SIMULATION FOR TESTS ON THE VALIDITY OF THE ASSUMPTION THAT THE UNDERLYING DISTRIBUTION OF LIFE IS EXPONENTIAL.

by

ANJO THOPPIL

B.E. (Production Engineering), University of Madras,

Madras, India, 1980

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements of the degree

MASTER OF SCIENCE

KANSAS STATE UNIVERSITY

Manhattan, Kansas

Approved by

Doris Grosh

Major Professor

TABLE OF CONTENTS

THIS BOOK CONTAINS NUMEROUS PAGES WITH THE ORIGINAL PRINTING BEING SKEWED DIFFERENTLY FROM THE TOP OF THE PAGE TO THE BOTTOM.

THIS IS AS RECEIVED FROM THE CUSTOMER.

THIS BOOK CONTAINS NUMEROUS PAGES WITH DIAGRAMS THAT ARE CROOKED COMPARED TO THE REST OF THE INFORMATION ON THE PAGE.

THIS IS AS RECEIVED FROM CUSTOMER.

ACKNOWLEDGEMENT

The author wishes to express his sincere appreciation to his major professor, Dr. Doris Grosh for her invaluble advice, creative guidance, and unfailing support during the course of this work; to his committee members, Dr.L.E.Grosh and Dr.A.P.Mathews for their advice and to the Department of Statistics for being so helpful on the many occasions that I had to be there.

SECTION 1A.

INTRODUCTION

From its beginnings in the early part of the eighteenth century, the growth of mathematical statistics has proved to be a powerful tool in the analysis of problems in the face of uncertainty. Although statistics is primarily used for the discriptive character and value that it imparts to a mass of unmanageble data, the science of inferential statistics has given character and definition to the theories of probability and estimation.

Fundamental to this study is the immediate application of probabilistic methods to two major areas -- those of Reliability and Quality Control.

Quality Control deals primarily with the inspection of new products, especially before they are put into service. It frequently addresses questions like: what fraction of the product is good; how much of the lot is to be inspected; and what fraction of the lot is acceptable?

Reliability deals primarily with the behaviour of components after they have been put into service. It often addresses problems like: how long can a component or unit be expected to last; what is the probabilty that a 500 hour mission, say, will succeed?

Stated in simple terms "Reliability is the capability of an equipment not to break down in operation and a measure of an equipment's reliability is the frequency with which failures occur in time". Reliability predicts mathematically the equipment's behaviour under expected operating conditions. It expresses in numbers the chance of the equipment to operate without failure for a given length of time in an environment for which it was assigned. By its most primitive definition "Reliability is the probability that no failure will occur in a given time interval of operation". A more formal definition of reliability states "Reliability is the probability of a device performing its purpose adequately for the period of time intended under the operating conditions encountered".

One of the prominent areas of reliability deals with life testing. The primary goal in life testing is estimating the mean life of a component, and a life test is usually conducted by subjecting one or more identical components to a set of operating conditions and noting the period of satisfactory performance, that is, the time to failure of each component. Failure is defined as the state in which the component fails to perform satisfactorily. Most of the literature in this area is based on CFR (constant failure rate) components, for reasons of mathematical tractability.

Of prime importance in the study of life testing is the exponential distribution, a continuous distribution with a constant hazard rate which is independent of time. The hazard rate is also called the instantaneous failure rate and its constancy means that the probability of failure is independent of age -- i.e. an old equipment that is still operating is just as good as a new one. The exponential failure law is used to predict the probability of survival of a part as a function of time. Its density function is of the form:

$$f(t) = \lambda e^{-\lambda t} \qquad t \geq 0 \qquad (1a)$$

where λ is the constant called the chance failure rate and 't' is the operating time. In its alternate form the model for component life 't' assumes the form :

$$f(t) = \frac{1}{\theta}\ e^{-(t/\theta)} \qquad t \geq 0 \qquad (2a)$$

where $\lambda = \frac{1}{\theta}$.

The area of life testing has been intensively explored by mathematical statisticians like Benjamin Epstein and Mark Sobel and many of their papers address the question of whether in the case of life-test data one is justified in assuming that the underlying distribution of life is exponential. In particular Epstein presents several procedures for evaluating the validity of the exponential assumption. His work can be divided into three major areas :

(1) A graphical procedure, one which is particularly useful if there are huge amounts of data.

(2) A test for abnormally early first failures.

(3) A test for abnormally late first failures.

Several techniques have been used by Epstein in each of these cases. However, the examples presented seem to be obviously contrived to fit the rejection tests that he recommends. In most of his examples it is obvious that a statistical test is not required to categorize a particular set of data and that even a relatively inexperienced observer can look at the data and come to a conclusion about its nature without resort to any of the statistical tests provided. The purpose of this report has been to investigate whether a visual examination of a set of data can reveal the nature of the distribution and aid the relatively inexperienced observer to categorize the data under the types mentioned without having to resort to any of these tests. In addition, a special case will be considered wherein a sample arising out of an exponential distribution with a certain mean life is contaminated or overlayed by a sample arising out of another exponential distribution having a different mean life. Under each of these cases a minimum of 30 samples has been generated, typical sample sizes being 5, 10, 20 and 50. Altogether, about 200 sets of data and an equal number of plots have been generated using simulation techniques. This report, however, contains only a few representative cases for the purposes of discussion and analysis. The intention has also been to use the data and plots for demonstration purposes in a typical reliability course where the inexperienced student can acquire the feel for the exponential distribution and the different considerations involved. Whereas the data, the statistical tests used and the plots will be described in the body of the report, the simulation programs themselves are presented in separate appendices. Appropriate explanations for each of the programs has also been included.

SECTION 1B

MATHEMATICAL BACKGROUND

It is critical to the reliability practitioner to make the correct choice of a failure law model for describing data. An attempt is being made here to describe the various important models and their relationships that apply to components subject to failure in a life-test. These are :

1) The Negative Exponential Distribution.

2) The Gamma Distribution

3) The Chi-Square Distribution and

4) Snedecor's *F-Distribution*

The Negative Exponential Distribution

Throughout its development the theory of reliability has been based heavily on the assumption of the negative exponential failure law. This is because of its mathematical tractability more than anything else. Its density function is of the form

$$f(t) = \lambda e^{-\lambda t} \qquad t \geq 0 \qquad (1b)$$

where λ is the constant called the chance failure rate and t is the operating time. In its alternate form the model for component life t assumes the form

$$f(t) = \frac{1}{\theta} e^{-(t/\theta)} \qquad (2b)$$

where θ is refered to as the mean life of the component. The cumulative distribution is

$$F(t) = \int_0^t \lambda e^{-\lambda x}\, dx = 1 - e^{-\lambda t} \qquad (3b)$$

Reliability is defined as the probability of failure after time t, so that in the exponential case we have

$$R(t) = \int_t^{\infty} f(w)\, dw = e^{-\lambda t} \qquad (4b)$$

or equivalently

$$R(t) = e^{-t/\theta}.$$

The CFR(constant failure rate) law is one of the most important in raliability work. In this work it will be convenient to use the notation

$$X \sim NGEX\ (\lambda) \quad \text{or} \quad X \sim CFR\ (\lambda) \qquad (5b)$$

to stand for a random variable which is distributed exponentially. The mean and the variance of the exponential distribution are

$$\mu = 1/\lambda \quad \text{and} \quad \sigma^2 = 1/\lambda^2$$

<u>The Gamma Distribution</u>

The gamma distribution derives its name from the well known gamma function defined by

$$\Gamma\ (\alpha) = \int_0^{\infty} x^{\alpha-1}\, e^{-x}\, dx \qquad \alpha>0 \qquad (6b)$$

The gamma distribution is defined by the density function

$$f(x) = Kx^{\alpha-1} e^{-x/\beta} \qquad \alpha,\beta>0,\ x>0 \qquad (7b)$$

where $K = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}$ is a normalizing constant chosen to give unit area under the curve. The special gamma distribution for which α is called the Negative Exponential Distribution.

The Chi-Square Distribution

The chi-square distribution is a special case of the gamma distribution. Let $X \sim G(\alpha,\beta)$ so that

$$f(x) = [\Gamma(\alpha)\beta^{\alpha}]^{-1} x^{\alpha-1} e^{-x/\beta} \qquad (8b)$$

now let $y=cx$, so that $\frac{dx}{dy} = \frac{1}{c}$. Then

$$g(y) = [\Gamma(\alpha)\ \beta^{\alpha}]^{-1} (\tfrac{y}{c})^{\alpha-1} e^{-\frac{(y/c)}{\beta}} \ (1/c)$$

$$= [\Gamma(\alpha)\ (c\beta)^{\alpha}]^{-1} y^{\alpha-1} e^{-\frac{y}{c\beta}} \qquad (9b)$$

The result therefore is that if $X \sim G(\alpha,\beta)$ then $Y=CX \sim G(\alpha,c\beta)$. The chi-square distribution is obtained by letting $\alpha = \nu/2$ and $\beta = 2$, where ν is a positive integer.

$$\text{If } X \sim G(\alpha,\beta) \text{ then } Y = \frac{2X}{\beta} \sim G(\alpha,2) \sim \chi^2(2\alpha)$$

The probability density so obtained is called the chi-square distribution with ν degrees of freedom.

The continuous random variable X has a chi-square distribution with ν_1 degrees of freedom, if its density function is given by

$$f(x) = \frac{1}{2^{\nu/2}\,\Gamma(\nu/2)}\; x^{\nu/2-1}\, e^{-x/2} \qquad (10b)$$

where ν is a positive integer. The mean and the variance of the chi-square distribution are given by

$$\mu = \nu \text{ and } \sigma^2 = 2\nu$$

<u>Snedecor's F-Distribution</u>

One of the most important distributions in applied statistics is the F-distribution. The statistic F is defined to be the ratio of two independent chi-square random variables, each divided by their degrees of freedom. Hence we can write

$$F = \frac{u/\nu 1}{v/\nu_2}$$

where u, v are independent random variables with ν_1 and ν_2 degrees of freedom, respectively. The distribution of the random variable

$$F = \frac{u/\nu_1}{v/\nu_2}$$

is given by

$$h(f) = \frac{\Gamma[(\nu_1 + \nu_2)/2] \quad (\nu_1/\nu_2)^{\nu_1/2}}{\Gamma(\nu_1/2) \quad \Gamma(\nu_2/2)} \quad \frac{f^{\,\nu_1/2 - 1}}{(1+\nu_1\; f/\nu_2)^{(\nu_1+\nu_2)/2}} \qquad (11b)$$

$$= 0 \text{ elsewhere}$$

In life testing, whether using single or multiple socket testing, the reliability engineer who is testing a good product may have a very long wait before the test results are available for analysis. This is because of the heavy tail of the exponential distribution. The heaviness of the tail means that, when a fairly large number of items are put on a life test, a few of them may last for a very long time and delay the test results inordinately. One dare not leave them out of the calculations, since to do so would bias the θ estimate. What is needed is a method of terminating testing before all elements have failed, and of correctly incorporating the incomplete information. The Type 2 censuring method is used with stopping rule :

" Put n items on test at time t=0 in separate testing stations(sockets). Record failure times. Terminate testing when 'r'(a predetermined number) of the items have failed."

Let t_i represent the i^{th} random failure time generated by the exponential distribution. Each t_i is a realization(sample observation) of a common random variable t for which we can either write $t \sim$ NGEX or $t \sim G(1,\theta)$. Let τ_i represent the i^{th} ordered failure time. Suppose that for the ease of notation the symbols $Z_1, Z_2, ---, Z_r$ are used for the ordered failure times $\tau_1, \tau_2, ---, \tau_r$. As soon as Z_r is observed, testing is terminated and there are (n-r) unfailed items with an unknown amount of residual life left in them. An estimate of mean life is required and the process is to form the joint density function of the first r order statistics.

$$g(Z_1, Z_2, ---, Z_r) = \frac{n!}{(n-r)!} \left(f(Z_1)\, f(Z_2) --- f(Z_r)\right) [1-F(Z_r)]^{n-r}$$

$$= \frac{n!}{(n-r)!} \theta^{-r} e^{-[Z_1 + Z_2 + --- Z_r + (n-r)Z_r]/\theta} \quad (12b)$$

The expression assumes simpler form if we define T as

$$T = Z_1 + Z_2 + --- + Z_r + (n-r)Z_r$$

$$= Z_1 + Z_2 + --- + Z_{r-1} + (n-r+1)Z_r$$

so that

$$g(Z_1, Z_2, ---, Z_r) = \frac{n!}{(n-r)!} \theta^r e^{-T/\theta} \quad (13b)$$

The variable T is seen to be the total life time of all items on the test. The right hand terms of 13b, though written in terms of T, is not in fact the density function of $\tilde{T}$, but is still the joint density of the order statistics $Z_1, Z_2, ---, Z_r$. However, it can be regarded as a likelihood function for θ in the current testing situation; so it can be written

$$L(\theta) = \frac{n!}{(n-r)!} \theta^{-r} e^{-r/\theta} \quad (14b)$$

Then the log likelihood function is

$$\ln L = \ln n! - \ln(n-r)! - r\ln\theta - T/\theta$$

i.e. $$\frac{\delta \ln L}{\delta\theta} = \frac{-r}{\theta} + \frac{T}{\theta^2}$$

and when set equal to zero, this yields a maximum likelihood estimator

$$\hat{\theta} = \frac{T}{r} \quad (15b)$$

$$= \frac{Z_1 + Z_2 + ---- + Z_r + (n-r)Z_r}{r}$$

It is of importance to find the density function for the random variable, and for this a change in variable is necessary, the old variable set $Z_1, Z_2, ---, Z_r$ being replaced by a new variable set defined as follows

$$
\begin{aligned}
\text{Let } w_1 &= nZ_1 \\
w_2 &= (n-1)(Z_2-Z_1) \\
w_3 &= (n-2)(Z_r-Z_2) \\
&\vdots \\
w_{r-1} &= (n-r-2)(Z_{r-1} - Z_{r-2}) \\
w_r &= (n-r+1)(Z_r - Z_{r-1})
\end{aligned}
\tag{16b}
$$

Since $Z_1 \leq Z_2 \leq --- \leq Z_r$, the w's are all non-negative. They can be seen to represent 'partial lives'. This is because all n components were put on test and all lived until time Z_1 and so nZ_1 is the accumulated group life until Z_1 . After the first failure, only n-1 components were alive and they lived $(Z_2 - Z_1)$ hours until the second failure. Hence the accumulated group life over the second interval is $(n-1)(Z_2 - Z_1)$. Similar reasoning obtains for the rest of the set. If all the equations of 16b are added, the right hand sides sum to

$$Z_1 + Z_2 + - - - + Z_{r-1} + (n-r+1)Z_r = T \tag{17b}$$

and thus

$$T = w_1 + w_2 + - - - + w_r \tag{18b}$$

as could be expected. The inverse transformation must be obtained from 16b and this is done in a sequential manner, yielding

$$
\begin{aligned}
Z_1 &= w_1/n \\
Z_2 &= w_1/n + w_2/n-1 \\
&\vdots \\
Z_r &= w_1/n + w_2/(n-1) + - - - + w_r/(n-r+1)
\end{aligned}
$$

the Jacobian of this transformation is

$$\begin{vmatrix} 1/n & & & & \\ 1/n & 1/(n-1) & & & \\ 1/n & 1/(n-1) & 1/(n-2) & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 1/n & 1/(n-1) & 1/(n-2) & \cdots & 1/(n-r+1) \end{vmatrix} = [n(n-1)(n-2)\ldots(n-r+1)]^{-1} = \frac{(n-r)!}{n!}$$

Now since the new density g^* is developed from the relationship

$$g^*(w_1, w_2, ---, w_r) = g(Z_1, Z_2, ---, Z_r)\ |J|$$

it could be written

$$g^*(w_1, w_2, ---, w_r) = \frac{n!}{(n-r)!}\ \theta^{-r}\ \frac{e^{-(w_1 + w_2 --- + w_r)}}{\theta}\qquad \frac{(n-r)!}{n!}$$

$$= \frac{e^{-w_1/\theta}}{\theta} \cdot \frac{e^{-w_2/\theta}}{\theta} \cdots \frac{e^{-w_r/\theta}}{\theta} \qquad (21b)$$

Each w_i is a random variable whose lower bound is zero(when Z_i occurs right after Z_{i-1}) and with no upper bound since can occur 'a long, long time' after Z_{i-1}. It can also be proved that the w_i are statistically independent. Then since $\tilde{w}_i \sim G(1,\theta)$ the total life $T = w_1+w_2+\ldots+w_r$ has the distribution

$$\tilde{T} \sim G(r,\theta)$$

SECTION 2.

