1970-04

Methods of estimating the two parameters of a life distribution characterized by a linear increasing failure rate, a + 2bt

Farmelo, Gene R.

Monterey, California; Naval Postgraduate School

https://hdl.handle.net/10945/15089

METHODS OF ESTIMATING THE TWO PARAMETERS OF A LIFE DISTRIBUTION CHARACTERIZED BY A LINEAR INCREASING FAILURE RATE, a + 2bt

Gene R. Farmelo

United States
Naval Postgraduate School

THESIS

METHODS OF ESTIMATING THE TWO PARAMETERS OF A LIFE DISTRIBUTION CHARACTERIZED BY A LINEAR INCREASING FAILURE RATE, a + 2bt

by

Gene R. Farmelo

April 1970

Methods of Estimating the Two Parameters of a Life Distribution Characterized by a Linear Increasing Failure Rate, a + 2bt

by

Gene R. Farmelo
Captain, United States Army
B.S., United States Military Academy, 1965

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
April 1970

ABSTRACT

The assumption of a linear increasing failure rate uniquely determines a life distribution which has mathematically tractable qualities. The pertinent features of this distribution are derived and listed in the paper. Three methods of estimating the two parameters of the linear increasing failure rate are derived. For each procedure a computer program is provided which performs the necessary calculations. Results utilizing simulated failure data are listed for two of the methods of parameter estimation.

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION

A. BACKGROUND

An important aspect of life testing and system reliability predictions is the proper selection of a life distribution that describes the failure data generated by the statistical model. An assumption which has wide application is that the life distribution is exponentially distributed. Several factors account for the extensive use of the exponential distribution. Benjamin Epstein and his associates have developed procedures and models which can be used when the life distribution of the failure data is assumed to be exponentially distributed. Further, the exponential distribution has mathematically tractable qualities which allow results to be obtained from relatively easy computations.

There are cases where it may be detrimental to assume that the failure data is exponentially distributed, particularly when the equipment being tested has an increasing failure rate; i.e. the instantaneous probability of failure increases with time or increased stress levels. Failure data of equipment which behaves in such a manner is not appropriately described by the exponential distribution; consequently, there are cases where the assumption of an exponential distribution may result in false conclusions about a system's reliability.

Many life distributions are characterized by an increasing failure rate and can be used to describe equipment which ages. Perhaps the most extensively used distribution in this class is the Wiebull distribution. In dealing with equipment which ages, the Wiebull distribution is a better selection than the exponential distribution.

Regardless of the parameters selected in the Wiebull distribution, the instantaneous probability of failure at time zero is zero. This precludes accurate description of equipment which has a strictly positive instantaneous probability of failure at time zero. This limitation of the Wiebull distribution is also experienced if the test is to begin after the equipment has been in operation for some time.

B. PURPOSE

The purpose of this paper is to develop and examine a life distribution function which is characterized by a linear increasing failure rate. It is felt that such a distribution will be appropriate when describing failure data of some equipment which ages with time or increased stress levels or both. The assumption of a linear increasing failure rate $(a + 2bt)$ may accurately describe equipment which has a strictly positive instantaneous probability of failure at time zero.

Accurate estimation of the parameters of the failure rate is paramount if meaningful predictions are to be made about a system's reliability. Hence, various methods of estimating the parameters were examined in the paper. All failure data was simulated by use of the computer. Computer programs (utilizing FORTRAN IV language) are provided for each method of parameter estimation examined.

II. SUMMARY AND CONCLUSIONS

A. SUMMARY

The assumption of a linear increasing failure rate uniquely defines a life distribution which has mathematically tractable qualities. The first and second central moments were derived and can be easily evaluated with the use of standardized normal tables.

Three methods of estimating the two parameters of the linear failure rate were examined. For each method and a procedure is outlined and computer programs are listed which perform the necessary calculations. Tables of results are shown for the method of moments and fitted truncated normal approaches to estimating the two parameter failure rate. These results were gained from failure data which was simulated by use of the computer. The method of generating simulated failure data from a life distribution with linear increasing failure rate is shown in Appendix C.

The third method examined was the maximum likelihood estimates of the parameters of the linear increasing failure rate. The derivation of the maximum likelihood estimates of the parameters is shown in Appendix B. Using simulated failure data, results obtained for this method were not accurate and are not listed. Limitations in obtaining the maximum likelihood estimates for the parameters of the linear failure rate are described.

B. CONCLUSIONS

The assumption of a linear increasing failure rate is more accurate than the assumption of a constant failure rate (exponential) when describing equipment which experiences aging. The important features of the life distribution characterized by a linear increasing failure rate are mathematically tractable and computational effort required is not excessive.

The correct assumption that failure data is from a life distribution with a linear failure rate has an advantage over the more general assumption of a Wiebull distribution. The advantage is that given the same failure data, tighter confidence intervals can be obtained for the parameters of the linear failure rate when compared to the confidence intervals for the parameters of the Wiebull distribution.

The method of moments approach is the simplest of the three methods examined and results (using simulated failure data) are quite accurate. A possible restriction in using the method of moments approach is that a large sample size (50 or more) is needed to attain a high degree of accuracy. However, for many types of equipment this may not be a serious restriction.

The maximum likelihood procedure is a mathematically cumbersome approach, primarily because the maximum likelihood estimates involve the solution of a n^{th} degree polynomial where n is the number of failure times contained in the data. This makes solving for the maximum likelihood estimates virtually impossible without the use of a computer. One reason for the inaccuracies in the maximum likelihood procedure is that a large sample size could not be used in order to properly simulate the life distribution.

