greater variance than the remaining convergent examples in a case. This contention is supported in tables 9 and 10.

The only attempt to compare the observed and asymptotic variances is table 8 which shows the ratio, the observed var($\hat{C}$) divided by the asymptotic var(C), for all 37 cases.

Tables 9, 10 and 11 show the effects of the parameters on the observed variances and all are in accord with the theory. The degree of effect on the variances cannot be validly assessed because of the divergent examples. The proportion of divergent examples for procedure II and the asymptotic variances are included in the tables so that the effects of the divergent examples

Table 8. The Ratios of Observed Variances of $\hat{C}$ to the Asymptotic Variances of C.

		B							
		100		500		1,000		10,000	
	1	5	.0219	4	.0928	3	.1786	3	.921
		8	-	5	.1093	5	.1712	5	.277
						8	.0644		
	2	5	.109	3	.402	3	4.198	3	3.769
		5	.315	4	.305	4	.824	5	.675
		8	.110	5	.542	5	.353		
				8	.881	8	.601		
KD/C	3	5	.0379	5	.511	5	1.431	5	.500
	4			3	1.192	3	.348	5	.857
				5	1.457	4	.846		
				8	1.043	5	.671		
						8	.603		
	8					3	7.33		
						4	32.50		
						8	14.00		

can be observed. The observed variances are less than the asymptotic variances in tables 9 and 10 when the proportion divergent is .20 or larger. Some interaction effects are

Table 9. The Effect of B on the Observed Variance of $\hat{C}$.

	B			
	100	500	1,000	10,000
Average var($\hat{C}$)	.2033	.1467	.1054	.0177
Average asymptotic var(C)	4.09	2.045	.4090	.0409
Proportion Divergent	.433	.233	.233	.200

included in table 11 but the relationship between the observed and the asymptotic variances for several levels of divergence can be seen.

The effects of B on the variance are the average of the variances for each of the cases where K = 5 and KD/C = 1,2,3, in order to eliminate interaction effects. Similarly the variances in table 10 are averages for all cases where K = 5 and for all levels of B.

Table 10. The Effect of KD/C on the Observed Variance of $\hat{C}$.

	KD/C			
	1	2	3	4
Average var($\hat{C}$)	.2454	.0888	.0175	.0088
Average Asymptotic var(C)	.5116	.0263	.0056	.0023
Proportion Divergent	.550	.175	.100	.025

Table 11. The Effect of K on the Observed Variance of $\hat{C}$.

	K			
	3	4	5	8
Average var($\hat{C}$)	.1233	.0285	.0985	.0321
Average Asymptotic var(C)	.6135	.0320	.2070	.3406
Proportion Divergent	.150	.066	.20	.30

THE ADVANTAGES OF THE MODIFIED ESTIMATION PROGRAM

As mentioned before, most of the time and effort was devoted toward reducing the number of divergent examples. Table 12 shows that a very significant reduction in number of divergent examples was obtained. A reduction in the number of iterations required for convergence was an additional advantage of the particular modifications that were utilized.

The number divergent for procedure II in table 13 is smaller than in table 12 because procedure III was not run for all cases. The proportion divergent in table 12 is larger than in table 13 because cases for which all 10 examples converged were not included.

Table 12. Proportion Divergent Examples for Procedures I and II.

	Procedure I (unmodified)	Procedure II (modified)
Number divergent	135	84
Total	360	360
Fraction Divergent	.375	.233

Table 13. Proportion Divergent Examples for Procedure III.

	Procedure II	Procedure III
Number Divergent	43	6
Total	120	120
Fraction Divergent	.358	.050

The method used to obtain the data for the comparisons in tables 12 and 13 was to run both the modified and unmodified estimation procedures for each of the cases number 1 through 36. Procedure III was used when procedure II was divergent and K $\geqslant 5$.

To compute the fraction divergent for the combined modified procedures, the fraction divergent for procedure III in table 13 is multiplied by the fraction divergent for procedure II in table 12.

Total Fraction Divergent for

Combined Modified Procedures = $(.233)\cdot(.05) = .012$

This figure is approximate since independence of the two procedures is assumed and the fraction divergent data are averages over observed results.

Only the examples were included in table 14 for which

Table 14. Number of Iterations Required for Convergence.