SAMPLES OF NEGATIVE EXPONENTIAL PLOTS OF TIME VS LOG(N+1)/(N+1-I).

Graphical procedure 1

Several statistical tests could be conducted for distinguishing effectively between hypotheses regarding the values of the intensity λ of tested systems.

Epstein (1,2,3a,3b,4,5) considers a graphical procedure. With a limited amount of initial data from reliability tests, each specific statistical test can, with a high degree of confidence, distinguish an exponential distribution from certain other distributions. There would always be the possibility of encountering distributions of the time of failure-free operation that differ significantly from an exponential distribution. This section demonstrates how a simulated set of life test data could reveal the nature of the underlying exponential distribution. The following is the data of failure times provided by Epstein in (3b) :

1.2	13.7	38.9	72.4	102.8	151.6	203.0
2.2	15.1	47.9	73.6	108.5	152.6	204.3
4.9	15.2	48.4	76.8	128.7	164.2	229.5
5.0	23.9	49.3	83.8	133.6	166.8	253.1
6.8	24.3	53.2	95.1	144.1	178.6	304.1
7.0	25.1	55.6	97.9	147.6	185.2	341.7
12.1	35.8	62.7	99.6	150.6	187.1	354.4

wherein the hypothesis tested is whether the underlying distribution of life is exponential.

The cumulative distribution function (c.d.f) F(t) of the exponential distribution is given by :

$$F(t) = 0 \qquad t < 0$$
$$= 1-e^{-(t/\theta)} \qquad t \geq 0$$

hence

$$t/\theta = -\log (1-F(t)),$$

the expected value of F(t) being $(\frac{i}{n+1})$. The above formula yields

$$t/\theta = \log (\frac{n+1}{n+1-i}) = y$$

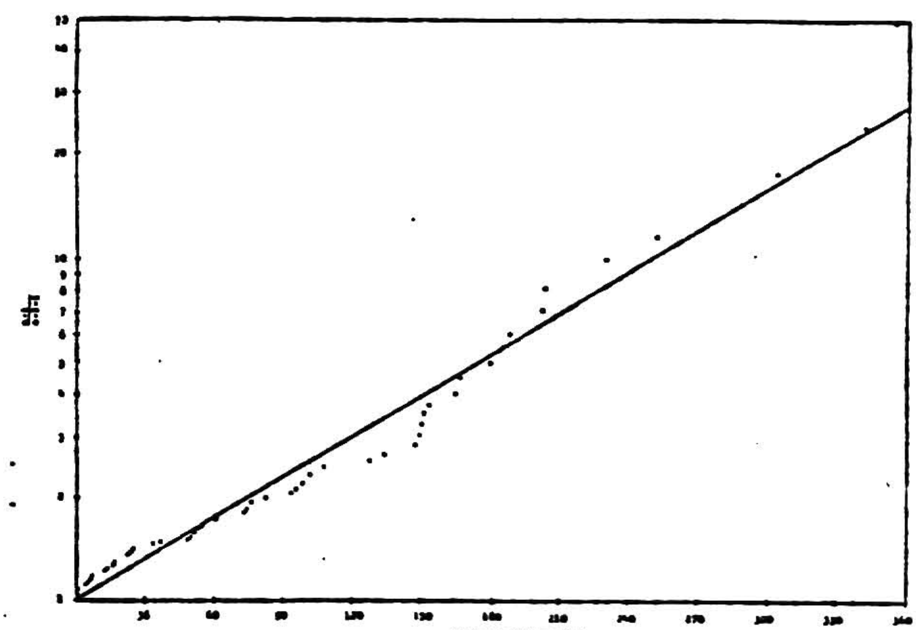
and when the values of y are plotted against t, a straight line with a slope $1/\theta$ is obtained. If the exponential assumption holds, then the plotted points can be fitted well by a straight line passing through the origin.

This section contains several simulated sets of data obtained by the process of generating random exponential deviates. Appendix 1 contains a SAS program that generates, sorts and plots the required values; it also explains the program. Tables 2.1.1 to 2.1.20 contain the data generated while Figures 2.1.1 to 2.1.20 display the plots. To test if sample size had any noticeable effect on linear trends, samples of size 5, 10, 20 and

50 were generated. At this point it is worth comparing Epstein's data with the failure times (x values) generated by the program. Figure 2.1a is the plot obtained using Epstein's data and this could be compared with the plots displayed in Figures 2.1.1 to 2.1.20.

DISCUSSION

Table 2.1.1 provides a sample of 5 randomly generated exponential deviates (X values), while Figure 2.1.1 plots these values against y, the logarithm of the reciprocal reliability function. It seems rather apparent that the trend is linear. Tables 2.1.2 to 2.1.5 display several such samples of size 5 while Figures 2.1.2 to 2.1.5 plot values of x with the corresponding values of y. Figures 2.1.2 and 2.1.3 show that the first few values are low and such low values are responsible for obscuring the tendency towards linearity. Higher initial values of x indicate the possibility of a better linear fit than smaller initial values do. Tables 2.1.5 to 2.1.20 provide several such samples of varying sizes, while Figures 2.1.5 to 2.1.20 plot the values. A comparison of the data with those used by Epstein indicate a very close similarity. The scatter of values obviously indicate the linear trend highlighted in the plots. It does seem evident that one is able to recognize the nature of a particular sample by a close examination of the respective data. The conclusion that one draws is that the experienced analyst could



—Graph of $y_i = \log (n + 1)/(n + 1 - i)$ against τ_i.
Data from Example 1 ($n = 49$).

FIGURE 2.1a Plot of Epstein's Data

easily detect the exponential nature of a certain set of life test data without having to resort to formal tests used by Epstein.

Table 2.1.1 Data Set 1 (sample size = 5)

SAS

OBS	X	Y
1	0.26453	0.18232
2	0.29462	0.40547
3	0.36803	0.69315
4	1.23092	1.09861
5	1.94155	1.79176

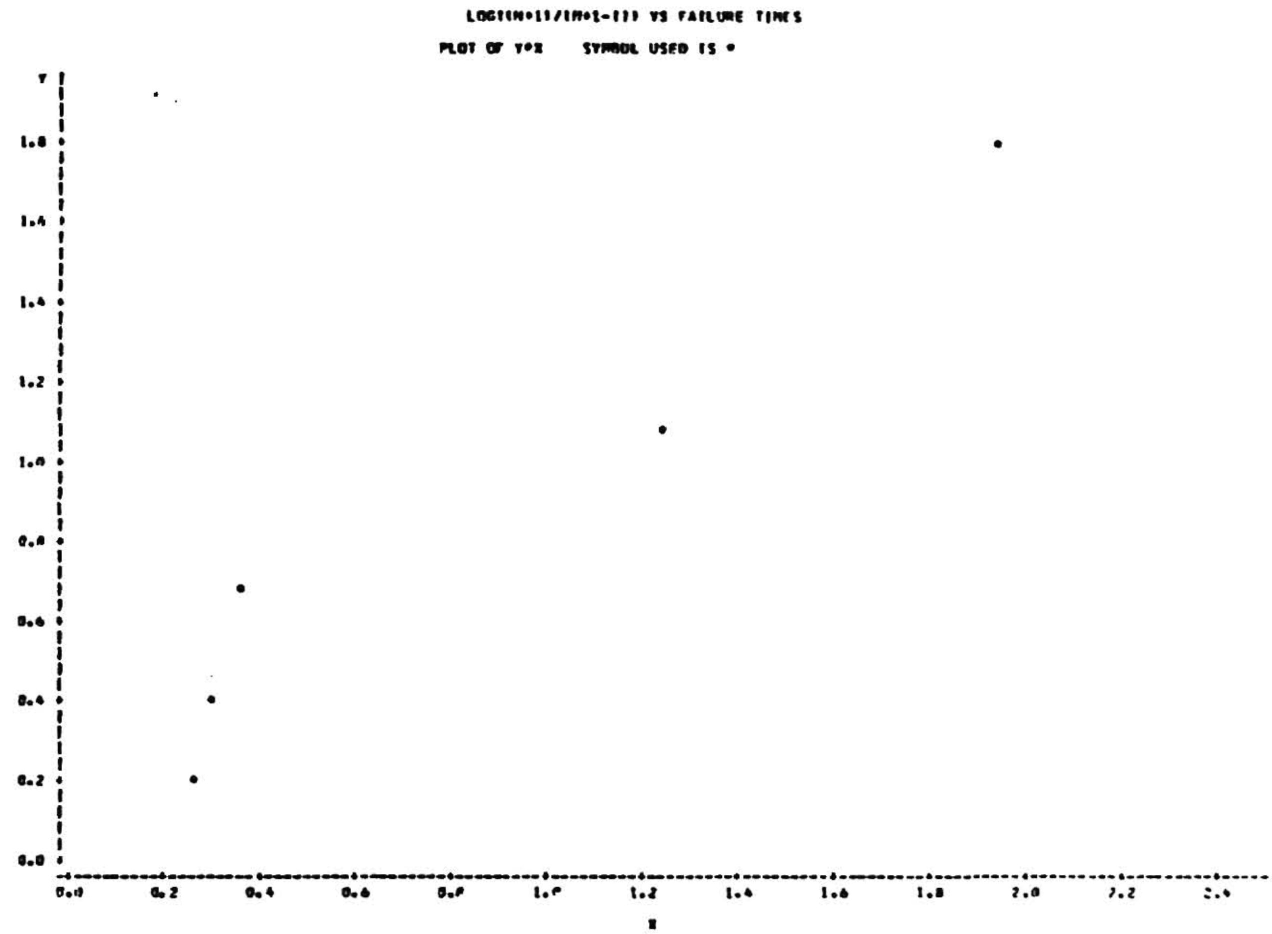


Figure 2.1.1 Plot of dataset 1 (sample size = 5)

Table 2.1.2 Data Set 2 (sample size = 5)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X1	Y
1	0.25873	0.18232
2	0.63471	0.40547
3	2.37601	0.69315
4	2.38712	1.09861
5	3.80697	1.79176

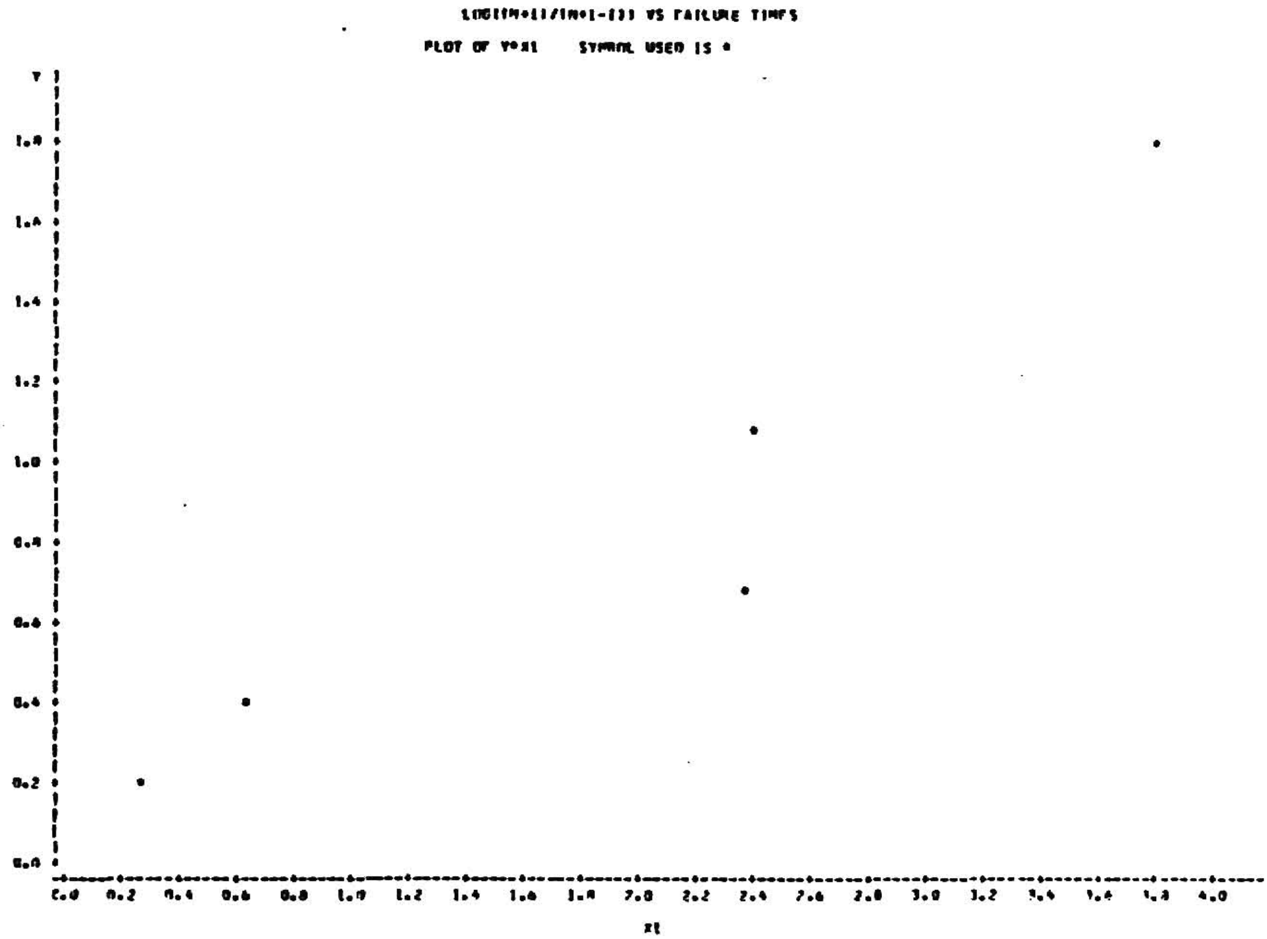


Figure 2.1.2 Plot of dataset 2 (sample size = 5)