By fitting the truncated normal distribution to failure data from a life distribution with a linear increasing failure rate, accurate estimation of the slope of the linear failure rate is obtained. Intercept estimates are not accurate, consequently once the slope is estimated, it is recommended that a relationship developed in the method of moments and maximum likelihood estimates approach be used to find an estimate of the intercept.

III. FAILURE RATE FUNCTIONS

A. CONSTANT FAILURE RATE

As mentioned earlier, the exponential distribution has been widely used in life testing studies and programs. Reasons for its wide application are that calculations are easily made and methods of parameter estimation are well known. In this section the exponential distribution will be defined and pertinent features of the distribution, as they apply to life testing, will be shown.

The exponential family of distributions has probability density functions (p.d.f.) of the following form

$$g(t;c) = \begin{cases} 0 & t < 0 \\ ce^{-ct} & t \geq 0 \end{cases}$$

where the parameter c is strictly positive. The cumulative distribution function (c.d.f.) for the exponential is

$$G(t) = \begin{cases} 0 & t < 0 \\ 1-e^{-\lambda t} & t \geq 0 \end{cases}$$

R(t) is the probability of survival to age t, which is 1 - G(t). It is assumed that components are tested until they fail; hence, R(t) is identical with component reliability. In the exponential case the reliability at time t is

$$R(t) = \begin{cases} 1 & t < 0 \\ e^{-\lambda t} & t \geq 0 \end{cases}$$

The failure rate at time t is denoted by z(t) and has a heuristic interpretation of the instantaneous probability of failure at a time t, given that the component has not failed prior to time t. z(t) is defined as the ratio of the p.d.f. to the component reliability, which in the exponential case is

$$z(t) = \frac{g(t;c)}{R(t)} = c$$

The exponential family of distribution is the only family of distributions which has a constant failure rate. The constant failure rate is frequently referred to as the "memoryless" property and has the following interpretation. The instantaneous probability of failure in some time interval $(t, t+\Delta t)$ is independent of t; consequently, when an item has been on test or in service for some time t and it has not failed, the item's instantaneous probability of failure is the same as when the item was new. For this reason care must be exercised when the assumption of an exponential distribution is made.

B. LINEAR INCREASING FAILURE RATE

The assumption of a linear increasing failure rate will be appropriate in describing some equipment which experiences aging or fatiguing. In this section the assumption of a linear increasing failure rate will be the starting point in deriving characteristics of the life distribution.

$$z(t) = \frac{-d \ln (R(t))}{dt} = a + 2bt$$

Integrating both sides of the equation yields

$$\ln(R(t)) = -\int_0^t (a+2bt)dt$$

$$e^{-\int_0^t (a+2bt)dt} = R(t)$$

By evaluating the integral, the reliability at time t is

$$R(t) = \begin{cases} 1 & t<0 \\ e^{-(at+bt^2)} & t\geq 0 \end{cases}$$

The c.d.f. is

$$F(t) = \begin{cases} 0 & t<0 \\ 1 - e^{-(at+bt^2)} & t\geq 0 \end{cases}$$

Since the distribution is strictly continuous the p.d.f. is

$$f(t;a,b) = \begin{cases} 0 & t<0 \\ (a+2bt)e^{-(at+bt^2)} & t\geq 0 \end{cases}$$

where parameters $\underline{a}$ and $\underline{b}$ are positive.

If the parameter $\underline{b}$ is allowed to be zero, the p.d.f. is recognized as the exponential distribution with parameter $\underline{a}$.

Also by observing the p.d.f. when the parameter a is allowed to be zero it is recognized as the Wiebull distribution with parameter $p = 2$ and $r = 1$ where the p.d.f. of the Wiebull distribution is defined as

$$h(t;p,r) = \begin{cases} 0 & t<0 \\ rpt^{r-1}e^{-pt^r} & t\geq 0 \end{cases}$$

The life distribution (denoted F) characterized by a linear increasing failure rate has an expected value of $e^{\frac{a^2}{4b}} \sqrt{\frac{\pi}{b}}\ \Phi\left(\frac{-a}{\sqrt{2b}}\right)$ where Φ is the c.d.f. of a standardized normal function. The second central moment is found to be $\frac{1}{b}$ (1-aE(T)). The derivation of E(T) and $E(T^2)$ are shown in Appendix A.

IV. ESTIMATES OF LINEAR FAILURE RATE

A. STATISTICAL MODEL

The assumed statistical model consisted of a number, n, of identical components which were put on test and remained on test until failure. Each failure time is a random variable T which has a density function $f_T(t)$ characterized by a linear increasing failure rate, $a + 2bt$. The data generated (failure times) was the data used to estimate the parameters of the linear failure rate. It was assumed that stress and environmental levels remained constant throughout testing.

B. METHOD OF MOMENTS

The approach for the method of moments technique of estimating the parameters is to use the failure data to estimate the first and second central moments of the distribution F. Equating these estimates to the expressions derived for the first and second central moments results in two equations with a and b as unknowns. Simultaneous solution of these two equations results in the estimates of a and b.

The first moment of the distribution is estimated by $M_1 = \frac{1}{n} \sum_{i=1}^{n} t_i$ where t_i represents the failure time of the i^{th} item in the random sample of size n. The second moment is estimated by $M_2 = \frac{1}{n} \sum_{i=1}^{n} t_i^2$.