Number of Iterations	Procedure I	Procedure II
1	2	11
2	9	55
3	33	62
4	50	21
5	38	27
6	26	16
7	24	6
8	16	4
9	4	2
10	3	2
11	1	1
12	1	0

both procedures were convergent. The average number of iterations required for procedure I is 5.1 and for procedure II is 3.6. Procedure II requires, on the average, 1.5 less iterations than procedure I.

Due to the fact that the average number of iterations is smaller for the modified procedure, the total program running time is less than for procedure I. The procedures I and II were timed for several cases and the observed times were averaged. When K = 5 the average times were 1.72 minutes and 1.54 for procedures I and II respectively. When K = 8 the average times were 2.61 and 2.29.

CHAPTER III

THE ESTIMATION EQUATIONS

The basic exponential density function that is usually assumed in radioactive decay problems is

$$p(t) = \frac{1}{C} e^{-t/C}$$

When a value for the mean life (C) is given the probability of a particle emission at any time t may be calculated. Multiplying p(t) by the mean number of particles contained in the source at the start of observation (B) gives q(t) with which the expected number of particles emitted from the source at any instant of time may be calculated.

$$q(t) = \frac{B}{C} e^{-t/C}$$

In practical problems there usually is a small relatively constant particle emission from objects in the vicinity of the experiment which will add to the total particle count. The average intensity of this background emission is A. The number of particles contained in the source is assumed to be Poisson distributed and independent of the background particles. Similarly the number of particles contained in the background is assumed to be Poisson distributed. The particles in each of the Poisson distributions are also assumed to be independent of each other.

When background emission is present the expected number of particles emitted at any instant of time is d(t).

$$d(t) = \frac{B}{C} e^{-t/C} + A$$

The probability density function of the time of arrival of a particle at the counter when the source is observed for a total time interval T is f(t). (6, p.25)

$$f(t) = \frac{\frac{B}{C} e^{-t/C} + A}{A \cdot T + B\left(1 - e^{-T/C}\right)}$$

The maximum liklihood estimates derived in (6) are obtained from the joint probability density function of the times $(t_1, t_2, \ldots, t_n)$ at which counts are recorded.

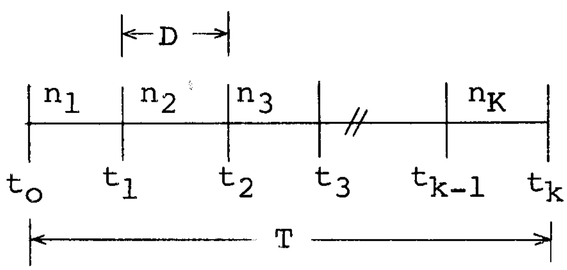
$$f(t_1, t_2, \ldots, t_n, n \mid A, B, C) = \prod_{i=1}^{n} \left\{ \frac{B}{C} e^{-t_i/C} + A \right\} \frac{e^{-\left[A \cdot T + B\left(1 - e^{-T/C}\right)\right]}}{n!}$$

where n is the total number of counts in interval T.

$$i = 1,2,\ldots,n$$

$$0 \leq t_i \leq T$$

The emitted particles are counted for discrete intervals of time. The time scale is divided in the following manner:



where $t_i - t_{i-1} = D$ is the interval length

$K*D = T$ is the total time of observation

n_i is the number of observed counts in interval i

The asymptotic variance of the estimates is MC^2/B where the multiplier M may be written as a function of several parameters M(K,T/C,AC/B,D,d). There are four tables in the paper (6) that give values for the multiplier for several values of T/C, AC/B, K and background level assumed known and unknown.

The initial estimates A_o, B_o and C_o are obtained from the following relations:

$$A_o = \frac{n_1}{D}$$

$$C_o = \frac{D + d}{Ln\left(\dfrac{n_1 - A_o \cdot D}{n_2 - A_o \cdot D}\right)}$$

$$B_o = \frac{n_1}{1 - e^{-D/C_o}}$$

These initial values are used in the iterative procedure to solve for improved estimates A_1, B_1 and C_1. The procedure consists of solving the following set of equations for A_1, B_1 and C_1 and it may be repeated until the desired accuracy is attained.