Table 2.1.3 Data Set 3 (sample size = 5)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X2	Y
1	0.66796	0.18232
2	0.78674	0.40547
3	0.98154	0.69315
4	1.38133	1.09861
5	2.14873	1.79176

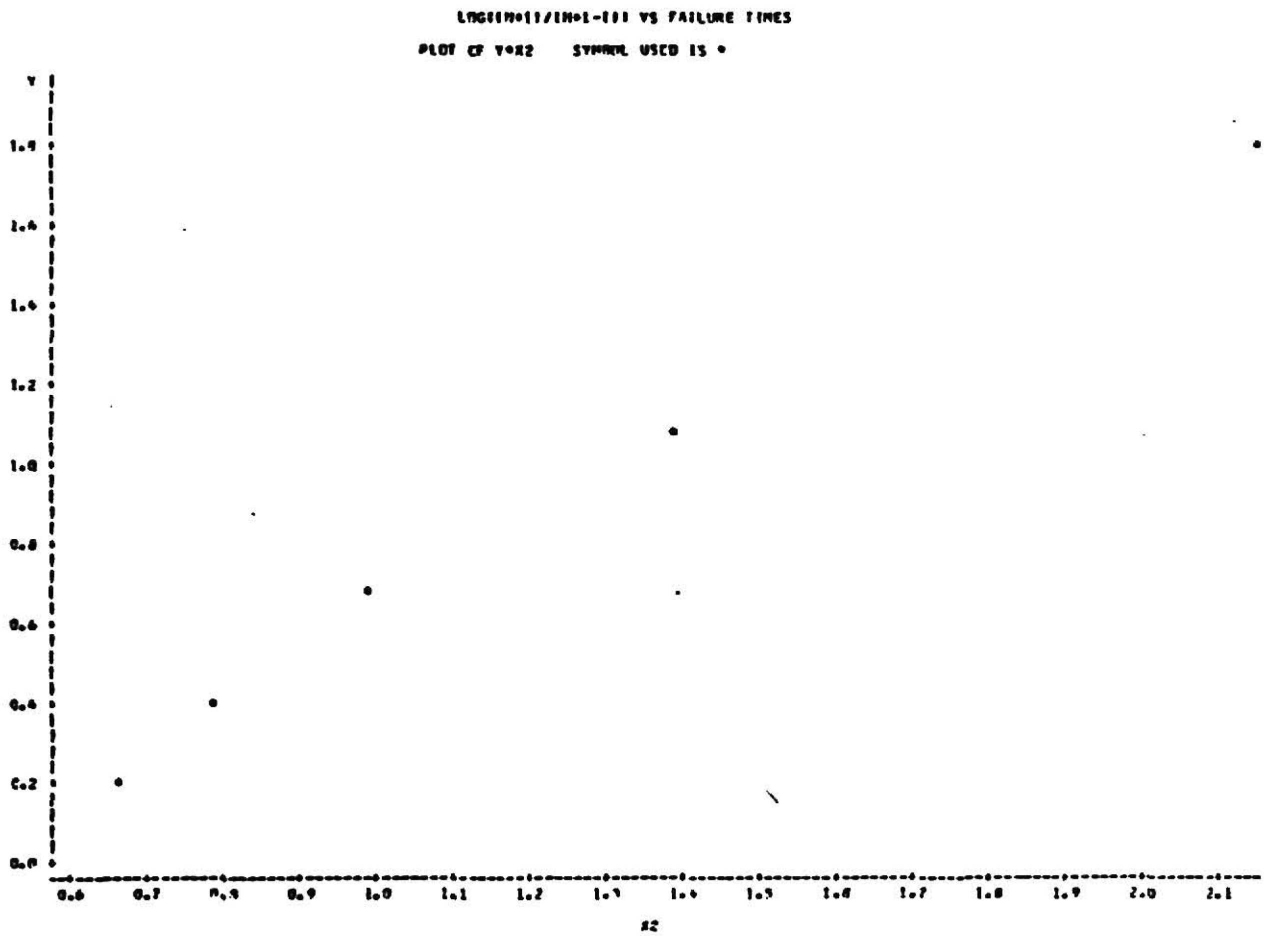


Figure 2.1.3 Plot of dataset 3 (sample size = 5)

Table 2.1.4 Data Set 4 (sample size = 5)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X3	Y
1	0.42353	0.18232
2	0.43274	0.40547
3	0.51307	0.69315
4	1.79673	1.09861
5	2.05447	1.79176

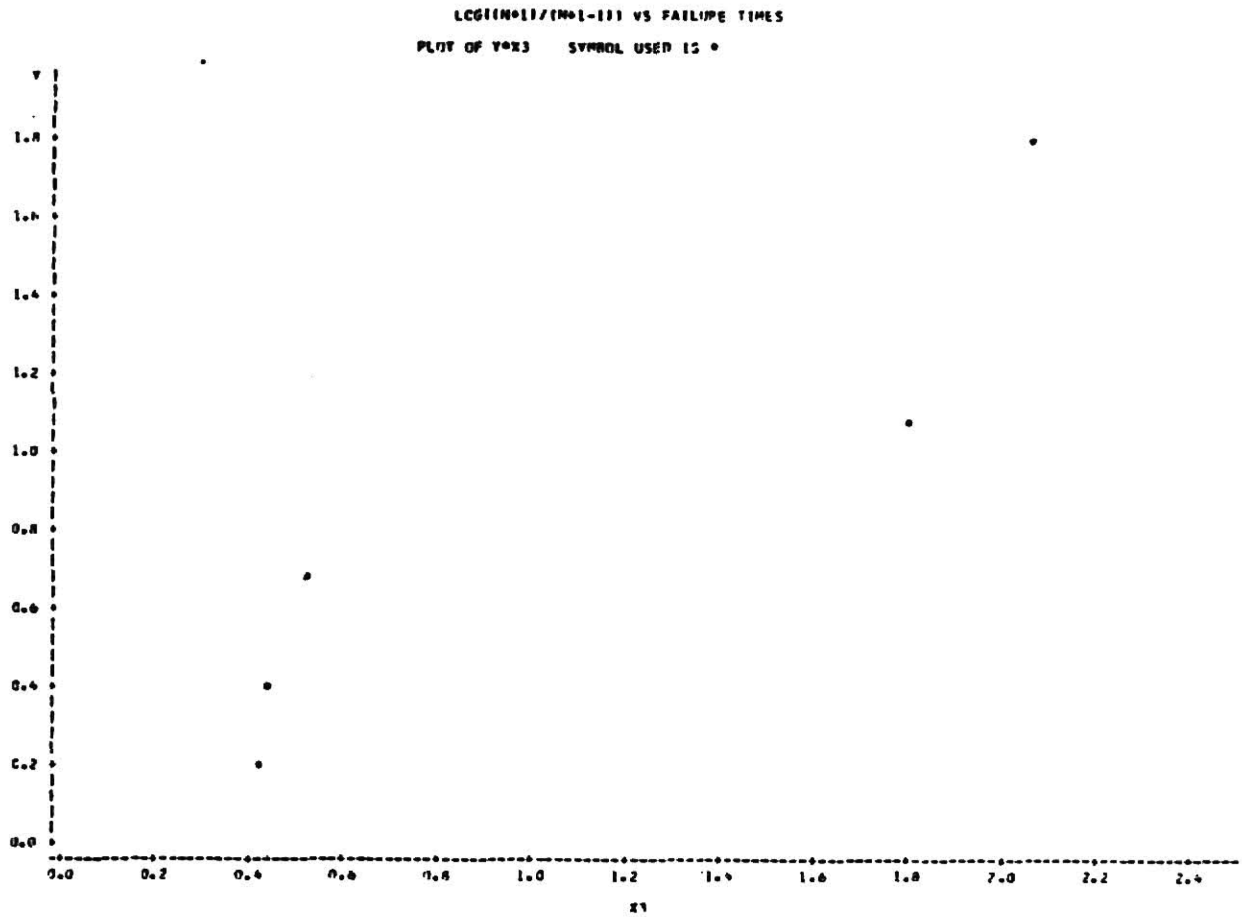


Figure 2.1.4 Plot of dataset 4 (sample size = 5)

Table 2.1.5 Data Set 5 (sample size = 5)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X4	Y
1	0.29167	0.18232
2	0.59451	0.40547
3	1.29709	0.69315
4	1.55636	1.09861
5	2.27235	1.79176

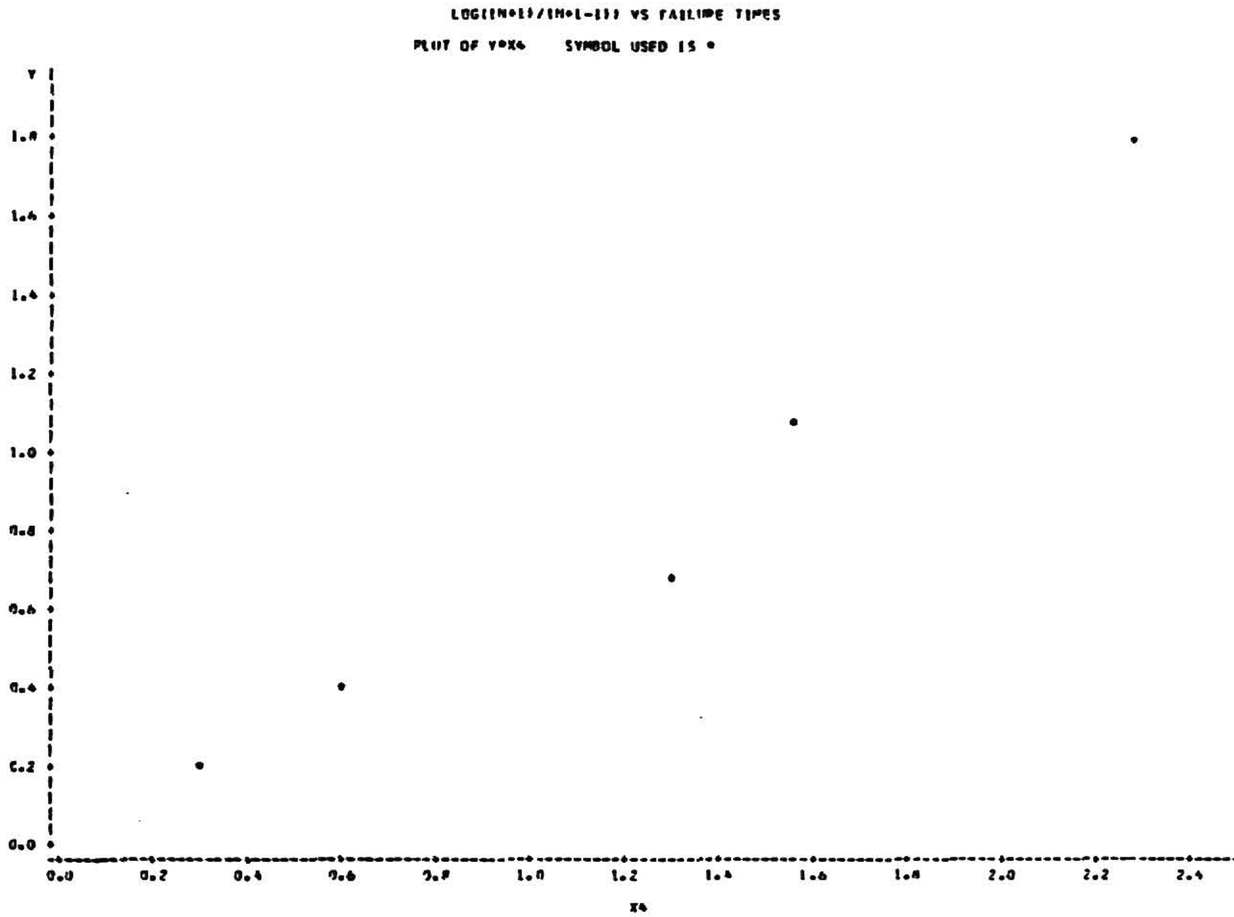


Figure 2.1.5 Plot of dataset 5 (sample size = 5)

Table 2.1.6 Data Set 1 (sample size = 10)

SAS

OBS	X	Y
1	0.05573	0.09531
2	0.12237	0.20067
3	0.13881	0.31845
4	0.21527	0.45199
5	0.32043	0.60614
6	0.50563	0.78846
7	0.66081	1.01160
8	1.17202	1.29929
9	1.57637	1.70475
10	2.52884	2.39790