Using these estimates and the expression for the first and second central moments the following two equations are obtained:

$$M_1 = e^{\frac{a^2}{4b}} \sqrt{\frac{\pi}{b}}\, \Phi\left(\frac{-a}{\sqrt{2b}}\right) \tag{1}$$

$$M_2 = \frac{1}{b} - \frac{a}{b} M_1$$

Simultaneous solution for a and b in these two equations result in the estimates of a and b. Since the equations are not in a closed form, an iterative technique was employed.

A computer program which performs the indicated calculations for the method of moments technique is shown at the end of this paper. Utilizing simulated failure data the technique was tested and results for interesting parameter values are shown in Table I.

C. FITTED TRUNCATED NORMAL DISTRIBUTION

The second technique examined in estimating the parameters a and b is to fit failure data to a truncated normal distribution. Motivation for this approach comes from the fact that for some parameter values of the truncated normal distribution the failure rate is nearly linear. This characteristic and other features of the truncated normal distribution are shown in an article by B. J. Flehinger and P. A. Lewis (Ref. 4). Consequently, it was felt that failure data from the life distribution F could be used to estimate the parameters μ and σ^2 of the truncated normal distribution. Utilizing these estimated parameter values the failure rate function of the truncated normal distribution could be evaluated to determine slope and intercept estimates.

The truncated normal distribution truncated at time zero has a p.d.f. defined as

$$h(t;\mu,\sigma^2) = \begin{cases} 0 & t<0 \\ \dfrac{1}{A\sqrt{2\pi\sigma^2}} \, e^{-\frac{1}{2\sigma^2}(t-\mu)^2} & t \geq 0 \end{cases}$$

where the truncated factor A is defined as

$$\int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} = \int_{-\frac{\mu}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} \, dw = \Phi(-\frac{\mu}{\sigma})$$

The c.d.f. of a truncated normal distribution is

$$H(t) = \begin{cases} 0 & t<0 \\ 1 - \Phi(\frac{t-\mu}{\sigma}) \Big/ \Phi(\frac{-\mu}{\sigma}) & t \geq 0 \end{cases}$$

and the reliability is defined as

$$R(t) = \begin{cases} 1 & t<0 \\ \Phi(\frac{t-\mu}{\sigma}) \Big/ \Phi(\frac{-\mu}{\sigma}) & t \geq 0 \end{cases}$$

Hence by definition the failure rate becomes

$$z(t) = \frac{h(t)}{R(t)} = \frac{1}{\Phi(\frac{t-\mu}{\sigma})\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} \qquad t \geq 0$$

Using simulated failure data from the life distribution F, estimates of μ and σ^2 of the truncated normal distribution were obtained by utilizing a method outlined by Cohen (Ref. 2). Basically, Cohen's method uses statistics $\bar{x}$ and s^2 which estimate the mean and variance of a normal distribution which is not truncated. A correction factor is then used on $\bar{x}$ and s^2. Once these are determined the estimates of μ and σ^2 of the truncated normal distribution are obtained by the following relationships

$$\hat{\mu} = \bar{x} + H(\bar{x} - x_o)$$

$$\hat{\sigma}^2 = s^2 - H(\bar{x} - x_o)^2$$

Where H is the correction factor which is a function of $\bar{x}$ and s^2 and x_o; x_o is the point at which the normal distribution is truncated ($x_o = 0$ for a life distribution). Cohen has tabulated the function (H) in his article.

Once maximum likelihood estimates of the truncated normal distribution are obtained from failure data, they can be used to estimate a and b in the following manner. The failure rate function of the truncated normal distribution is evaluated at time zero: this becomes the estimate of a, the intercept. Next a time t_o is selected and the failure rate function is evaluated at time t_o. The estimate of b is determined by evaluating

$$b = \frac{z(t_o) - z(0)}{zt_o}$$

This procedure was programmed and the program is supplied at the end of the paper. Results of testing the procedure utilizing simulated failure data are shown in Table II.

D. MAXIMUM LIKELIHOOD ESTIMATES

A third method of estimating the parameters a and b of the life distribution F is the maximum likelihood estimates (MLE). The mathematics of the procedure becomes quite cumbersome and MLE solutions are nearly impossible without the aid of a computer. The derivation of the MLE of a and b is shown in Appendix B. The remaining part of this section will show the results of the derivation and indicate some of the difficulties and restrictions when obtaining MLE of a and b.

Let T_i, i=1,2,...,n be a random sample of size n from the life distribution F with parameters a and b. The joint p.d.f. of this random sample is

$$f_{T_i}(t_i) = \prod_{i=1}^{n} (a+2bt_i)\, e^{-(at_i+bt_i^2)}$$

The values of b which are necessary conditions for maximizing the joint p.d.f. are the solutions to the following equation

$$\sum_{i=1}^{n} \frac{1}{n+b(2t_i \sum_{J=1}^{n} t_J - \sum_{J=1}^{n} t_J{}^2)} = 1 \qquad (1)$$

Upon expanding this equation and solving for $\hat{b}$, the MLE, the degree of difficulty becomes apparent, specifically; the solution $\hat{b}$ requires solving a polynomial of degree n (sample size). Not obvious is the fact that the coefficients of this polynomial become quite large as the sample size is increased.