$$-F = (C_1 - C_o) \cdot M + (B_1 - B_o) \cdot N + (A_1 - A_o) \cdot S$$

$$-G = (C_1 - C_o) \cdot N + (B_1 - B_o) \cdot P + (A_1 - A_o) \cdot Q$$

$$-H = (C_1 - C_o) \cdot S + (B_1 - B_o) \cdot Q + (A_1 - A_o) \cdot R$$

Where

$$F = \sum_{i=1}^{K} W_i \left(\frac{n_i - m_i}{m_i}\right)$$

$$G = \sum_{i=1}^{K} X_i \left(\frac{n_i - m_i}{m_i}\right)$$

$$H = \sum_{i=1}^{K} Y_i \left(\frac{n_i - m_i}{m_i}\right)$$

$$M = \sum_{i=1}^{K} \left\{ Z_i \left(\frac{n_i - m_i}{m_i}\right) - W_i^2 \cdot \frac{n_i}{m_i^2} \right\}$$

$$N = \sum_{i=1}^{K} \left\{ E_i \left(\frac{n_i - m_i}{m_i}\right) - W_i \cdot X_i \cdot \frac{n_i}{m_i^2} \right\}$$

$$S = \sum_{i=1}^{K} - W_i Y_i \frac{n_i}{m_i^2}$$

$$P = \sum_{i=1}^{K} - X_i^2 \cdot \frac{n_i}{m_i^2}$$

$$Q = \sum_{i=1}^{K} - X_i \cdot Y_i \cdot \frac{n_i}{m_i^2}$$

$$R = \sum_{i=1}^{K} - Y_i^2 \frac{n_i}{m_i^2}$$

$$D_i = \frac{B}{C^4} e^{-(i-1)\frac{D+d}{C}} \left\{ \left[(i-1)(D+d)\right]^2 - e^{-D/C} \left[(i-1)D+d+D\right]^2 \right\} - \frac{2W_i}{C}$$

$$E_i = \frac{W_i}{B}$$

$$A_i = \frac{B}{C^2} e^{-(i-1)\frac{D}{C}} \left\{ \left[(i-1)D+d\right] - e^{-D/C} \left[(i-1)(D+d)+D\right] \right\}$$

$$m_i = BX_i + A \cdot D$$

$$Y_i = D$$

$$X_i = e^{-(i-1) \cdot \frac{D+d}{C}} \left(1 - e^{-D/C}\right)$$

The quantity d has been assumed to be zero in all cases presented in this thesis. This quantity is an interval of non-observation of the source between each interval of observation and it is assumed to be of constant length in the equations just presented. It may be desirable to include d in the estimation procedure if sources of extremely long half lives are to be observed. It is a simple task to insert d in the appropriate program statements if it is needed.

GENERAL PROGRAM DEVELOPMENT

The computer used in the calculation of all the cases was the I.B.M. 1620 with a memory capacity of 40,000 digits and card and typewriter input-output. All of the supporting card handling equipment was also available. The computer programs were all written in FORTRAN.

The estimation program was first written directly from the estimation equations in the preceding section. This program exceeded the computer memory capacity, and it was modified to require less memory space. The program was then checked by running 30 cases. In these cases the frequency counts were calculated from the function d(t) in the previous section.

In order to investigate the behavior of the estimates, frequency count data were simulated by a Monte Carlo procedure. The fact that the number of counts in an interval is Poisson distributed enabled the use of one random number for simulation of each frequency count. The Poisson distribution was used to convert random numbers to frequency counts when the expected count for the interval was six or less. The normal approximation to the Poisson distribution was used to convert random numbers to frequency counts when the expected count was greater than six.

The computer program and detailed operating instructions may be obtained from the Statistics Computing Laboratory, Oregon State University.

MODIFICATIONS TO THE ESTIMATION PROCEDURE

A number of modifications were made to the estimation procedure most of which were aimed toward reducing the number of divergent examples. The most significant modification was the improvement of the initial estimates A_o, B_o and C_o. The original initial estimates are:

$$A_o = \frac{n_k}{D}$$

$$C_o = \frac{D}{\ln\left(\dfrac{n_1 - A_o \cdot D}{n_2 - A_o\ D}\right)}$$

$$B_o = \frac{n_1}{1 - e^{-D/C_o}}$$

These initial estimates were modified to the following:

$$A_o = \frac{n_k}{D} \tag{1}$$

$$A_o'' = A_o \tag{2}$$

$$C_o' = \frac{D}{\ln\left(\dfrac{n_1 - A_o \cdot D}{n_2 - A_o \cdot D}\right)} \tag{3}$$

$$C_o'' = \frac{D}{\ln\left(\dfrac{n_2 - A_o D}{n_3 - A_o D}\right)} \tag{4}$$

$$C_o = \frac{C_o' + C_o''}{2} \qquad (5)$$

$$B_o = \frac{n_1 - A_oD}{1 - e^{-D/C_o}} \qquad (6)$$

$$A_o = A_o'' + \frac{B_o}{D}\left(e^{-KD/C_o} - e^{-(K-1)D/C_o}\right) \qquad (7)$$

Steps 3,4,5,6 and 7 are repeated once more letting $A_o' = A_o$.