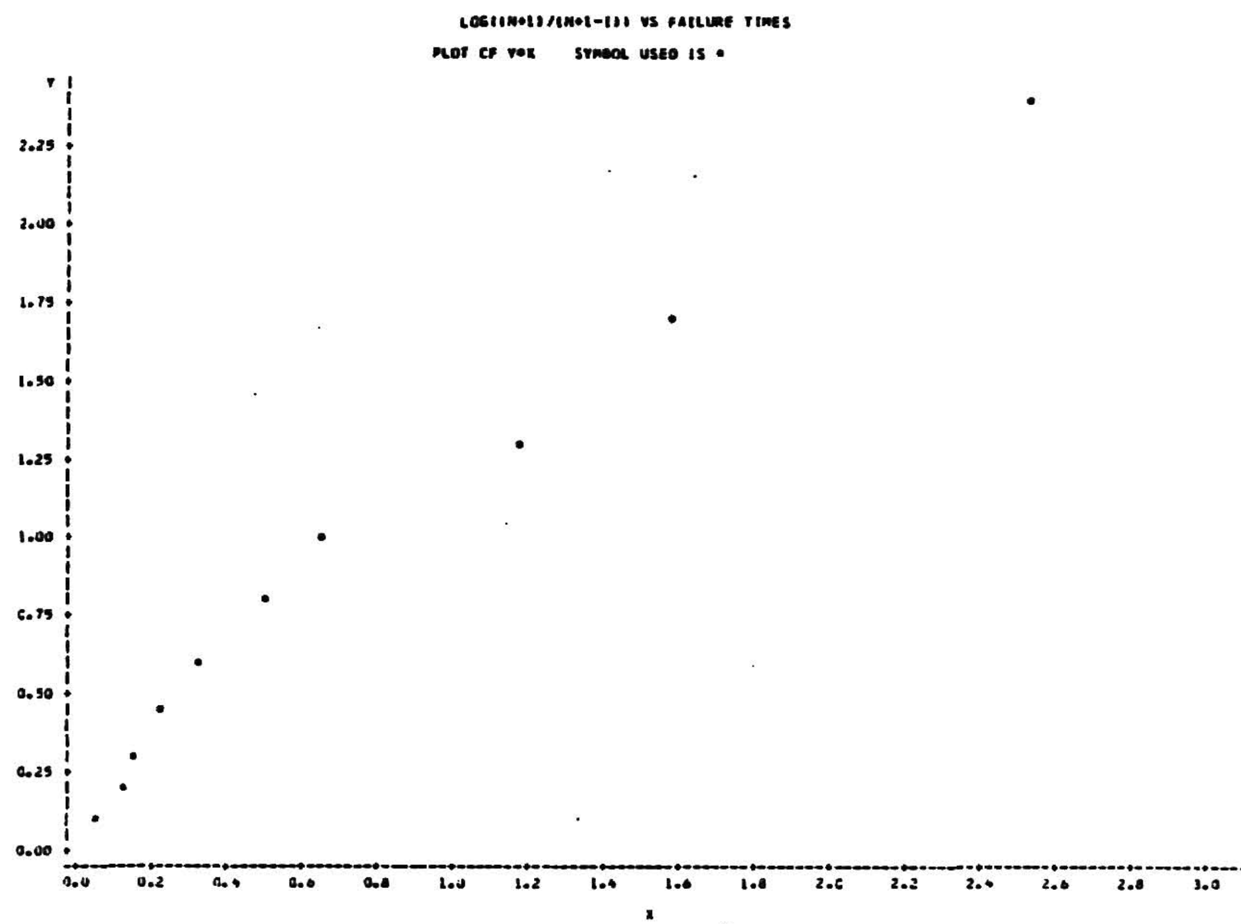


Figure 2.1.6 Plot of dataset 1 (sample size = 10)

Table 2.1.7 Data Set 2 (sample size = 10)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X1	Y
1	0.05495	0.09531
2	0.23518	0.20067
3	0.29197	0.31845
4	0.39147	0.45199
5	0.40227	0.60614
6	0.45576	0.78846
7	0.61274	1.01160
8	0.95455	1.29928
9	1.89141	1.70475
10	2.11337	2.39790

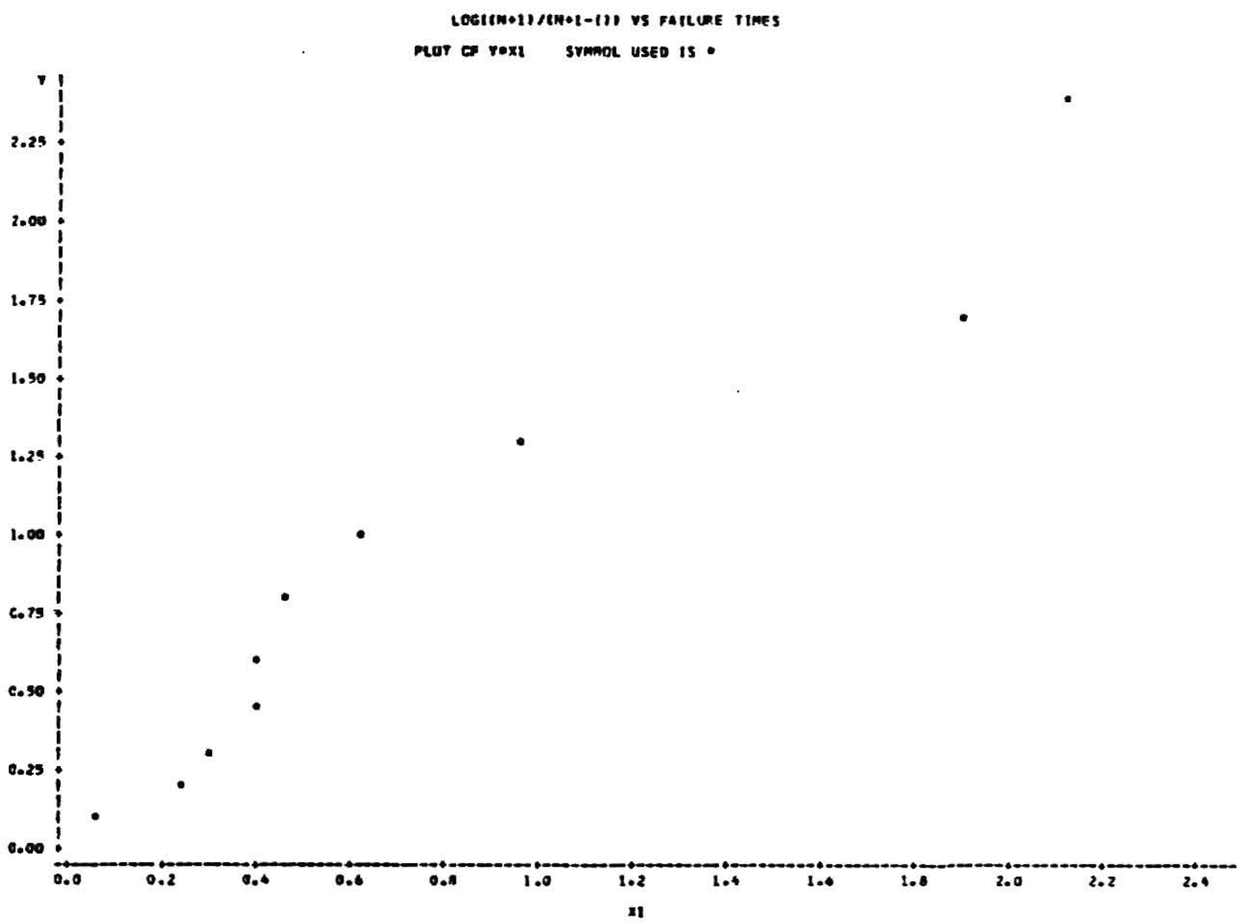


Figure 2.1.7 Plot of dataset 2 (sample size = 10)

Table 2.1.8 Data Set 3 (sample size = 10)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X2	Y
1	0.23721	0.09531
2	0.24346	0.20067
3	0.33995	0.31845
4	0.57777	0.45199
5	0.70250	0.60614
6	0.79783	0.78846
7	1.03134	1.01160
8	1.06963	1.29928
9	1.18968	1.70475
10	2.77565	2.39790

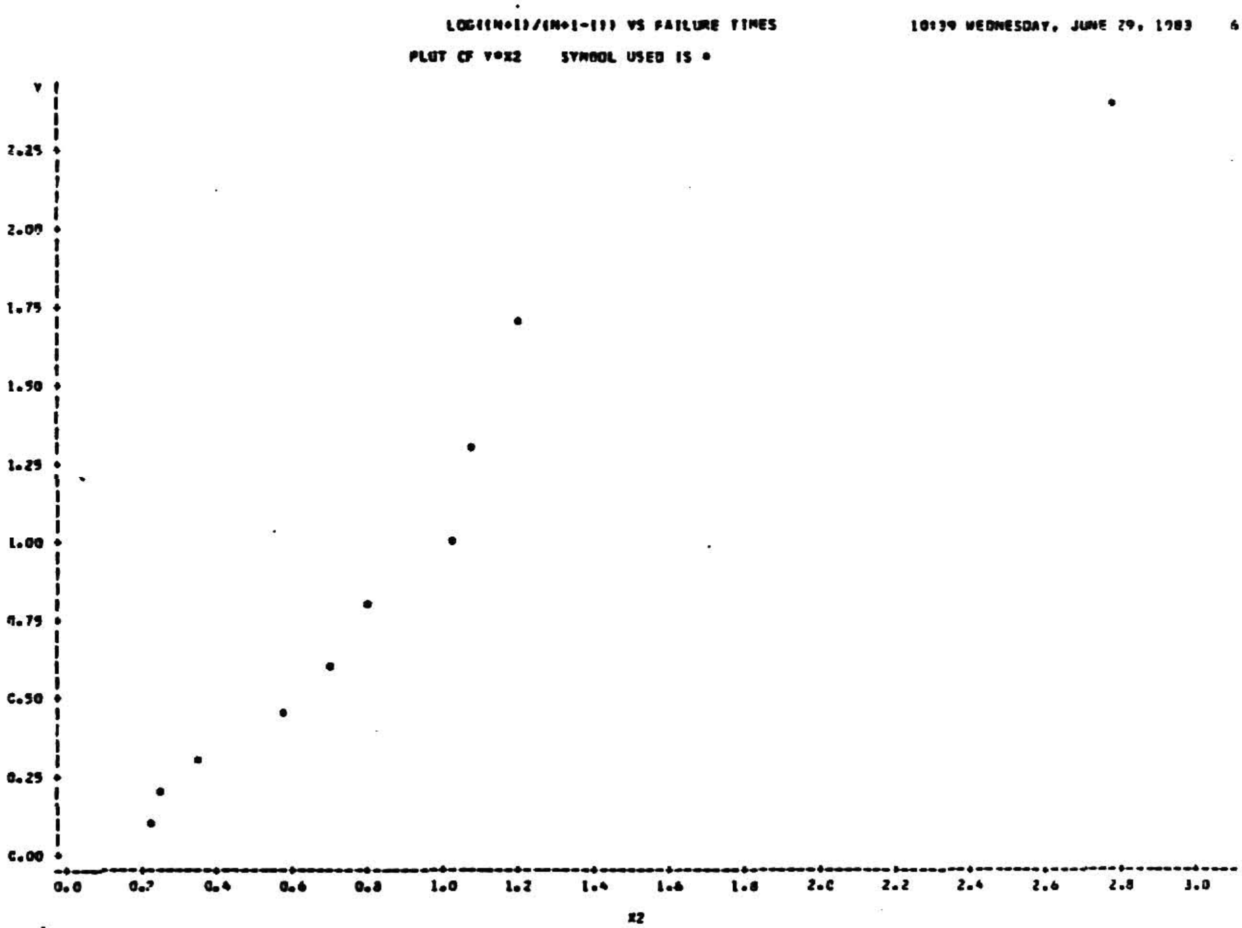


Figure 2.1.8 Plot of dataset 3 (sample size = 10)

Table 2.1.9 Data Set 4 (sample size = 10)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X3	Y
1	0.15921	0.09531
2	0.30354	0.20067
3	0.38203	0.31845
4	0.68658	0.45199
5	0.71078	0.60614
6	1.19662	0.78846
7	1.29859	1.01160
8	1.46325	1.29928
9	1.51292	1.70475
10	2.73611	2.39790

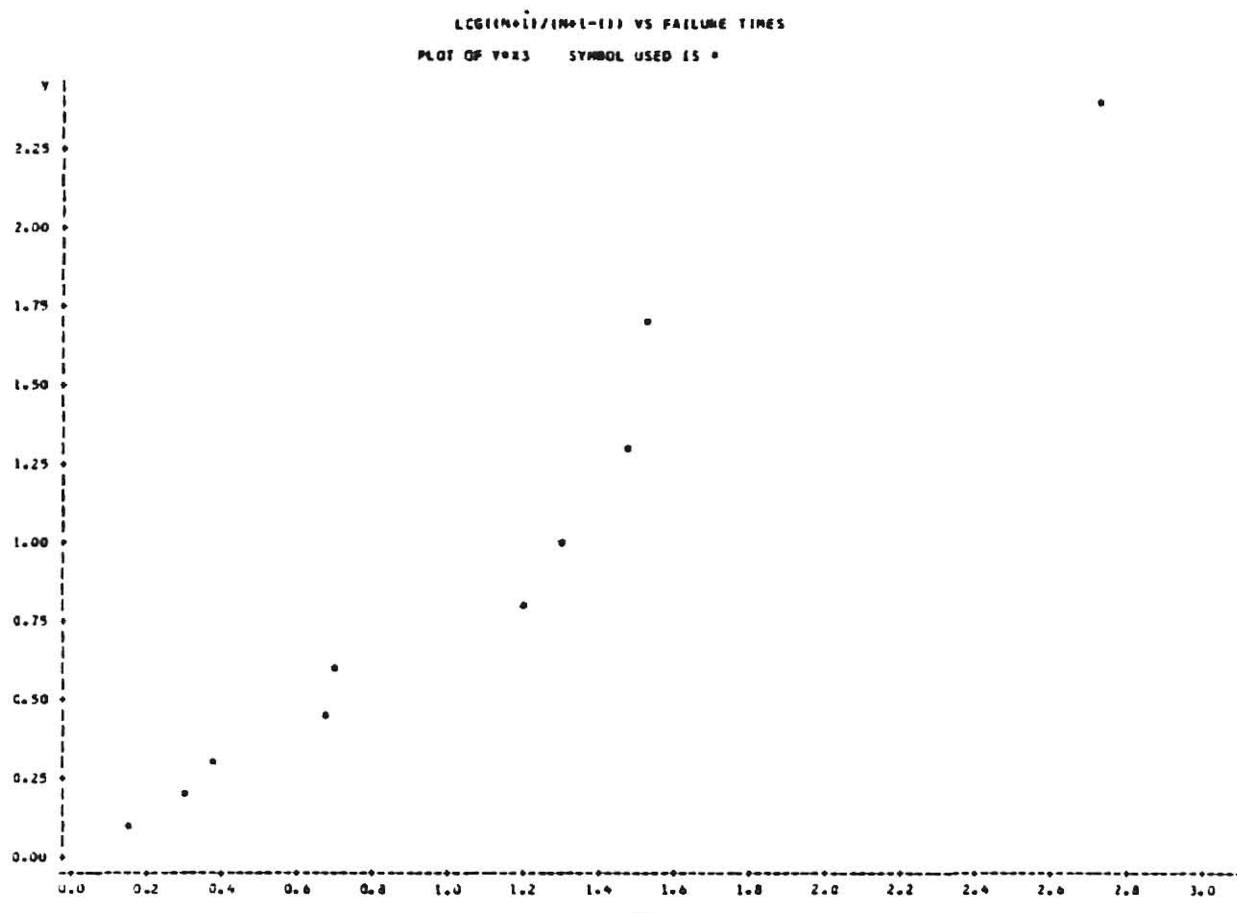


Figure 2.1.9 Plot of dataset 4 (sample size = 10)

Table 2.1.10 Data Set 5 (sample size = 10)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X4	Y
1	0.01158	0.09531
2	0.06377	0.20067
3	0.07894	0.31845
4	0.14470	0.45199
5	0.63795	0.60614
6	0.76206	0.78846
7	1.11758	1.01160
8	1.12161	1.29928
9	1.20767	1.70475
10	4.91842	2.39790