Sufficient conditions for a maximum are not derived. Rather, the joint density function is to be evaluated for each positive value of b which is a root of equation (1). The value of b which maximizes the joint density functions is the MLE. Once $\hat{b}$ is obtained, $\hat{a}$ is solved for from the relationship which is derived in Appendix B, namely;

$$a \sum_{J=1}^{n} t_J + b \sum_{J=1}^{n} t_J^2 = n$$

A program which performs the calculations in obtaining the MLE of a and b is supplied at the end of the paper. The solution of the n^{th} degree polynomial is obtained by the Newton-Rhapston method. The sub-routine which calculates the roots cannot solve polynomials of degree 48 or higher, hence, this is an upper limit on the sample size which can be evaluated. The sub-routine used was IBM sub-routine RTBLSP and is not listed. A second and more restrictive constraint in using the MLE technique is that the coefficients of the polynomial become exceedingly large and exceed the number 2^{31} which is the largest number which can be stored in an IBM 360 computer. This constraint becomes active when the sample size is approximately 25.

Results of testing the MLE procedure using simulated failure data are not encouraging. It is felt that a major factor for the inaccurate estimates is that a large enough sample size could not be generated in order to accurately simulate the life distribution. Results are not shown.

E. TABLES OF RESULTS

Tables I and II contain the results of estimating a and b using the method of moments (Table I) and fitted truncated normal (Table II) techniques. For each combination of parameters a and b, the programs would use the value azero and bzero to generate simulated failure times from the life distribution F. Once failure data was generated, the program would estimate the parameters from the simulated failure data (failure times). These estimates are designated ahat and bhat. For example, the first case in Table I shows that bzero and azero are .005 and .010 respectively. Estimates of these parameters are .005 and .013. For each case 100 failure times were generated and used in estimating the parameters.

The results in Table I indicate that the method of moments is an accurate way of estimating the parameters a and _b_. Estimates of the slope (bhat) are exact for the ten cases with azero equal to .010. As azero increases to .030 estimates of the slope become high by as much as 16% in cases 26 and 27. Intercept estimates (ahat) are not as accurate as the slope estimates, however, they are acceptable. Estimates of the intercept are better when azero is .030 as compared to the smaller value of 0.10 for azero.

Table II contains the results of the fitted truncated normal technique. Slope estimates are more accurate than intercept estimates. It can be seen that in many cases slope estimates (bhat) are exact. Bhat is 18% higher than bzero in the worst case (case 30). Intercept estimates (ahat) are not accurate and in many cases the estimates are off by as much as a factor of ten.

TABLE I

METHOD OF MOMENTS

n = 100

CASE	BZERO	BHAT	AZERO	AHAT	CASE	BZERO	BHAT	AZERO	AHAT
1	.005	.005	.010	.013	16	.030	.032	.020	.022
2	.010	.010	.010	.014	17	.035	.038	.020	.021
3	.015	.015	.010	.014	18	.040	.043	.020	.028
4	.020	.020	.010	.015	19	.045	.047	.020	.027
5	.025	.025	.010	.015	20	.050	.052	.020	.027
6	.030	.030	.010	.016	21	.005	.007	.030	.029
7	.035	.035	.010	.017	22	.010	.013	.030	.030
8	.040	.040	.010	.017	23	.015	.018	.030	.034
9	.045	.045	.010	.019	24	.020	.024	.030	.030
10	.050	.050	.010	.019	25	.025	.029	.030	.028
11	.005	.006	.020	.022	26	.030	.035	.030	.032
12	.010	.012	.020	.028	27	.035	.040	.030	.031
13	.015	.017	.020	.024	28	.040	.044	.030	.030
14	.020	.022	.020	.022	29	.045	.049	.030	.030
15	.025	.027	.020	.022	30	.045	.054	.030	.033

TABLE II

FITTED NORMAL APPROXIMATION TECHNIQUE

$n = 100 \quad t_o = 1.5$

CASE	BZERO	BHAT	AZERO	AHAT	CASE	BZERO	BHAT	AZERO	AHAT
1	.005	.005	.005	.025	16	.030	.033	.015	.064
2	.010	.010	.005	.035	17	.035	.038	.015	.068
3	.015	.015	.005	.042	18	.042	.046	.015	.073
4	.020	.020	.005	.046	19	.045	.051	.015	.077
5	.025	.027	.005	.050	20	.050	.056	.015	.081
6	.030	.031	.005	.056	21	.005	.006	.025	.043
7	.035	.036	.005	.060	22	.010	.011	.025	.053
8	.040	.044	.005	.064	23	.015	.017	.025	.059
9	.045	.049	.005	.068	24	.020	.023	.025	.063
10	.050	.054	.005	.071	25	.025	.028	.025	.070
11	.005	.006	.015	.035	26	.030	.033	.025	.074
12	.010	.010	.015	.043	27	.035	.038	.025	.079
13	.015	.016	.015	.048	28	.040	.044	.025	.082
14	.020	.020	.015	.055	29	.045	.050	.025	.083
15	.025	.027	.015	.060	30	.050	.059	.025	.088

APPENDIX A

DERIVATION OF MEAN AND $E(T^2)$

In order to determine the first central moment of the life distribution F a well known theorem was used, namely,

$$E(T) = \int_0^\infty t\ f(t)\ dt = \int_0^\infty R(t)\ dt$$

In the case of the distribution F, the expected value is

$$\int_0^\infty e^{-(at+bt^2)}\ dt$$

by factoring a b and completing the square in the exponent, the integral becomes

$$e^{\frac{a^2}{4b}} \int_0^\infty e^{-b(t+\frac{a}{2b})^2}\ dt$$

Except for the constant term this integral has the general form of a normal density function. Multiplying by one in the form $\sqrt{\frac{\pi}{b}} \Big/ \sqrt{\frac{\pi}{b}}$ yields

$$e^{\frac{a^2}{4b}} \sqrt{\frac{\pi}{b}} \int_0^\infty \sqrt{\frac{b}{\pi}}\ e^{-b(t+\frac{a}{2b})^2}\ dt$$

where now the integrand is recognized as the normal density function with mean $-\frac{a}{2b}$ and variance $\frac{1}{2b}$. Making the necessary transformation to standardize the normal density and recalling that