The initial estimate of A in both the modified and the unmodified procedures is: $A_o = n_k/D$. This estimate for A will usually be too large and, in some cases, it may be much too large. The quantity n_k/D is an estimate of A if the number of counts from the source in interval K is zero, however this is seldom true and there may be an appreciable number of counts from the source in interval K. The estimate of A_o is improved by subtracting the theoretical count for interval K from A_o. The theoretical count for interval K is an estimate since it is calculated with B_o and C_o.

The initial estimate C_o is obtained from the average of C_o' and C_o'' which are estimates of C_o obtained from the ratio n_1/n_2 and n_2/n_3 respectively. The reason for using the average of two rather than one estimate is that when n_2 is close to n_1 in size the estimate C_o' is too large. The denominator is the logarithm of the ratio of the frequency counts which becomes very small as the ratio approaches one. There are two reasons for choosing the ratios of these particular intervals. The first reason is that differences between frequency counts in the first few intervals are usually greater than in the remaining intervals. The second reason is that random variations in the frequency counts of adjacent pairs will cancel out in many cases. For instance, if n_2 assumes a high value the difference between n_1 and n_2 is small, however the difference between n_2 and n_3 will likely be larger and thereby effectively compensate for the high value of n_2.

The estimates of C may still cause trouble, in particular if $n_2 \geq n_1$ or $n_3 \geq n_2$ the estimate of C_o will be negative or undefined and the procedure will not converge. The following alternative estimates of B_o and C_o are used in these situations.

$$B_o = n$$

$$C_o = \frac{2D}{\ln\left(\frac{n_1 - A_oD}{n_3 - A_oD}\right)}$$

The total frequency count over all intervals of observation is n. The derivations of the modified initial estimates are given in appendix B.

The estimates occasionally converge with $A < 0$ especially when the actual background count is very slight. Since none of the parameters can have negative values, the computer print out of the estimate of A in these cases was changed to zero. Because of the relationship between the estimates $\hat{A}$ and $\hat{B}$, $\hat{B}$ is always larger than its true value when $A < 0$, therefore the modification $\hat{B} = \hat{B} - |\hat{A}|$ was made for computer print out. Due to these changes, the accuracy of the estimates $\hat{A}$ and $\hat{B}$, when the background is light, is better than that indicated by theory.

All the modifications discussed to this point are included in procedure II. From table 12, the proportion divergent for procedure II is 23.3% which was still an uncomfortably large percentage of the examples.

It was observed that a large number of the divergent examples occurred when the interval width was short and the differences between the frequency counts were small. Table 6 shows that the interval width relative to the mean life (D/C) has a very adverse effect upon convergence.

Procedure III was devised to enable examples to converge when the interval width is too small for convergence by procedure II. This procedure forms new frequency counts (n_i) by adding adjacent pairs of frequencies.

$$n'_{i'} = n_{2i'} + n_{2i'-1} \qquad i' = 1,2,\ldots,K'$$

The new interval width (D') is doubled and the number of intervals is halved.

$$D' = 2D \qquad K' = K/2 \qquad \text{for K even}$$

$$K' = (K + 1)/2 \quad \text{for K odd}$$

A frequency count must be created for the K + 1 interval. This is accomplished by calculating an estimate of the theoretical frequency count for the interval using the modified initial estimates A_o, B_o and C_o.

$$n'_{K'} = n_{2K'} + n_{2K'-1}$$

$$n'_{K'} = n_K + n_{K+1} \qquad \text{for K odd}$$

$$n'_{K'} = n_K + \int_{t_{K+1}}^{t_K} \left[\frac{B_o}{C_o} e^{-t/C_o} + A_o \right] dt \qquad \text{for K odd}$$

It may not be desirable to create artificial data; however the expected count in interval K + 1 should be small relative to the other intervals and the error in estimating this small frequency is assumed to be small. Furthermore the estimation error is reduced relative to the total count in the interval when n_k and n_{k+1} are added together. The use of the artificial data should be avoided whenever possible by selecting an even number of intervals. Procedure III can be used only when $K \geq 5$. The reason for this is that initial estimates C_o and B_o cannot be obtained when $K \geq 5$.