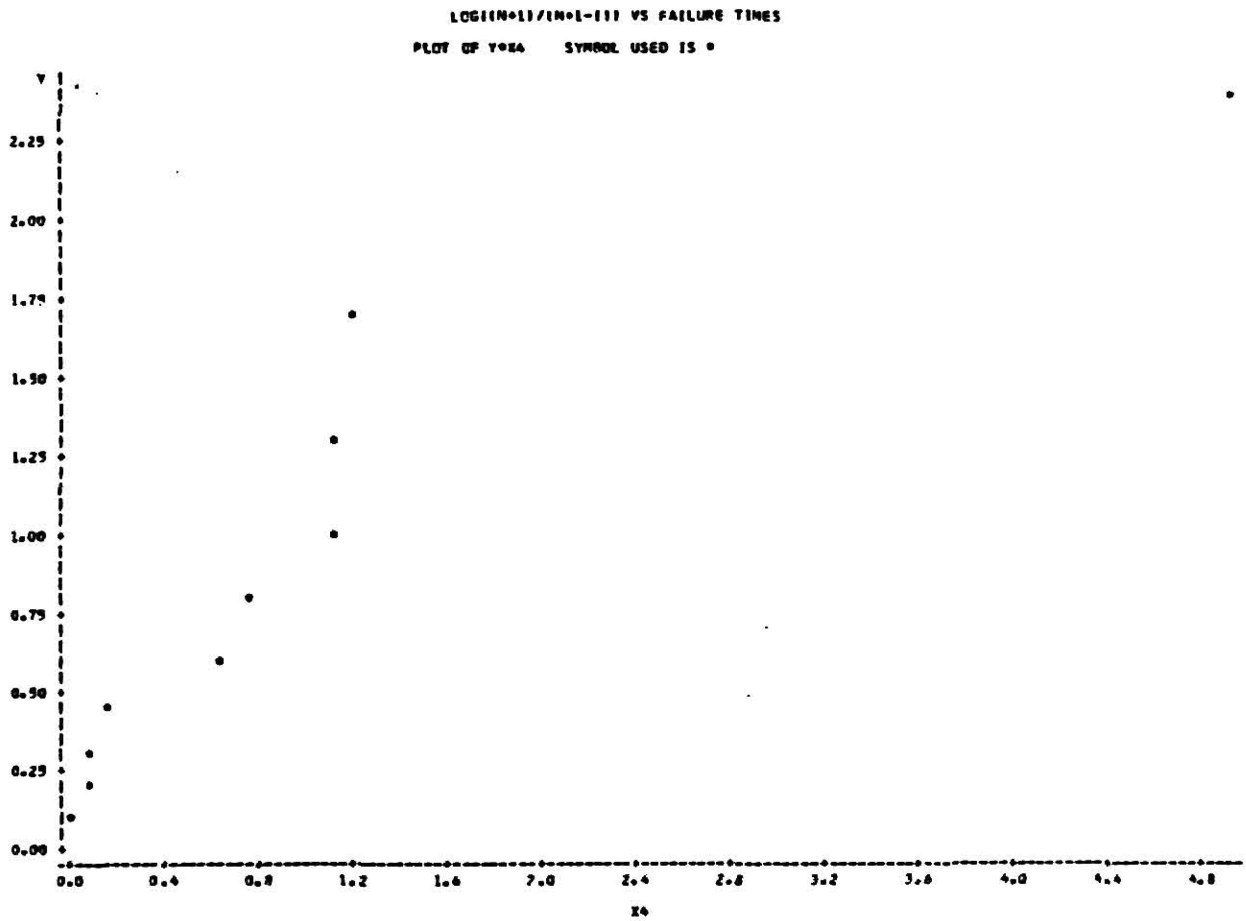


Figure 2.1.10 Plot of dataset 5 (sample size = 10)

Table 2.1.11 Data Set 1 (sample size = 20)

SAS

OBS	X	Y
1	0.07450	0.04879
2	0.15479	0.10008
3	0.17448	0.15415
4	0.18569	0.21131
5	0.33687	0.27193
6	0.60145	0.33647
7	0.68936	0.40547
8	0.85824	0.47957
9	1.15500	0.55962
10	1.15895	0.64663
11	1.23669	0.74194
12	1.42042	0.84730
13	1.46797	0.96508
14	1.56148	1.09961
15	1.85983	1.25276
16	1.87211	1.43508
17	2.12075	1.65823
18	2.29668	1.94591
19	2.57600	2.35138
20	3.27994	3.04452

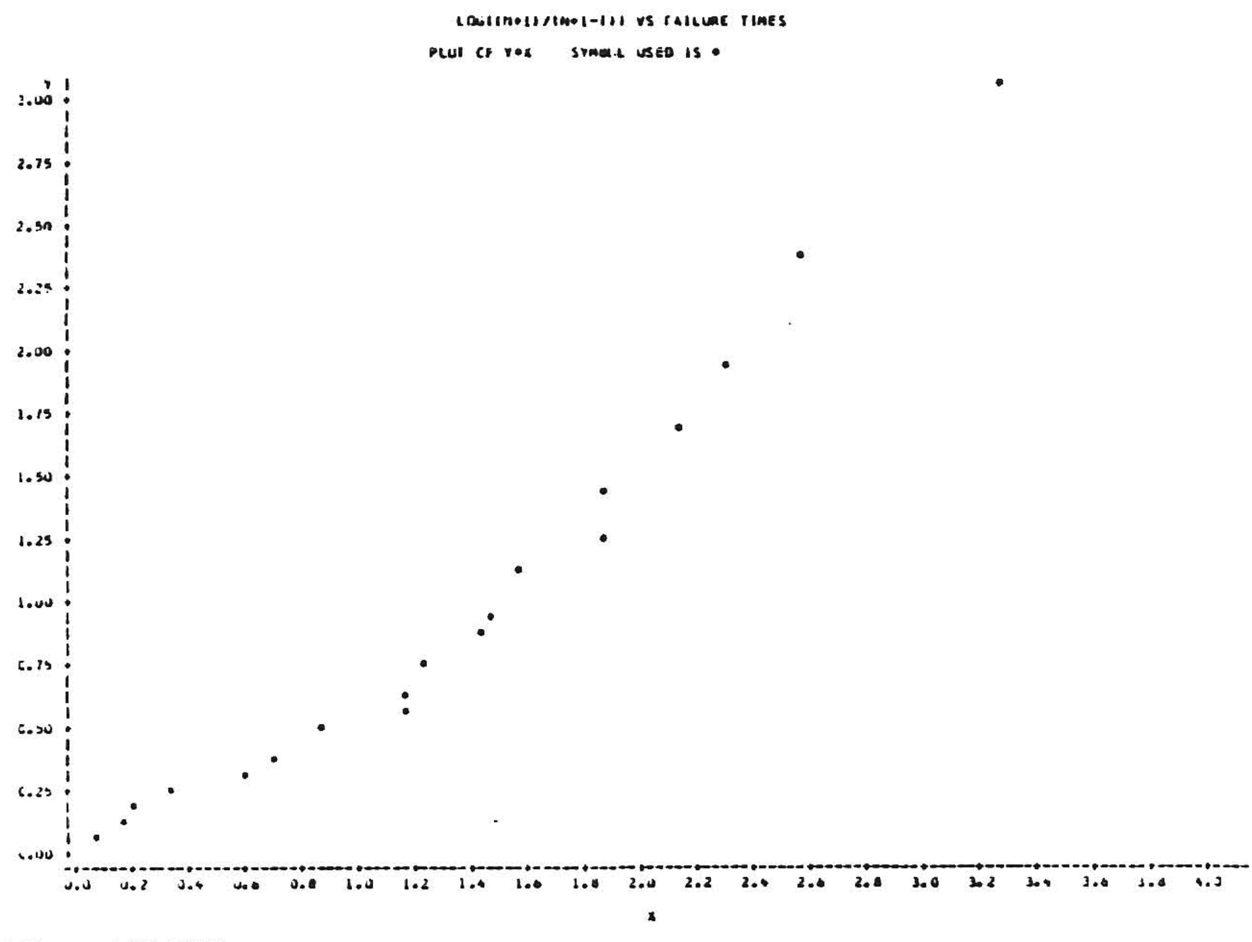


Figure 2.1.11 Plot of dataset 1 (sample size = 20)

Table 2.1.12 Data Set 2 (sample size = 20)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X1	Y
1	0.09759	0.04879
2	0.14978	0.10008
3	0.24262	0.15415
4	0.25482	0.21131
5	0.30623	0.27193
6	0.42030	0.33647
7	0.43796	0.40547
8	0.44440	0.47957
9	0.61280	0.55962
10	0.70223	0.64663
11	0.70274	0.74194
12	0.78791	0.84730
13	0.87805	0.96508
14	0.87811	1.09861
15	1.17480	1.25276
16	1.53174	1.43508
17	1.73053	1.65823
18	1.75933	1.94591
19	1.82737	2.35138
20	2.47328	3.04452

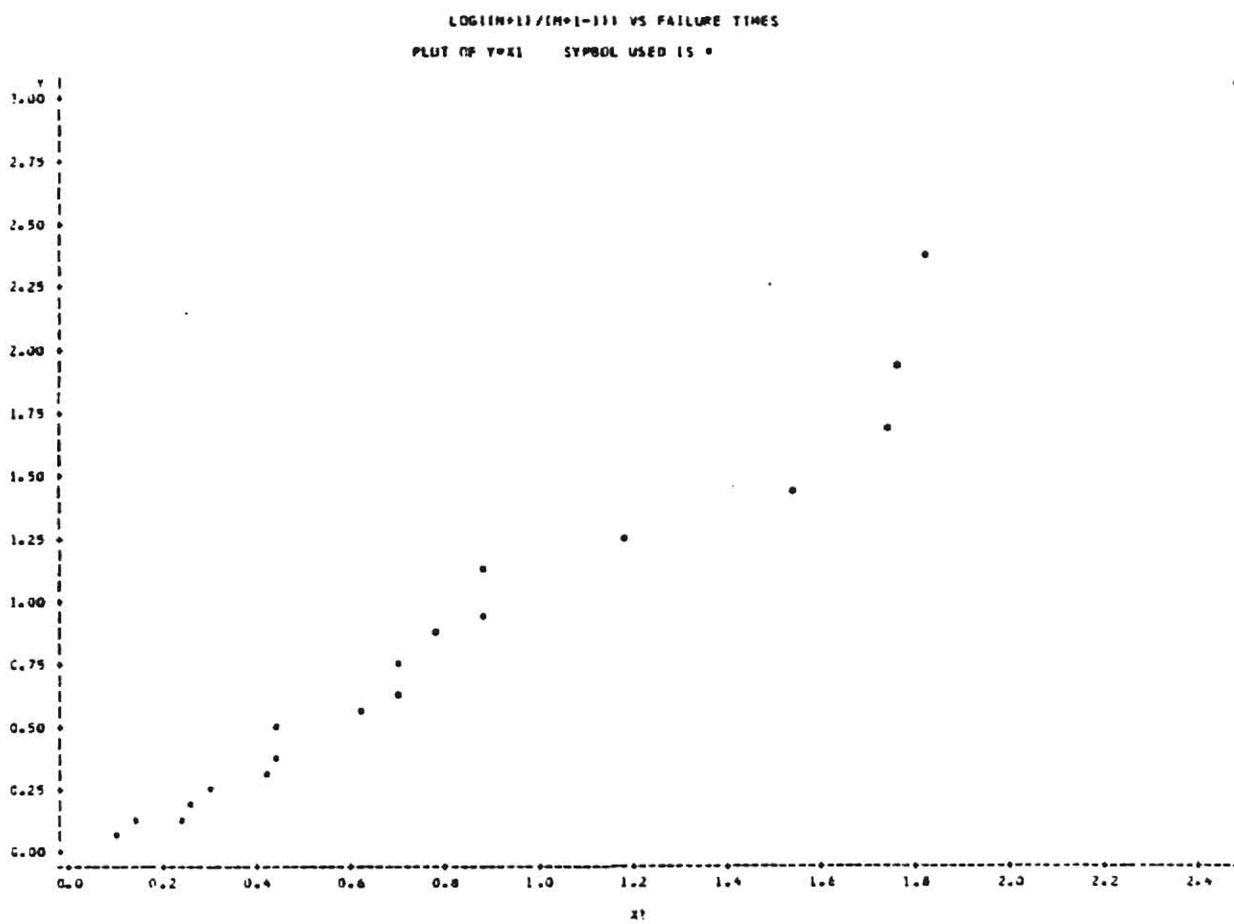


Figure 2.1.12 Plot of dataset 2 (sample size = 20)

Table 2.1.13 Data Set 3 (sample size = 20)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X2	Y
1	0.00373	0.04879
2	0.04471	0.10008
3	0.12168	0.15415
4	0.22960	0.21131
5	0.40349	0.27193
6	0.48447	0.33647
7	0.51098	0.40547
8	0.70911	0.47957
9	0.71056	0.55962
10	0.77185	0.64663
11	0.87342	0.74194
12	0.89705	0.84730
13	0.90190	0.96508
14	0.99498	1.09861
15	1.01964	1.25276
16	1.25267	1.43508
17	1.32734	1.65823
18	1.45699	1.94591
19	1.96850	2.35138
20	3.14073	3.04452