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw$$

the mean of the distribution F becomes

$$E(T) = e^{\frac{a^2}{4b}} \sqrt{\frac{\pi}{b}} \left(1 - \Phi\left(\frac{a}{\sqrt{2b}}\right)\right)$$

Since the standardized normal function is symmetric about zero, the expected value can be written as

$$e^{\frac{a^2}{4b}} \sqrt{\frac{\pi}{b}} \Phi\left(-\frac{a}{\sqrt{2b}}\right)$$

The second central moment is by definition

$$E(T^2) = \int_0^{\infty} t^2(a+2bt)\ e^{-(at+bt^2)} dt$$

which can be evaluated by the integration by parts method. Let $u = t^2$ and $dV = (a+2bt)e^{-(at+bt^2)}dt$, hence $du = 2tdt$ and $V = -e^{-(at+bt^2)}$

$$E(T^2) = t^2\ e^{-(at+2bt)} \Big|_0^{\infty} + \int_0^{\infty} 2t\ e^{-(at+bt^2)} dt$$

By using L'Hospital's rule, the first term on the right side of the equation may be shown to be zero, hence

$$E(T^2) = \int_0^\infty 2t\ e^{-(at+bt^2)}\ dt$$

$$= \frac{1}{b}\int_0^\infty 2bt\ e^{-(at+bt^2)}dt + \int_0^\infty \left(\frac{a}{b} - \frac{a}{b}\right)e^{-(at+bt^2)}dt$$

$$= \frac{1}{b}\int_0^\infty (a+2bt)\ e^{-(at+bt^2)} - \frac{a}{b}\int_0^\infty e^{-(at+bt^2)}\ dt$$

The first integral is one, since it is the density function of F and the second integral is the E(F) derived above. Therefore

$$E(T^2) = \frac{1}{b} - \frac{a}{b}\ E(T)$$

APPENDIX B

MAXIMUM LIKELIHOOD ESTIMATES

Let T_i, $i = 1,2,\ldots,n$ be a random sample of failure times from the life distribution F which has a linear increasing failure rate. The joint p.d.f. of the failure times is the product of the marginal density functions

$$f_{T_i}(t_i;a,b) = \prod_{i=1}^{n} (a+2bt_i)\, e^{-(at_i+bt_i^2)} \qquad (1)$$

The logarithm of the joint p.d.f. is

$$\ln(f_{T_i}(t_i)) = L = \sum_{i=1}^{n} \ln(a+2bt_i) - a\sum_{i=1}^{n} t_i - b\sum_{i=1}^{n} t_i^2$$

The necessary conditions for a maximum are found by taking the partial derivatives and equating them to zero.

$$\frac{\partial L}{\partial a} = \sum_{i=1}^{n} \frac{1}{a+2bt_i} - \sum_{i=1}^{n} t_i = 0 \qquad (2)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{n} \frac{2t_i}{a+2bt_i} - \sum_{i=1}^{n} t_i^2 = 0 \qquad (3)$$

Multiplying equation (2) by a and equation (3) by b and adding the two equations results in

$$n = a\sum_{i=1}^{n} t_i + b\sum_{i=1}^{n} t_i^2 \qquad (4)$$

Solving for a and substituting into equation (2) gives an equation in b which will yield necessary conditions for maximizing the likelihood function. The resulting equation after simplification is

$$\sum_{i=1}^{n} \frac{1}{n+b(2t_i \sum_J t_J - \sum_J t_J^2)} = 1 \qquad (5)$$

When solving for values of b which are roots to this equation, it becomes apparent that it is necessary to solve a n^{th} degree polynomial in b. For example, if n = 3, the equation after expanding and simplifying became

$$t_1t_2t_3b^3 + 2(t_1t_2+t_1t_3+t_2t_3)b^2 + 3(t_1+t_2+t_3)b = 0$$

The equation for a general n can be written

$$\sum_{k=0}^{n-1} \binom{n}{n-k}_{t_i} b^{n-k} (n^k - kn^{k-1}) = 0 \qquad (6)$$

where $\binom{n}{n-k}_{t_i}$ denotes the sum of the product of combinations of n failure times taken n - k numbers at a time, e.g., if n is four and k is one it denotes

$$\binom{4}{3}_{t_i} = t_1t_2t_3 + t_1t_2t_4 + t_1t_3t_4 + t_2t_3t_4$$

Another example is if n is six and k is zero

$$\binom{6}{6}_{t_i} = t_1t_2t_3t_4t_5t_6$$

The n^{th} degree polynomial in b which is equation (5) or (6) can be solved by the use of an IBM program which is called RTPLSB. This program utilized the Newton-Rhapston method and solves polynomials up to degree 48.

The positive roots of equation (5) can be substituted into equation (1) and the root which maximizes equation (1) becomes the maximum likelihood estimate for b. The MLE of a is determined from equation (4) with $\hat{b}$ being substituted for b.