CHAPTER IV

TWO SIMULATED EXPERIMENTS

These experiments demonstrate the repeatability of the estimation procedure to obtain accurate estimates of the parameters. The experiments also demonstrate the use of recommended parameter values and experiment 2 demonstrates an additional advantage of procedure III. The true values of the parameters are denoted by A, B and C. The estimates of the parameters are denoted by $\hat{A}$, $\hat{B}$ and $\hat{C}$. For comparisons, the true values of the parameters and the asymptotic variance of C are listed beside each set of estimates in the experiments.

It is assumed that a particle counter is available that will accurately count 5,000 particles per unit of time. It is also assumed that the experimenter has a sufficient quantity of the radioactive material in question.

EXPERIMENT 1

This experiment assumes that the experimenter made a poor initial guess that $K = 4$ and $D = .25$. The experimenter obtained the frequency count data and used the estimation program.

$$\hat{A}_1 = 408 \qquad A_1 = 0$$

$$\hat{B}_1 = 250 \qquad B_1 = 1{,}000$$

$$\hat{C}_1 = .401 \qquad C_1 = 1.000$$

$$\mathrm{var}(C) = 1.690$$

A value for KD/C of 10 was chosen and a value for D of .40. The value for D was chosen to make the ratio D/C approximately unity. The value for K must be calculated from the relation KD/C=10 using $\hat{C}_1$ and D.

New data were obtained using the new values for K and D, and the source size was doubled. The new estimates are:

$$\hat{A}_2 = 6 \qquad A_2 = 0$$

$$\hat{B}_2 = 1962 \qquad B_2 = 2{,}000$$

$$\hat{C}_2 = .935 \qquad C_2 = 1.000$$

$$\mathrm{var}(C) = .0005$$

New values were chosen that were KD/C=10, D=.90 and the source size again doubled.

$$\hat{A}_3 = 1 \qquad A_3 = 0$$

$$\hat{B}_3 = 3{,}990 \qquad B_3 = 4{,}000$$

$$\hat{C}_3 = .992 \qquad C_3 = 1.000$$

$$\mathrm{var}(C) = .0002$$

The experiment could be continued for a few more trials until B is at a maximum of 5,000 and the estimates exhibit very little change.

The conservative increase of B should be noted. If the source for the second part of the experiment were increased to its maximum on the basis of the estimate $\hat{B}_1$, then the true value of B_2 would be 20,000 and this would introduce considerable counter error into the estimates.

EXPERIMENT 2

The experimenter in this case felt that he could not make a sufficiently accurate guess of the values for K and D to enable the estimation procedure to converge. He would utilize procedure III of the program by selecting a large even number for K and a small number for D. This enables procedure III to be used several times to locate a good combination of K and D without repeating the counting data.

The first values chosen were K = 12 and D = .20 and the estimates did not converge. Procedure III was used which changed the values to K = 6 and D = .40 for which the estimates did converge.

$$\hat{A}_1 = 20 \qquad A_1 = 100$$

$$\hat{B}_1 = 1229 \qquad B_1 = 1{,}000$$

$$\hat{C}_1 = 1.192 \qquad C_1 = 1.000$$

$$\text{var}(C) = .0302$$

The new values $KD/C = 10$ and $D = 1.0$ were chosen and the source was increased four times its original size.

$$\hat{A}_2 = 97 \qquad A_2 = 100$$

$$\hat{B}_2 = 4{,}030 \qquad B_2 = 4{,}000$$

$$\hat{C}_2 = .999 \qquad C_2 = 1.000$$

$$\text{var}(C) = .0002$$

CHAPTER V

SUMMARY

The estimates $\hat{C}$ are biased and the amount of bias increases with the asymptotic variance. The bias was judged to be not excessive since there was not a significant amount when $KD/C \geq 3$.

Divergence of the estimates was considered to be the most undesirable property. Through program modifications, the number of divergent examples was reduced from 37.5% to 1.2% with $K \geq 5$. A minimum value for D/C is recommended to be .50 when $B = 1,000$. This minimum value may be smaller when $B > 1,000$ but it should be larger when $B < 1,000$.