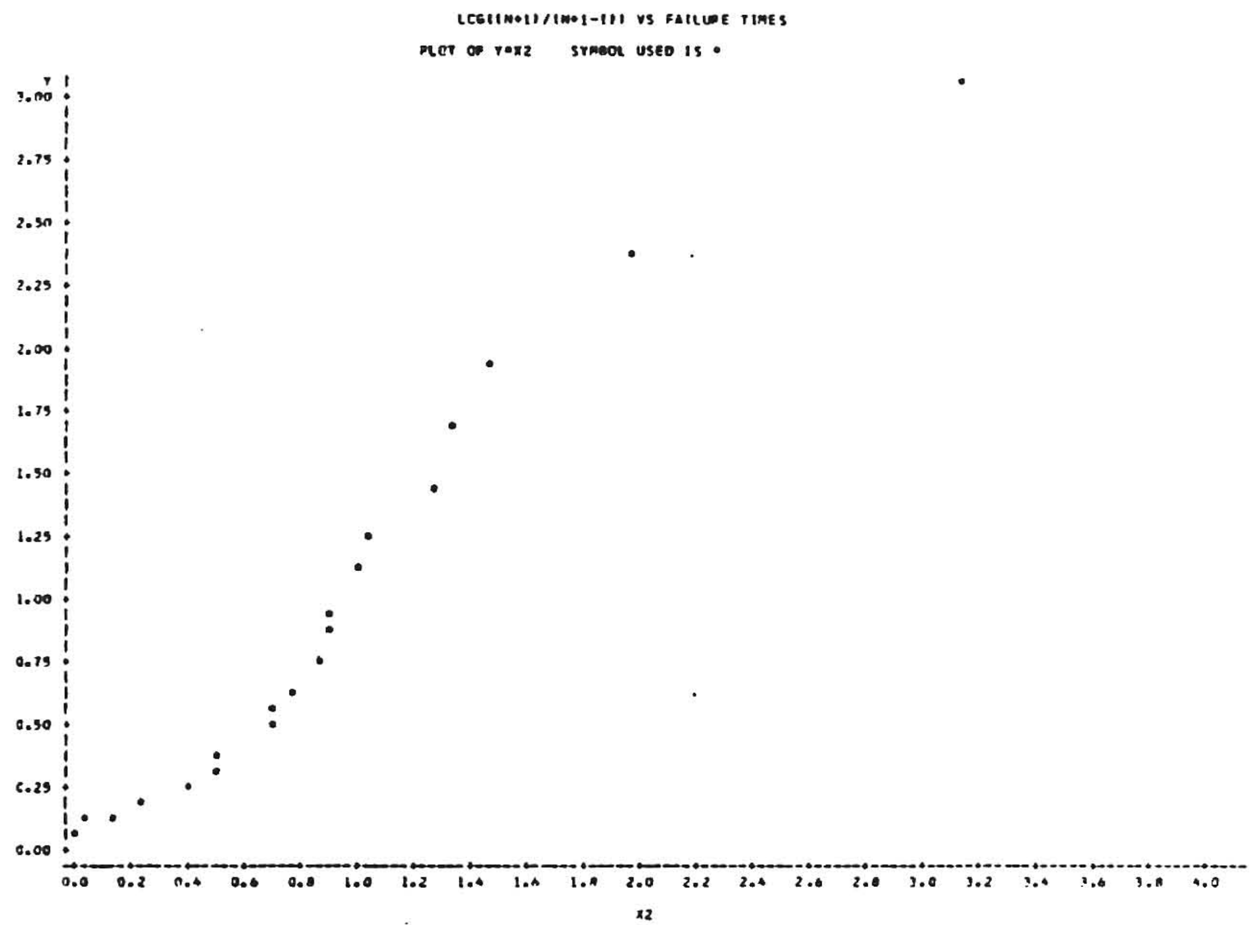


Figure 2.1.13 Plot of dataset 3 (sample size = 20)

Table 2.1.14 Data Set 4 (sample size = 20)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X3	Y
1	0.02963	0.04879
2	0.07923	0.10008
3	0.09385	0.15415
4	0.18247	0.21131
5	0.31156	0.27193
6	0.47182	0.33647
7	0.63395	0.40547
8	0.71902	0.47957
9	0.85693	0.55962
10	0.86134	0.64663
11	0.88718	0.74194
12	0.89024	0.84730
13	0.91802	0.96508
14	0.92154	1.09861
15	0.97415	1.25276
16	1.22978	1.43508
17	1.34009	1.65823
18	1.75460	1.94591
19	2.55893	2.35138
20	2.60805	3.04452

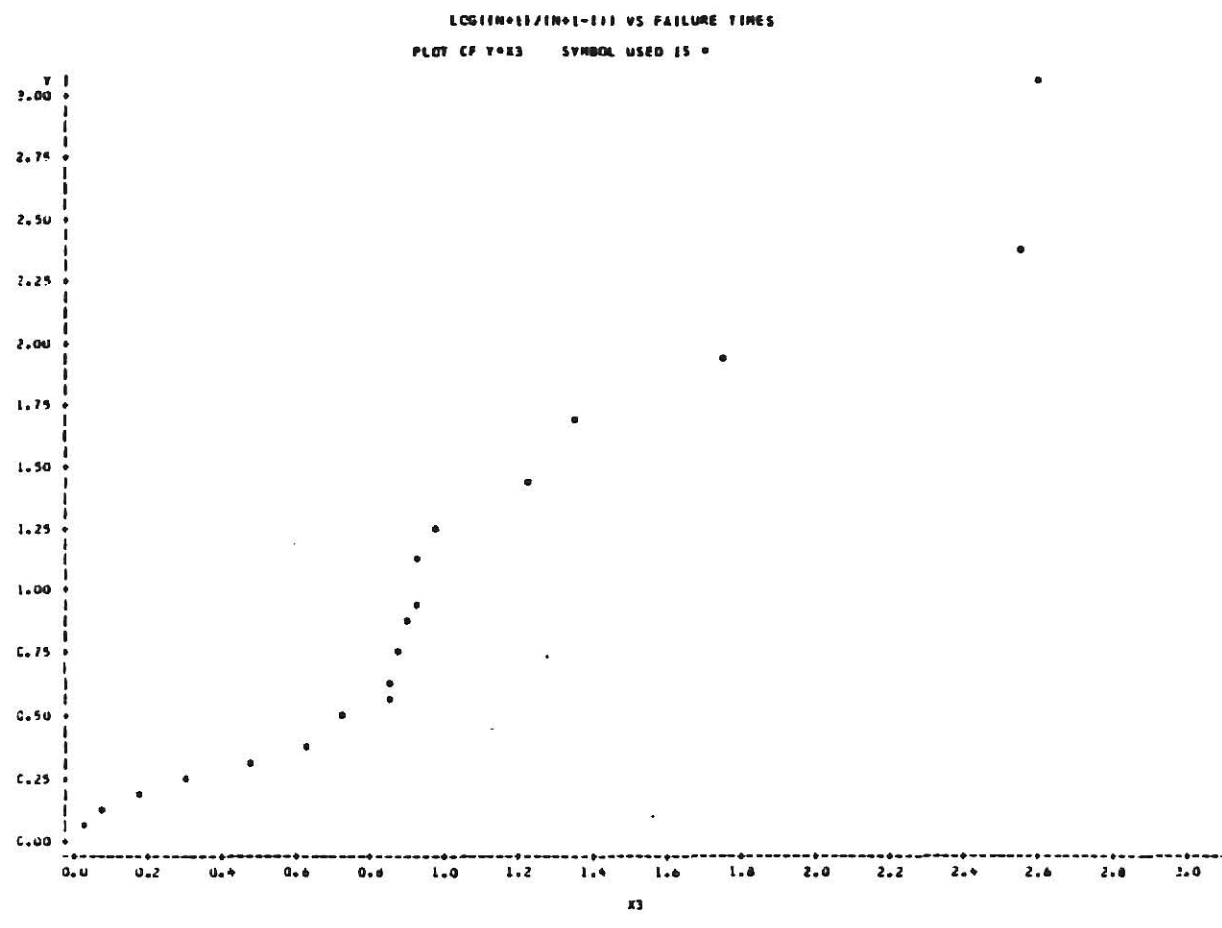


Figure 2.1.14 Plot of dataset 4 (sample size = 20)

Table 2.1.15 Data Set 5 (sample size = 20)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X4	Y
1	0.04097	0.04879
2	0.05076	0.10008
3	0.17806	0.15415
4	0.18115	0.21131
5	0.18625	0.27193
6	0.24255	0.33647
7	0.40793	0.40547
8	0.42037	0.47957
9	0.47640	0.55962
10	0.47909	0.64663
11	0.56075	0.74194
12	0.82576	0.84730
13	0.87677	0.96508
14	1.18562	1.09861
15	1.22336	1.25276
16	1.37578	1.43508
17	1.49870	1.65823
18	1.49936	1.94591
19	1.54362	2.35138
20	1.63969	3.04452

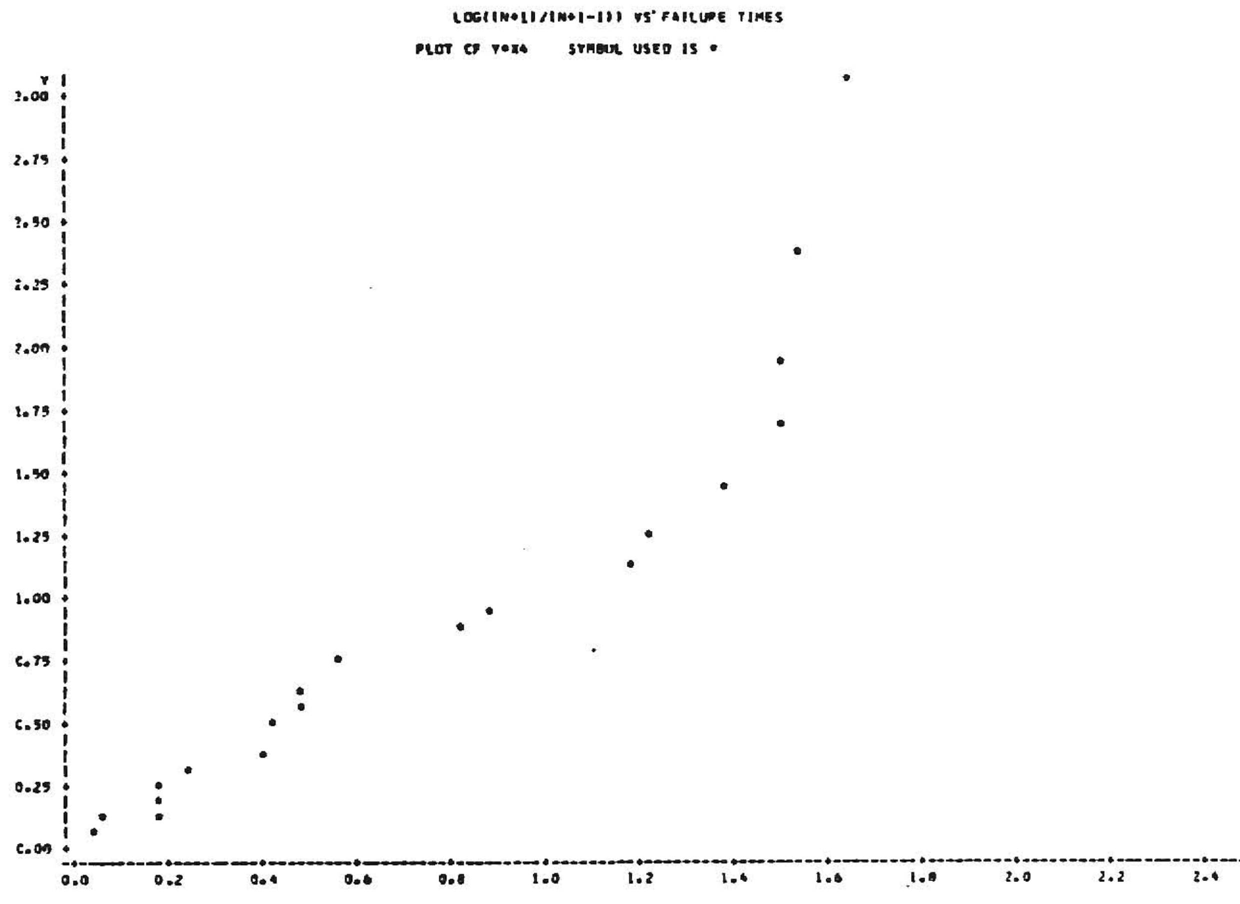


Figure 2.1.15 Plot of dataset 5 (sample size = 20)

Table 2.1.16 Data Set 1 (sample size = 50)

SAS

OBS	X	Y
1	0.02113	0.01980
2	0.02409	0.04001
3	0.02688	0.06062
4	0.03368	0.08168
5	0.06400	0.10318
6	0.07837	0.12516
7	0.11027	0.14764
8	0.11618	0.17063
9	0.12580	0.19416
10	0.14974	0.21825
11	0.19213	0.24295
12	0.23244	0.26826
13	0.24976	0.29424
14	0.27486	0.32091
15	0.41976	0.34831
16	0.44003	0.37648
17	0.44324	0.40547
18	0.52165	0.43532
19	0.55171	0.46609
20	0.57586	0.49784
21	0.71129	0.53063
22	0.71155	0.56453
23	0.83012	0.59962
24	0.99339	0.63599
25	0.99474	0.67373
26	0.99596	0.71295
27	1.03180	0.75377
28	1.12924	0.79633
29	1.20422	0.84078
30	1.22876	0.88730
31	1.24740	0.93609
32	1.31144	0.98739
33	1.36626	1.04145
34	1.47903	1.09861
35	1.49978	1.15924
36	1.56238	1.22378
37	1.57750	1.29277
38	1.67690	1.36688
39	1.89913	1.44692
40	2.26144	1.53393
41	2.43108	1.62924
42	2.52184	1.73460
43	2.72629	1.85238
44	2.74447	1.98592
45	3.15556	2.14007
46	3.30853	2.32239
47	3.50445	2.54553
48	3.91535	2.73321
49	4.06478	3.23868
50	4.37295	3.93183

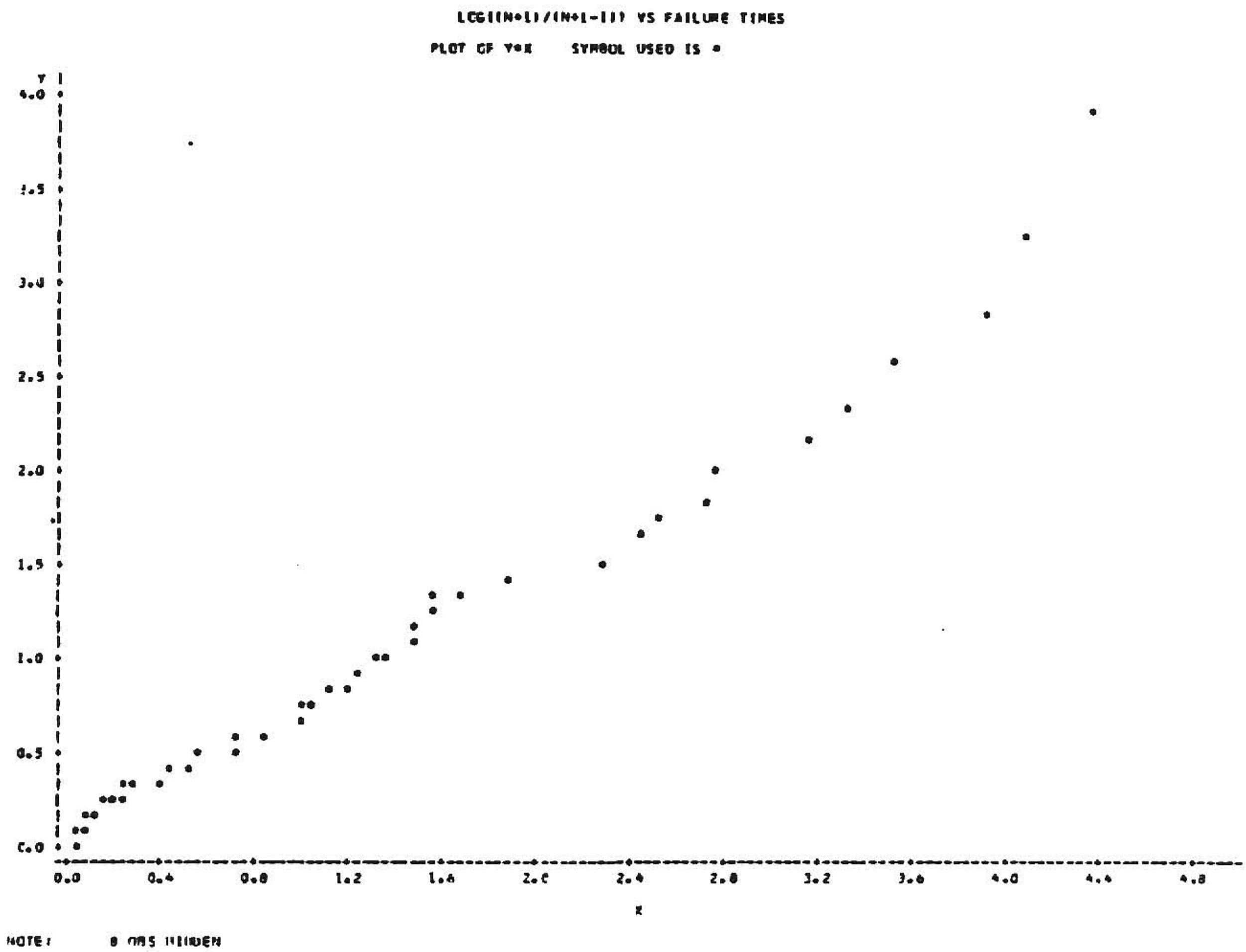


Figure 2.1.16 Plot of dataset 1 (sample size = 50)