APPENDIX C

FAILURE TIME SIMULATION

In order to obtain simulated failure times from a life distribution which has a linear increasing failure rate, a theorem which is stated in Lindgren (Ref. 5, p. 274) was used: if T is a continuous random variable with distribution function $G(x)$, then $U = G(T)$ and $G(u) = P[G(T) \leq u] = u$ for $0 \leq u \leq 1$. That is, U is distributed uniformly over the interval (0,1). It can also be shown that if U is uniform (0,1) then 1 - U is also uniform (0,1). Consequently, in the case of the distribution with the linear increasing failure rate the following relationship is true

$$
\begin{aligned}
P[R(T) > x] &= x \quad & 0 \leq x \leq 1 \\
&= 1 \quad & x < 0 \\
&= 0 \quad & x > 1
\end{aligned}
$$

where T is a R.V. with distribution $F(t)$, and, $R(t) = 1 - F(t)$.

If x_o is a random number from the interval (0,1) then a simulated failure time can be obtained by solving

$$R(T) = e^{-(a_o T + b_o T^2)} = x_o$$

where a_o and b_o are pre-selected parameter values. Solving for T yields

$$T = \frac{-a_o \pm \sqrt{a_o^2 - 4b_o \ln(x_o)}}{2b_o}$$

Since $\ln x_o$ is negative and the parameters a_o and b_o are always positive there will be one positive and one negative solution to the equation. The negative failure time can be disregarded.

METHOD OF MOMENTS TECHNIQUE

```
      IMPLICIT REAL*8(A-H),REAL*8(P-Z)
      DIMENSION PHE(40),T(500)
      DATA PHE/.4801,.4404,.4013,.3632,.3264,.2912,.2578,
     1.2266,.1917,.1711,.1469,.1251,.1056,.0885,.0735,.0606,
     1.0495,.0401,.0322,.0256,.0202,.0158,.0122,.0100,.0080,
     1.0054,.0040,.0029,.0022,.0016/
100   READ(5,10) AZERO,BZERO
10    FORMAT(2F5.4)
      IF(AZERO.GE..50)GO TO 1000
C
C   ***CALCULATE THE EXPECTED VALUE OF T
C
      EXPONT=AZERO**2/(4.*BZERO)
      ARGNOR =AZERO**2/(2.*DSQRT(BZERO))
      INORM=100.*ARGNOR/10
      IF(INORM.LT.1)INORM=1
      IF(INORM.GT.30)INORM=30
      TMEAN=DEXP(EXPONT)*DSQRT(3.1417/BZERO)*PHE(INORM)
C
C   ***CALCULATE THE ESTIMATED 1ST AND 2ND MOMENTS
C
      IX=65539
      SUMM=0.
      N=100
      DO 200 I=1,N
      CALL RANDU(IX,IY,YFL)
      IX=IY
      TEMP=DLOG(YFL)
      ARG=AZERO*AZERO-4.*BZERO*TEMP
      RAD=DSQRT(ARG)
      T(I)=(-AZERO+RAD)/(2.*BZERO)
200   CONTINUE
      XM1=0.
      XM2=0.
      DO 300 I=1,N
      XM1=XM1+T(I)
      XM2=XM2+T(I)*T(I)
300   CONTINUE
      XN=N
      XM1=XM1/XN
      XM2=XM2/XN
C
C   ***CALCULATE EXPECTED VALUE OF T SQUARED
C
      SECMOM=1./BZERO-(AZERO*TMEAN)/BZERO
C
C   ***SOLVE FOR BHAT
C
      BHAT=.001
400   CONTINUE
      DUMM=((1.-BHAT*XM2)/XM1)**2
      ARGPHE=DUMM/(2.*DSQRT(BHAT))
      IARG=100.*ARGPHE/10
      IARG=IARG+1
      IF(IARG.GT.30)IARG=30
      IF(IARG.LT.1)IARG=1
      POWER=DUMM/(4.*BHAT)
      IF(POWER.GT.160.)POWER=160.
      EQN=DEXP(POWER)*DSQRT(3.1417/BHAT)*PHE(IARG)
      IF(EQN.LT.XM1)GO TO 500
      BHAT=BHAT+.001
      IF(BHAT.GE.25)GO TO 999
      GO TO 400
500   CONTINUE
C
```

```
C     ***CALCULATE AHAT
C
        AHAT=(1.-BHAT*XM2)/XM1
C
        GO TO 100
999     WRITE(6,30)
30      FORMAT('BHAT REACHED .25')
1000    STOP
        END
```