Comparisons between observed variances and the asymptotic variances were hampered by the effects of the divergent examples. The contention that the observed variances of $\hat{C}$ were smaller when two or more examples of a case were divergent is supported in tables 9 and 10. The observed variances were in good agreement with the theory when the number of divergent examples were one or zero. The effects of the parameters B, KD/C and K were in accord with the theory with respect to direction of change caused by them.

The modified procedure was better than the unmodified procedure in all respects. As mentioned above, the number of divergent examples was reduced considerably. The number of iterations required for convergence was 1.5 less for the modified procedure. When K = 8 the modified procedure required, on the average, one-third of a minute less time to run than the unmodified procedure. The recommended values for the parameters are summarized below.

KD/C will generally be obtained from the tables in (6) and it should always be greater than or equal to 3.

K must be greater than or equal to 5.

D/C should be .50 when B = 1,000. It may be smaller when $B > 1,000$, but it should be larger when $B < 1,000$.

B should always be as large as possible.

BIBLIOGRAPHY

1. Bennett, Carl A. and Norman L. Franklin. Statistical analysis in chemistry and the chemical industry. New York, John Wiley and Sons Inc., 1954. 711p.

2. Cochran, William G. and Gertrude M. Cox. Experimental designs. New York, John Wiley and Sons Inc., 1957. 611p.

3. Crow, Edwin L., Francis A. Davis, Margaret W. Maxfield. Statistics manual. New York, Dover, 1960. 288p.

4. Fraser, D. A. S. Statistics; an introduction. John Wiley and Sons Inc., 1958. 398p.

5. International Business Machines Corporation. Reference Manual I. B. M. 1620 FORTRAN. 1962. 94p.

6. Link, Richard F. On the estimation of the mean life of a radioactive source. (Mimeographed) Corvallis, Statistics Department, Oregon State University. n.d. 28p.

7. McCracken, Daniel D. A guide to FORTRAN programming. New York, John Wiley and Sons, Inc., 1961. 87p.

APPENDIXES

APPENDIX A

Case	B	K	C	KD/C	No. Divergent by Procedure I	II	III	Max. $\hat{C}$	Min. $\hat{C}$	Max. $\hat{C}$ minus Min. $\hat{C}$	Avg. $\hat{C}$	var $(\hat{C})$	$M\frac{C^2}{B}$ Asymptotic Variance
1	500	5	1	1	8	7	2	.534	.326	.208	.459	.3357	3.070
2	500	4	1	1	10	7	-	1.012	.311	.701	.546	.3137	3.380
3	10000	3	1	1	10	5	-	.845	.514	.331	.642	.1414	.2340
4	1000	3	1	1	10	3	-	.593	.183	.410	.361	.4342	2.430
5	1000	5	1	1	10	4	-	.864	.358	.506	.527	.2628	1.535
6	10000	5	1	1	10	6	-	.850	.681	.169	.794	.0469	.1535
7	100	8	1	1	10	10	-	-	-	-	-	-	-
8	100	5	1	1	4	1	-	.419	.419	0	.419	.3364	3.070
9	10000	3	1	2	0	0	-	1.127	.910	.217	.977	.0049	.0126
10	1000	3	1	2	2	2	-	1.261	.627	.634	.869	.0529	.1260
11	500	4	1	2	3	3	-	1.118	.621	.497	.833	.0539	.1762
12	100	5	1	3	6	3	-	1.275	.634	.641	1.038	.0394	.1750

Appendix A - Continued

Case	B	K	C	KD/C	No. Divergent by Procedure I	II	III	Max. $\hat{C}$	Min. $\hat{C}$	Max. $\hat{C}$ minus Min. $\hat{C}$	Avg. $\hat{C}$	var $(\hat{C})$	$M \frac{C^2}{B}$ Asymptotic Variance
13	500	5	1	4	1	1	-	1.261	.866	.395	1.002	.0204	.0141
14	10000	5	1	4	2	0	-	1.027	.964	.063	.993	.0006	.0007
15	1000	5	1	4	2	0	-	1.118	.914	.204	1.011	.0047	.0070
16	100	5	1	4	2	0	-	1.672	.779	.893	1.121	.0872	.0700
17	500	5	1	2	1	0	-	1.611	.583	1.028	.975	.0864	.1600
18	10000	5	1	2	0	0	-	1.120	.905	.215	1.003	.0054	.0080
19	1000	5	1	2	2	2	-	1.218	.744	.474	.934	.0282	.0800
20	100	8	1	2	8	5	-	1.629	.964	.665	1.272	.1402	.6680
21	100	5	1	2	4	5	0	1.519	.341	1.178	.685	.2341	.8000
22	500	5	1	3	2	0	-	1.202	.877	.325	1.066	.0280	.0340
23	1000	5	1	3	1	1	0	1.432	.907	.525	.942	.0252	.0170
24	10000	5	1	3	0	0	-	1.032	.934	.098	.993	.0009	.0017