Table 2.1.17 Data Set 2 (sample size = 50)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X1	Y
1	0.01852	0.01980
2	0.01902	0.04001
3	0.02281	0.06062
4	0.02327	0.08168
5	0.02845	0.10318
6	0.15754	0.12516
7	0.16435	0.14764
8	0.17140	0.17063
9	0.22621	0.19416
10	0.23057	0.21825
11	0.24858	0.24295
12	0.32050	0.26826
13	0.39907	0.29424
14	0.45751	0.32091
15	0.46537	0.34831
16	0.47904	0.37648
17	0.49944	0.40547
18	0.53319	0.43532
19	0.53728	0.46609
20	0.55087	0.49784
21	0.65173	0.53063
22	0.74825	0.56453
23	0.82875	0.59962
24	0.83169	0.63599
25	0.87043	0.67373
26	0.92906	0.71295
27	1.02647	0.75377
28	1.08663	0.79633
29	1.12286	0.84078
30	1.18616	0.88730
31	1.19689	0.93609
32	1.22043	0.98739
33	1.29238	1.04145
34	1.33868	1.09861
35	1.37051	1.15924
36	1.37197	1.22378
37	1.41462	1.29277
38	1.48398	1.36688
39	1.49245	1.44692
40	1.53748	1.53393
41	1.58167	1.62924
42	1.63671	1.73460
43	1.67498	1.85238
44	1.98709	1.98592
45	2.16141	2.14007
46	2.53872	2.32239
47	3.32474	2.54553
48	3.76999	2.83321
49	4.32203	3.23868
50	4.86960	3.93183

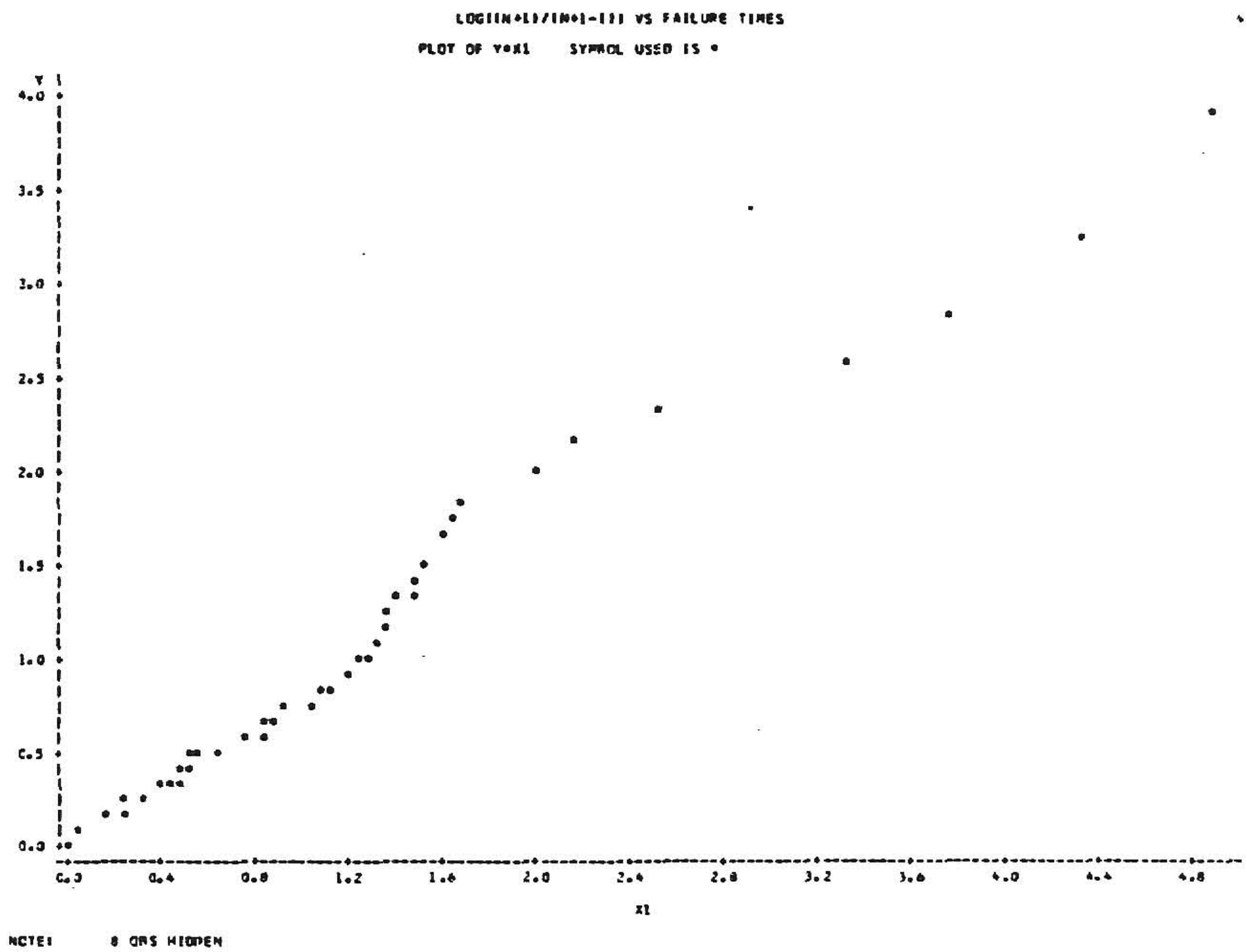


Figure 2.1.17 Plot of dataset 2 (sample size = 50)

Table 2.1.18 Data Set 3 (sample size = 50)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X2	Y
1	0.02182	0.01980
2	0.05353	0.04001
3	0.06026	0.06062
4	0.08104	0.08168
5	0.09522	0.10318
6	0.10538	0.12516
7	0.11184	0.14764
8	0.11766	0.17063
9	0.12080	0.19416
10	0.16213	0.21825
11	0.16961	0.24295
12	0.19741	0.26826
13	0.24590	0.29424
14	0.29166	0.32091
15	0.33722	0.34831
16	0.35136	0.37648
17	0.39199	0.40547
18	0.41594	0.43532
19	0.43049	0.46609
20	0.47681	0.49784
21	0.49916	0.53063
22	0.53482	0.56453
23	0.53952	0.59962
24	0.55699	0.63599
25	0.63400	0.67373
26	0.63641	0.71295
27	0.69829	0.75377
28	0.71638	0.79633
29	0.80364	0.84078
30	0.82270	0.88730
31	0.82486	0.93609
32	0.84901	0.98739
33	0.85973	1.04145
34	0.90453	1.09861
35	0.95816	1.15924
36	0.96159	1.22378
37	1.00246	1.29277
38	1.11868	1.36688
39	1.29177	1.44692
40	1.41691	1.53393
41	1.55573	1.62924
42	1.60208	1.73460
43	1.61488	1.85239
44	1.64903	1.98592
45	1.71609	2.14007
46	1.89214	2.32239
47	1.92130	2.54553
48	1.97827	2.83321
49	2.07728	3.23868
50	2.55205	3.93183

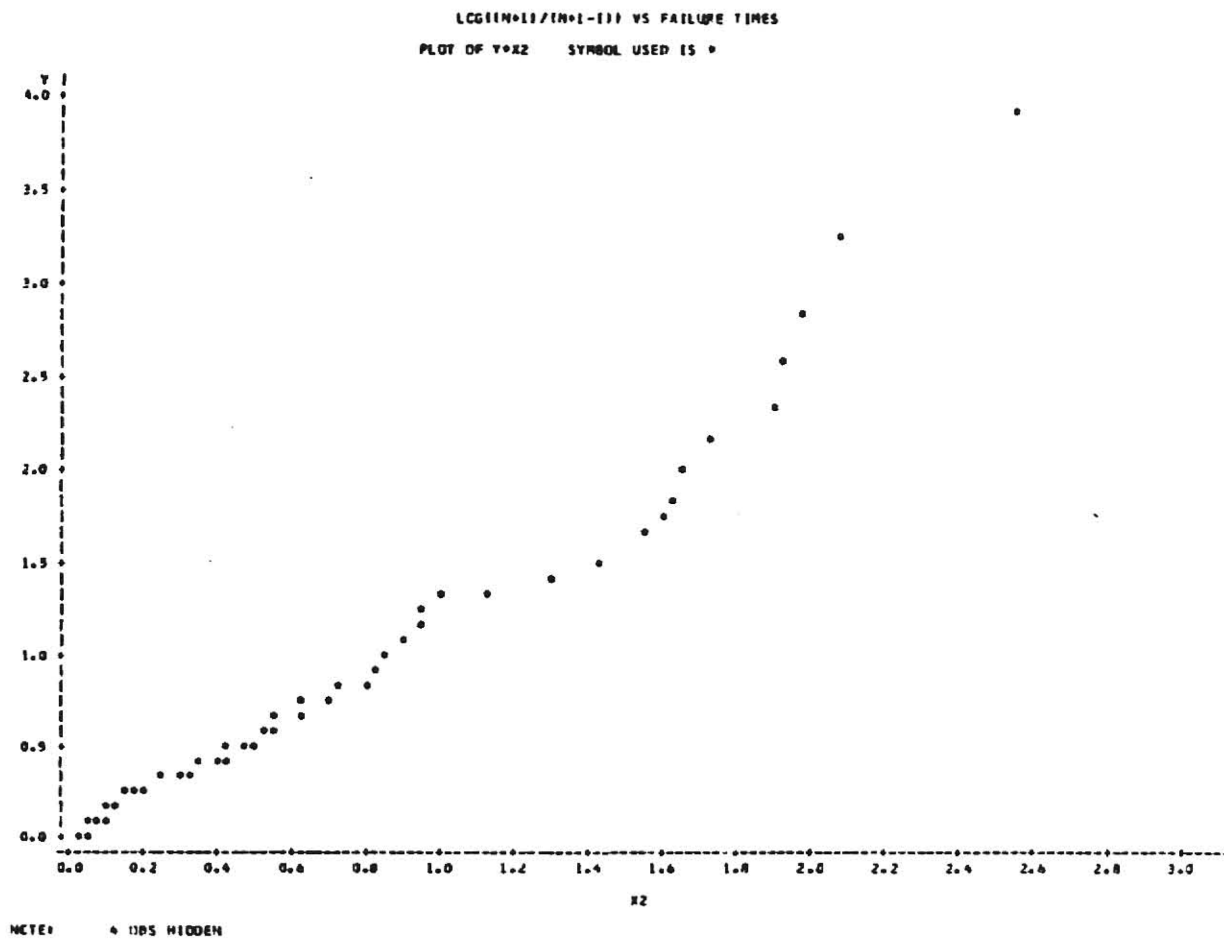


Figure 2.1.18 Plot of dataset 3 (sample size = 50)

Table 2.1.19 Data Set 4 (sample size = 50)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X3	Y
1	0.00210	0.01980
2	0.01023	0.04001
3	0.02795	0.06062
4	0.04604	0.08168
5	0.08457	0.10318
6	0.08751	0.12516
7	0.10336	0.14764
8	0.15296	0.17063
9	0.21063	0.19416
10	0.22225	0.21825
11	0.29797	0.24295
12	0.34192	0.26826
13	0.34703	0.29424
14	0.35079	0.32091
15	0.35634	0.34831
16	0.45052	0.37648
17	0.46028	0.40547
18	0.51310	0.43532
19	0.51511	0.46609
20	0.59966	0.49784
21	0.60269	0.53063
22	0.60326	0.56453
23	0.75932	0.59962
24	0.76728	0.63599
25	0.80802	0.67373
26	0.81352	0.71295
27	0.86330	0.75377
28	0.89288	0.79633
29	0.91031	0.84078
30	0.96316	0.88730
31	1.03832	0.93609
32	1.19460	0.98739
33	1.20143	1.04145
34	1.34552	1.09861
35	1.52964	1.15924
36	1.60801	1.22378
37	1.80204	1.29277
38	1.85579	1.36688
39	1.88467	1.44692
40	1.93825	1.53393
41	1.96809	1.62924
42	1.98569	1.73460
43	2.41192	1.85238
44	2.52613	1.98592
45	2.54387	2.14007
46	2.68461	2.32239
47	3.11769	2.54553
48	3.22413	2.83321
49	3.23058	3.23868
50	6.95263	3.93183