FITTED TRUNCATED NORMAL TECHNIQUE

```
      IMPLICIT REAL*8(A-H),REAL*8(O-Z)
      DIMENSION T(500),PHE(30),CORR(88)
101   READ(5,10)AZERO,BZERO
10    FORMAT(2F5.4)
      DATA PHE/.4801,.4404,.4013,.3632,.3264,.2912,.2578,
     1.2266,.1917,.1711,.1469,.1251,.1056,.0885..0735,.0606,
     1.0495,.0401,.0322,.0256,.0202,.0158,.0122,.0100,.0080,
     1.0054,.0040,.0029,.0022,.0016/
      DATA CORR/.0001,.0001,.0001,.0001,.0001,.0001,.0001
     1,.0001,.0001,.0011,.0014,.0024,.0034,.0052,.0070
     1,.0091,.0099,.0148,.0183,.0225,01268,.0319,.0378,.0445
     1,.0514,.0598,.0683,.0778,.0856,.1000,.1108,.1224,.1377
     1,.1517,.1635,.1864,.2042,.2234,.2439,.2705,.2945,.3201
     13473,.3764,.4074,.4402,.4751,.5121,.5593,.6012,.6453
     1,.6918,.7509,.8030,.8577,.9268,.9875,1.064,1.144
     1,1.230,1.319,1.413,1.512,1.633,1.742,1.881,1.981
     1,1.999,2.301,2.411,2.651,2.901,3.171,3.321,3.361
     1,3.921,4.251,4.961,5.331,5.942,6.361,7.031,7.741
     1,8.491,9.281,10.41,11.61/
      N=500
      IF(AZERO.GT..45)GO TO 1000
C
C *** GENERATE FAILURES USING PARAMETERS AZERO AND BZERO
C
      IX=65539
      DO 100 I=1,N
      CALL RANDU(IX,IY,YFL)
      IX=IY
      TEMP=DLOG(YFL)
      ARG=AZERO*AZERO-4.*BZERO*TEMP
      RAD=DSQRT(ARG)
      T(I)=(-AZERO+RAD)/(2.*BZERO)
100   CONTINUE
C
C *** CALCULATE THE SAMPLE MEAN AND VARIANCE
C
      XBAR=0.
      SBAR=0.
      XN=N
      DUMM=0.
      DO 200 I=1,N
      XBAR=XBAR+T(I)
200   CONTINUE
      XBAR=XBAR/XN
      DO 300 I=1,N
      DUMM=(T(I)-XBAR)**2
      SBAR=SBAR+DUMM
      DUMM=0.
300   CONTINUE
      SBAR=SBAR/XN
C
C ** CALCULATE THE ESTIMATE CORRECTION FACTOR
C
      THETA=SBAR/(XBAR*XBAR)
      JARG=1000.*THETA/10
      XMEAN=XBAR-CORR(JARG)*XBAR
      XVAR=SBAR+CORR(JARG)*XBAR*XBAR
      STDEV=DSQRT(SBAR)
C
C *** EVALUATE FAILURE RATE FUNCTION AT TIME ZERO(AHAT)
C
      TIME=0.
1     CONTINUE
      ARGPHE=(TIME-XMEAN)/STDEV
      IARG1= 100.*ARGPHE/10
```

```
      IARG=IABS(IARG1)
      IARG=IARG+1
      IF(IARG.LT.1.AND.IARG.GT.0)IARG=1
      IF(IARG.GE.30)IARG=30
      XNORM=PHE(IARG)
      IF(IARG1.LT.0)GO TO 3
2     CONTINUE
      EXARG=-1.*(TIME-XMEAN)**2/(2.*XVAR)
      ZFUNCT=(1.0/(XNORM*DSQRT(6.28*XVAR)))*DEXP(EXARG)
      IF(TIME.GT.0)GO TO 4
      AHAT=ZFUNCT
      TIME=1.
      GO TO 1
3     XNORM= XNORM + .5
      GO TO 2
4     CONTINUE
C
C *** CALCULATE ESTIMATE OF BZERO
C
      BHAT=(ZFUNCT-AHAT)/TIME
      GO TO 101
1000  STOP
      END
```