Appendix A - Continued

Case	B	K	C	KD/C	No. Divergent by Procedure I	II	III	Max. $\hat{C}$	Min. $\hat{C}$	Max. $\hat{C}$ Minus Min. $\hat{C}$	Avg. $\hat{C}$	var($\hat{C}$)	$M \frac{C^2}{B}$ Asymptotic Variance
25	1000	4	1	1	5	5	-	.887	.295	.592	.477	3.1883	1.690
26	1000	8	1	1	3	4	3	.971	.487	.484	.731	.0832	1.290
27	1000	4	1	2	0	0	-	1.557	.647	.910	.974	.0726	.0881
28	1000	8	1	2	6	7	1	1.268	.594	.674	1.041	.0420	.0668
29	1000	3	1	4	0	0	-	1.062	.941	.121	.990	.0039	.0112
30	1000	4	1	4	1	0	-	1.161	.917	.244	1.017	.0066	.0078
31	1000	8	1	4	4	1	0	1.070	.879	.191	.992	.0035	.0058
32	1000	3	1	8	4	1	-	1.088	.953	.135	1.027	.0022	.0032
33	1000	4	1	8	4	2	-	1.104	.872	.232	.958	.0065	.0021
34	1000	8	1	8	0	0	-	1.073	.940	.133	.960	.0014	1.290
35	500	8	1	2	5	5	0	1.430	.410	1.020	.893	.1177	.1336
36	500	3	1	4	1	0	-	1.561	.744	.817	1.047	.0444	.0224

Appendix A - Continued

Case	B	K	C	KD/C	No. Divergent by Procedure I	II	III	Max. $\hat{C}$	Min. $\hat{C}$	Max. $\hat{C}$ minus Min. $\hat{C}$	Avg. $\hat{C}$	var($\hat{C}$)	$M\frac{C^2}{B}$ Asymptotic Variance
37	500	8	1	4	2	3	0	1.203	.901	.302	1.025	.0121	.0116
38	500	3	1	2	3	2	-	1.346	.498	.848	.816	.1013	.2520

Note: $\frac{AC}{B} = 0$ for cases 1 through 38.

Case	A	B	C	K	KD/C	Max. $\hat{C}$	Min. $\hat{C}$	Max. $\hat{C}$ minus Min. $\hat{C}$	Avg. $\hat{C}$	var($\hat{C}$)	$M\frac{C^2}{B}$
39	200	1000	1	5	2	1.398	.613	.785	.958	.0529	.0800
40	200	1000	1	3	2	1.263	.569	.694	.848	.0750	.1260
41	125	500	1	5	4	1.291	.695	.596	1.031	.0274	.0140
42	125	500	1	3	4	1.478	.670	.808	.994	.0610	.0124
43	125	500	1	5	2	1.636	.297	1.339	.786	.1634	.0800
44	125	500	1	3	2	1.623	.867	.756	1.144	.0775	.1260

Appendix A - Continued

Case	A	B	C	K	KD/C	Max. $\hat{C}$	Min. $\hat{C}$	Max. $\hat{C}$ minus Min. $\hat{C}$	Avg. $\hat{C}$	var($\hat{C}$)	$M \frac{C^2}{B}$
45	500	5000	2	5	5	2.219	1.839	.380	2.013	.0127	.0008
46	500	5000	2	3	5	2.205	1.848	.357	2.003	.0102	.0013
47	500	5000	2	20	5	2.125	1.906	.219	1.986	.0047	.0006
48	500	5000	2	10	5	2.113	1.917	.196	1.996	.0370	.0006
49	0	500	1	12	6	1.110	.920	.190	.985	.0044	.0010
50	0	500	1	8	6	1.236	.953	.283	1.087	.0201	.0011
51	0	500	1	3	6	1.166	.942	.224	1.004	.0096	.0225
52	0	500	1	4	6	1.152	.890	.262	.982	.0083	.0015
53	0	500	1	6	6	1.161	.935	.226	.999	.0067	.0011

APPENDIX B

DERIVATION OF MODIFIED INITIAL ESTIMATES

The function d(t) from chapter III is integrated between the limits t_i and t_{i-1} to give the expected count (n_i) for the interval i.

$$n_i = \int_{t_{i-1}}^{t_i} \left[\frac{B}{C} e^{-t/C} + A\right] dt \tag{1}$$

$$n_i = B\left[e^{-t_{i-1}/C} - e^{-t_i/C}\right] + A\left(t_i - t_{i-1}\right) \tag{2}$$

Letting i = 1 in equation (2) and solving for B gives the modified initial estimate B_o.

$$n_1 = B\left(1 - e^{-D/C}\right) + A \cdot D \tag{3}$$

$$B_o = \frac{n_1 - A \cdot D}{1 - e^{-D/C}} \tag{4}$$

Using equation (3) and obtaining an additional equation by letting i = 2 in equation (2) yields the following pair of equations.

$$n_1 = B\left(e^{-t_1/C} - e^{-t_2/C}\right) + A \cdot D \tag{5}$$

$$n_2 = B\left(e^{-t_2/C} - e^{-t_3/C}\right) + A \cdot D \tag{6}$$

The substitution of $i \cdot D$ for t_i is made in both equations.

$$n_1 = B\left(1 - e^{-D/C}\right) + A \cdot D \tag{7}$$

$$n_2 = B\left(e^{-D/C} - e^{-2D/C}\right) + A \cdot D \tag{8}$$

The ratio of equations (7) and (8) is formed and the result is solved for C.

$$\frac{n_1 - A \cdot D}{n_2 - A \cdot D} = \frac{B\left(1 - e^{-D/C}\right)}{Be^{-D/C}\left(1 - e^{-D/C}\right)}$$

$$\frac{n_1 - A \cdot D}{n_2 - A \cdot D} = e^{D/C}$$

$$\frac{D}{C} = \ln\left(\frac{n_1 - A \cdot D}{n_2 - A \cdot D}\right)$$

$$C_o' = \frac{D}{\ln\left(\frac{n_1 - A \cdot D}{n_2 - A \cdot D}\right)} \tag{9}$$

An estimate for C_o'' was found in the same manner from the ratio of the equations for n_2 and n_3. The modified initial estimate for C_o is the arithmetic average of C_o' and C_o''.

The initial estimate of A in both the modified and unmodified procedures is $A_o = \frac{n_K}{D}$. The initial estimate of A is improved in the modified procedure by subtracting the estimated count from the source from the estimate A_o. Starting from equation (2) and letting i = k gives the total expected count for interval i. The substitution of (i-1)·D is then made for t_i.

$$n_K = B\left(e^{-t_K/C} - e^{-t_{K-1}/C}\right) + A\cdot D$$

$$n_K = B\left(e^{-K\cdot D/C} - e^{-(K-1)\cdot D/C}\right) + A\cdot D$$

The equation is divided by D and solved for A.

$$\frac{n_K}{D} = \frac{B}{D}\left(e^{-K\cdot D/C} - e^{-(K-1)\cdot D/C}\right) + A$$

$$A = \frac{n_K}{D} - \frac{B}{D}\left(e^{-K\cdot D/C} - e^{-(K-1)\cdot D/C}\right)$$

$$A_o' = A_o - \frac{B}{D}\left(e^{-KD/C} - e^{-(K-1)D/C}\right) \qquad (10)$$

The modified initial estimate of A is A_o'.

The initial estimate C_o in cases when $n_1 \geqslant n_2$ is the same as for the regular estimate of C_o' except that the ratio of n_1 to n_3 is formed rather than n_1 to n_2.

$$n_1 - A\cdot D = B\left(1 - e^{-D/C}\right)$$

$$n_3 - A \cdot D = Be^{-2D/C}\left(1 - e^{-D/C}\right)$$

$$\frac{n_1 - A \cdot D}{n_3 - A \cdot D} = e^{2D/C}$$

$$C_o = \frac{2 \cdot D}{\ln\left(\frac{n_1 - A \cdot D}{n_3 - A \cdot D}\right)} \qquad K \geqslant 3 \qquad (11)$$

The restriction $k \geqslant 3$ is necessary since $n_3 - A \cdot D = 0$ when $K = 3$.