MAXIMUM LIKELIHOOD ESTIMATES TECHNIQUE

```
      IMPLICIT REAL*8(A-H),REAL*8(O-Z)
      DIMENSION A(54),U(54),V(54),CONV(54),T(50),TRA(54)
     1TRB(54),AA(50),CORR(50),ICOUNT(40),XJOINT(50)
1     READ(5,10)N,AZERO,BZERO
      IF(N.GE.50)GO TO 1000
      IFLAG=0
      A(N+1)=0.
      AA(N+1)=0.
C
C   ***GENERATE FAILURE TIMES
C
      IX=659
      DO 100 I=1,N
      CALL RANDU(IX,IY,YFL)
      IX=IY
      TEMP=DLOG(YFL)
      ARG=AZERO*AZERO-4.*BZERO*TEMP
      RAD=DSQRT(ARG)
      T(I)=(-AZERO+RAD)/(2.*BZERO)
100   CONTINUE
C
C   ***CALCULATE SUM AND SUM OF SQUARED FAILURE TIMES
C
      SUM=0.
      SUMSQ=0.
      DO 200 I=1,N
      SUM=SUM+T(I)
      SUMSQ=SUMSQ+T(I)**2
200   CONTINUE
C
C   ***CALCULATE COEFFICIENTS FOR AHAT AND BHAT
C
      SUMA=0.
      SUMB=0.
      PRODA=1.
      PRODB=1.
      DO 300 I=1,N
      TRB(I)=2.*T(I)*SUM-SUMSQ
      TRA(I)=(SUMSQ/(2.*T(I)))-SUM
      SUMA=SUMA+TRA(I)
      SUMB=SUMB+TRB(I)
      PRODB=PRODB*TRB(I)
      PRODA=PRODA*TRA(I)
300   CONTINUE
      A(1)=PRODB
      AA(1)=PRODA
      DO 400 J=1,N
      K=N-J
      IF(K.EQ.1)GO TO 20
      CALL CHOOZ(N,K,TRB,TRA,X,Z)
      A(J+1)=X
      AA(J+1)=Z
400   CONTINUE
      GO TO 30
20    CONTINUE
30    A(N)=SUMB
      AA(N)=SUMA
C
C   **CALCULATE CORRECTION FACTOR FOR COEFFICIENTS
C
      LN=N-1
      DO 500 I=1,LN
      M=I-1
      CORR(I+1)=N**I-I*N**M
500   CONTINUE
```

```
      CORR(1)=1.
      DO 600 I=1,N
      A(I)=A(I)*CORR(I)
      AA(I)=AA(I)*CORR(I)
600   CONTINUE
C
C   ***CALCULATE ROOTS OF NTH DEGREE POLYNOMIAL
C
      CALL RTPLSB(N,A,U,V,CONV,IER)
      WRITE(6,40)IER
      WRITE(6,60)(U(I),V(I),CONV(I),I=1,N)
      IF(IFLAG.EQ.1)GO TO 1
C
C *** CALCULATE AHAT
C
      IFLAG=1
      DO 101 I=1,N
      A(I)=AA(I)
101   CONTINUE
      GO TO 202
C
C   ***EVALUATE JOINT P.D.F. FOR EACH POSITIVE ROOT
C
555   CONTINUE
      SUMM=0.
      POWER=0.
      DO 700 I=1,N
      IF(U(I).LE.0.)GO TO 800
      DO 800 J=1,N
      POWER=-1.*(AZERO*T(J)+U(I)*T(J)**2)
      IF(POWER.GE.170.)POWER=170.
      SUMM=SUMM+(AZERO+2.*U(I)*T(J))*DEXP(POWER)
      POWER=0.
800   CONTINUE
      XJOINT(I)=SUMM
      SUMM=0.
700   CONTINUE
      GO TO 1
1000  STOP
      END
      IMPLICIT REAL*8(A-H),REAL*8(O-Z)


      SUBROUTINE CHOOZ(N,K,TRB,TRA,X,Z)
      DIMENSION TRB(45),TRA(45),IA(40),IB(40)
      NN=N
      KK=K
      L=1
      X=0.
      Z=0.
      IA(1)=1
      PRA=1.
      PRB=1.
3     IB(L)=IA(L)
      IF(L.EQ.KK)GO TO 12
      IA(L+1)=IA(L)+1
      IF(IA(L+1).EQ.NN+1)GO TO 9
      L=L+1
      GO TO 3
9     L=L-1
      IF(L.EQ.0)GO TO 21
      GO TO 13
12    CONTINUE
      DO 40 I=1,KK
      PRA=PRA*TRA(IB(I))
      PRB=PRB*TRB(IB(I))
40    CONTINUE
      X=X+PRB
      Z=Z+PRA
      PRB=1.
      PRA=1.
13    IA(L)=IA(L)+1
```

```
      IF(IA(L).EQ.NN+1)GO TO 9
      GO TO 3
21    RETURN
      END
```

LIST OF REFERENCES

1. Barlow, Richard E. and Prochan, Frank, Mathematical Theory of Reliability, John Wiley and Son, Inc., 1965.

2. Cohen, A. Clifford Jr., "Simplified Estimates for the Normal Distribution When Samples are Singly Censored or Truncated," Technometrics, Vol. 1, No. 3, August, 1959.

3. Epstein, Benjamin and Sobel, Milton, "Life Testing," Journal of American Statistical Association, Vol. 48, No. 263, September, 1953.

4. Flehinger, B. J. and Lewis, P. A., "Two-Parameter Lifetime Distributions for Reliability Studies of Renewal Processes," IBM Journal of Research and Development, Vol. 3, No. 1, January, 1959.

5. Lindgren, B. W., Statistical Theory, MacMillan Company, 1962.

6. Lloyd, D. K. and Lipow, Myron, Reliability: Management, Methods and Mathematics, Prentice Hall, Inc., 1962.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	20
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Civil Schools Branch Office of Personnel Operations Washington, D. C. 20315	1
4.	Assoc.Professor W. Max Woods, Code 55Wo Department of Operations Analysis Naval Postgraduate School Monterey, California 93940	4
5.	Assoc. Professor D. Barr, Code 55Bn Department of Operations Analysis Naval Postgraduate School Monterey, California 93940	1
6.	CPT Gene R. Farmelo 10 Khakum Wood Greenwich, Conn. 06830	4
7.	Department of Operations Analysis Naval Postgraduate School Monterey, California 93940	1

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

METHODS OF ESTIMATING THE TWO PARAMETERS OF A LIFE DISTRIBUTION CHARACTERIZED BY A LINEAR INCREASING FAILURE RATE, a + 2bt

4. DESCRIPTIVE NOTES *(Type of report and, inclusive dates)*

Master's Thesis; April 1970

5. AUTHOR(S) *(First name, middle initial, last name)*

Gene R. Farmelo, Captain, United States Army

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 1970	43	6

8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

The assumption of a linear increasing failure rate uniquely determines a life distribution which has mathematically tractable qualities. The pertinent features of this distribution are derived and listed in the paper. Three methods of estimating the two parameters of the linear increasing failure rate are derived. For each procedure a computer program is provided which performs the necessary calculations. Results utilizing simulated failure data are listed for two of the methods of parameter estimation.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
RELIABILITY						
LINEAR INCREASING FAILURE RATE						
METHOD OF MOMENTS						
LIFE DISTRIBUTIONS						
MAXIMUM LIKELIHOOD ESTIMATES						
CONSTANT FAILURE RATE						
WIEBULL DISTRIBUTION						
EXPONENTIAL DISTRIBUTION						
TRUNCATED NORMAL DISTRIBUTION						

28FEB72 19793
20 FEB 73 19319
3 OCT 73 S11115

Thesis
F226
c.1

118488

Thesis
F226
c.1

Farmelo

Methods of estimating the two parameters of a life distribution characterized by a linear increasing failure rate, a+2bt