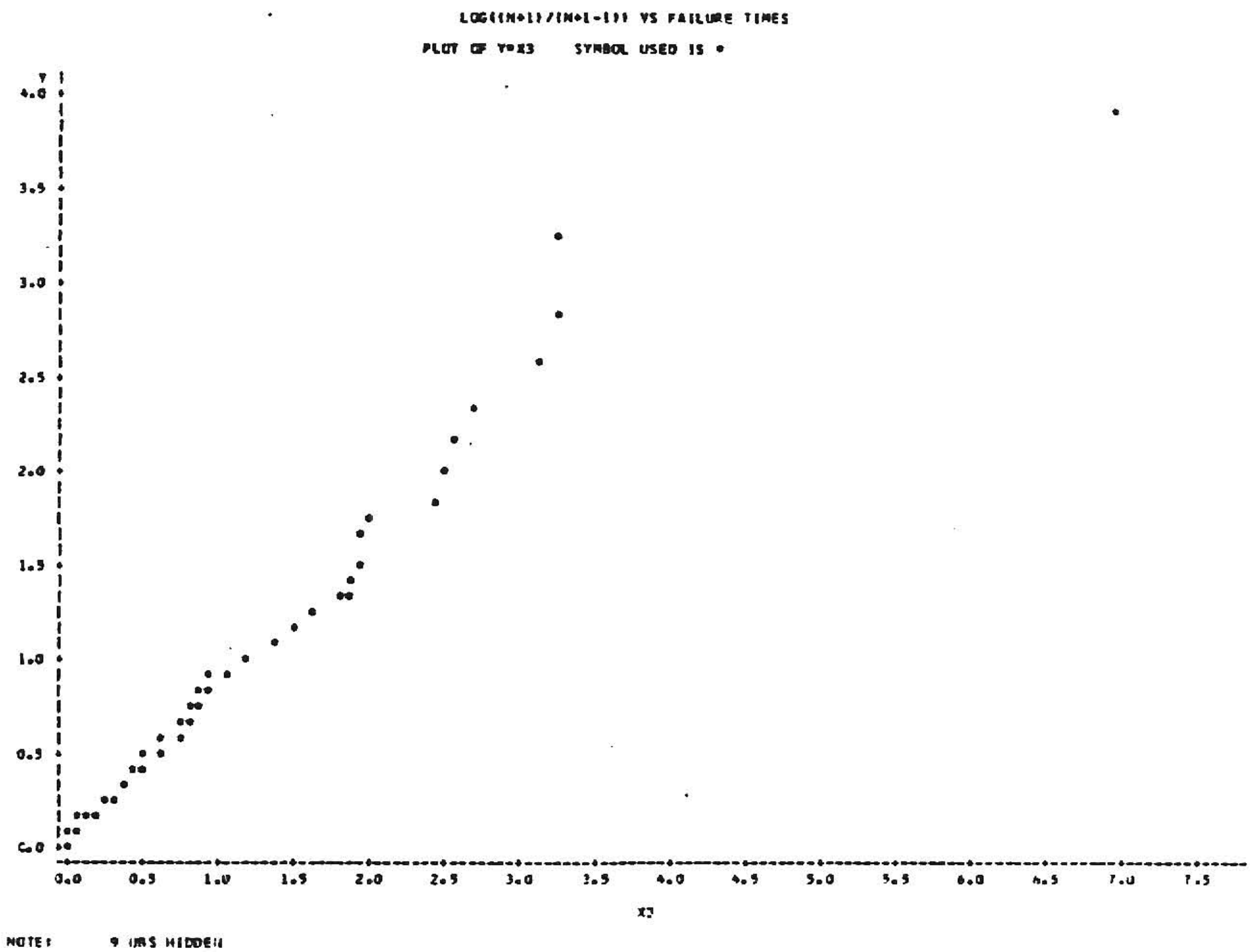


Figure 2.1.19 Plot of dataset 4 (sample size = 50)

Table 2.1.20 Data Set 5 (sample size = 50)

LOG((N+1)/(N+1-I)) VS FAILURE TIMES

OBS	X4	Y
1	0.00387	0.01980
2	0.00686	0.04001
3	0.04405	0.06062
4	0.05044	0.08168
5	0.07194	0.10318
6	0.07280	0.12516
7	0.08546	0.14764
8	0.14757	0.17063
9	0.14874	0.19416
10	0.15656	0.21825
11	0.25896	0.24295
12	0.26980	0.26826
13	0.27810	0.29424
14	0.33537	0.32091
15	0.40046	0.34831
16	0.41449	0.37648
17	0.43765	0.40547
18	0.44207	0.43532
19	0.46167	0.46609
20	0.47080	0.49784
21	0.49578	0.53063
22	0.56190	0.56453
23	0.56313	0.59962
24	0.57376	0.63599
25	0.66395	0.67373
26	0.66509	0.71295
27	0.74696	0.75377
28	0.77879	0.79633
29	0.83914	0.84078
30	0.85240	0.88730
31	0.94614	0.93609
32	0.95521	0.98739
33	1.00163	1.04145
34	1.09660	1.09861
35	1.14675	1.15924
36	1.29058	1.22378
37	1.46936	1.29277
38	1.51610	1.36688
39	1.73552	1.44692
40	2.16483	1.53393
41	2.23380	1.62924
42	2.45509	1.73460
43	2.50319	1.85238
44	2.66732	1.98592
45	2.85948	2.14007
46	3.15164	2.32239
47	3.31353	2.54553
48	4.25355	2.83321
49	6.12146	3.23869
50	6.81957	3.93183

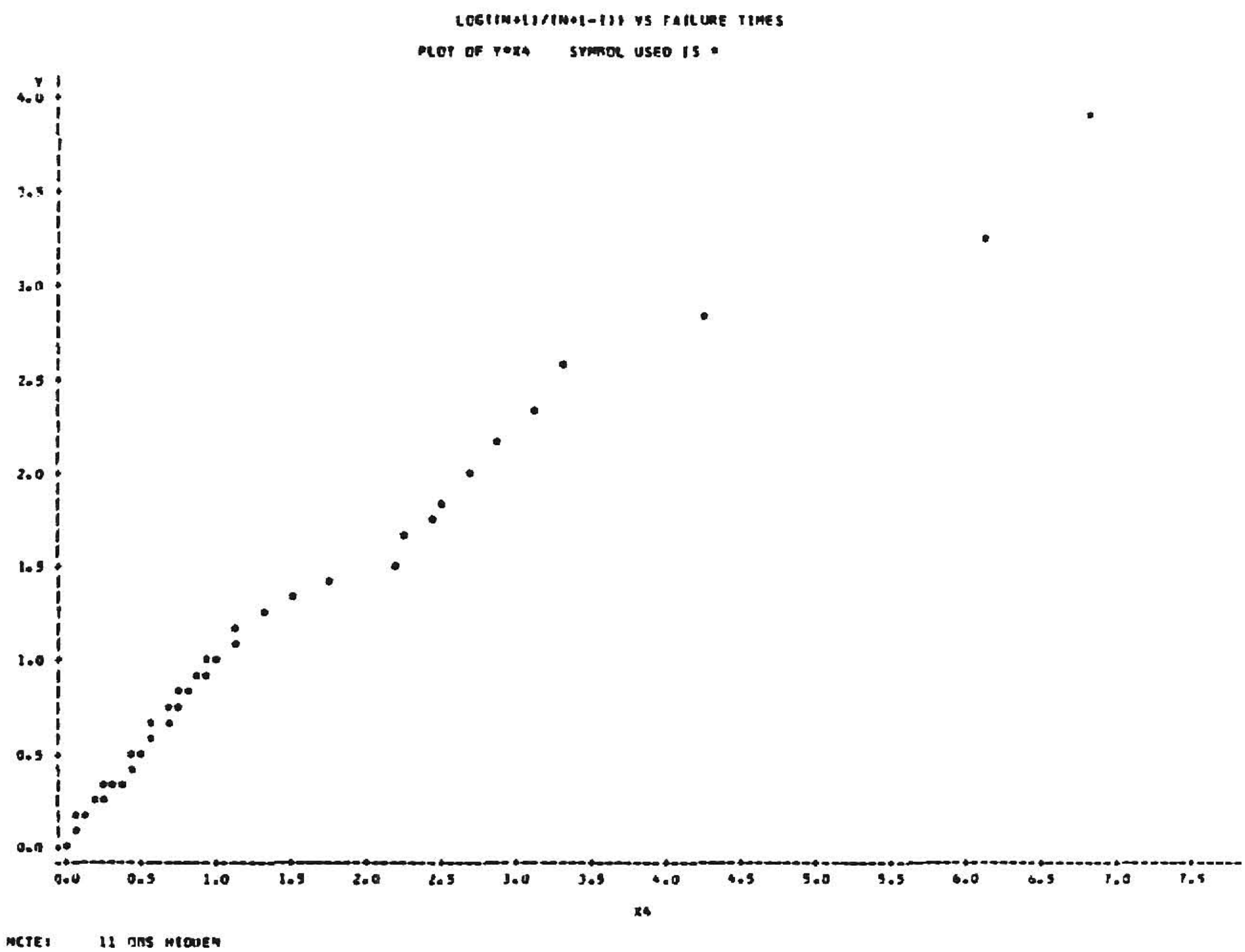


Figure 2.1.20 Plot of dataset 5 (sample size = 50)

SECTION 3.

A TEST FOR ABNORMALLY EARLY OR LATE FIRST FAILURE.

In life testing it is possible that the underlying distribution of life is exponential. However, it is of interest to observe if the first few failures occur abnormally early or late. Two situations are considered. Let 'A' be a constant, a value by which a particular exponential distribution is shifted. The two situations involve the cases where :

1) A=0, the one parameter exponential.
2) A>0, the two parameter exponential.

In the first of these cases several simulated samples are studied to study the occurence of abnormally early failures. An attempt is made to compare and contrast them with reference to the theories propounded by Epstein.

Let t_i represent the i^{th} random failure time generated by the exponential distribution. Each t_i is a realization(sample observation) of a common random variable t for which we can write either $t \sim NGEX(\lambda)$ or $t \sim G(1,\theta)$. Let τ_i represent the i^{th} ordered failure time. Suppose that

$$\tau_1 \leq \tau_2 \leq - - - + \tau_n$$

are the first r ordered failure times and we wish to test if τ_1 is abnormally small. The total test time is

$$T = \tau_1 + \tau_2 + - - - + \tau_n$$

From (1) and (2) Mark Sobel and Benjamin Epstein prove that if all the t_i are drawn from a common exponential distribution, then $T(\tau_1)$ (the total life in the interval $(0,\tau_1)$) and $T(\tau_r - \tau_1)$ (the total remaining life of all the components after the first failure), are distributed independently of each other . T has the distribution

$$T \sim G(n, \theta)$$

and hence it can be proved that

$$\frac{2T}{\theta} \sim G(n,2) = \chi^2(2n)$$

It can also be proved that

$$\frac{2T(\tau_1)}{\theta} \sim \chi^2(2)$$

and $$\frac{2T(\tau_r - \tau_1)}{\theta} \sim \chi^2(2r-2)$$

The ratio of two χ^2 distributions gives an F distribution and the ratio

$$F_c = \frac{(r-1)T(\tau_1)}{2T(\tau_r - \tau_1)}$$

is distributed as $F(2,2r-2)$. It can therefore reasonably be asserted that τ_1 is abnormally small if this ratio is too small, i.e. less than the critical point $F^{**} = F_{1-\alpha}(2,2r-2)$. Similarly, τ_1 can be judged abnormally large if F_c exceeds the critical value $F^* = F_\alpha(2,2r-2)$. If the value of F_c is less than the critical F^{**} value then there is sufficient evidence to accept the hypothesis that τ_1 is abnormally small. Consequently smaller ratios tend to support the fact that τ_1 is indeed an early failure.

DISCUSSION

Appendix 2 contains a program written in Fortran G that generates random exponential deviates. These values, along with the corresponding values for the log function, are passed into a SAS program which plots their values. The F value comparisons based on the above theory are incorporated in the Fortran program and the p values(also called $1 - \hat{\alpha}$)that give the left hand tail area of an F distribution with n1 and n2 degrees of freedom have also been determined. The (1-p) value gives the critical level at which the observation would be marginally significant. Tables 3.1.1a to 3.9.2a provide samples of size 10, 20 and 50 that are typical of the situation discussed here.

CASE 1. A=0.

All deviates were generated from an exponential distribution with θ=100. Considering the first data set that was generated (Tables 3.1.1a and 3.1.2a), the F_c value of .1668 is much less than F* = 3.55. The value at which F would be marginally significant would be at the (1-p)=(1-0.1523)=0.8467 percent level. The first failure time of 1.2899 does not provide enough evidence that an abnormally a late failure has occurred. The F** value is (1/3.55)=0.282. The F_c value is also less than the F** value which provides sufficient evidence to accept the hypothesis that τ_1 is abnormally small. We conclude that an abnormally early failure has indeed occurred. Consider data set 2 with sample size 10 (Table 3.2.1a). An examination of the first failure time of 53.3 and the corresponding values beyond it gives rise to the suspicion that an abnormally late first failure has occurred. The ratio F_c is 4.2190. This F_c value is greater than $F^*=3.55$ though not very much so. Clearly the hypothesis of an abnormally early failure must be rejected. The associated p value is 0.9686 and the value of F would be marginally significant at $(1-p)=(1-0.9686)=0.0314$ percent level. A visual examination of the data indicates however that the first failure time of 53.265 was seemingly very late and any inexperienced observer could look at it and in comparison with the rest of the data pronounce that a late failure has indeed occurred. It seems to be rather obvious that a rejection by the F test proposed by Epstein could be brought about only by virtue of the first failure time being extremely large. Such a situation could be easily recognized by a fairly experienced analyst without resort to the test that Epstein recommends. Several more samples indicate a similar situation. Varying the sample size using 20 and then 50 did not prove to be any different from the situations cited above.

CASE 2. A>0.

In the second of the cases discussed in this section letting A be greater than zero gives rise to the two parameter exponential whose density function is of the form :

$$f(t,A) = 1/\theta \, e^{-(t-A)/\theta} \qquad t \geq A \geq 0$$

$$= 0 \qquad \text{elsewhere}$$

and the associated cumulative distribution function is

$$F(t;A) = 0 \qquad t < 0$$

$$= 1-e^{-(t-A)/\theta} \qquad t \geq 0$$

It follows that

$$\log \left(\frac{1}{1-F(t;A)}\right) = \frac{(t-A)}{\theta}$$

Using the same reasoning as in section 2, the values of the individual τ are plotted against

$$y = \log[\,(n+1)/(n+1-i)\,]$$

For the two parameter exponential a straight line shifted A units to the right on the abscissa is obtained.

As mentioned in section 2 Epstein goes on to show that the ratio

$$\frac{(r-1)\ T\ (\tau_1)}{2T(\tau_r - \tau_1)}$$

is distributed as F (2,2r-2) and it can reasonably be asserted that τ_1 is abnormally large if this ratio is too large. More precisely if α is the significance level, the computed ratio F_c could be compared with an $F^* = F_\alpha(2,2r-2)$. If the value of the ratio is greater than F* then the hypothesis (that τ_1 is abnormally large) is supported. Smaller values of the F_c ratio is an indication that an abnormally large first failure is not present.

DISCUSSION

Appendix 3 contains the program that generates random exponential deviates with $\theta = 100$. The first value is incremented by the constant A for purposes of simulating an abnormally late first failure. The value of A ranges from 20 to 60, although only values from 20 to 40 have been displayed here for the purpose of convenience. Here again samples of 10, 20 and 50 were considered. Appendix 3 also has a SAS program that prints and plots these values against the log function y. Table 3.1.1b is the Fortran output that displays a sample of size 10. The first failure time is 21.2899 and the ratio of the first partial life to the total remaining life suitably weighted gives an F_c value of 2.7524. This value is lower than the F* value of 3.55 with 2 and (2r-2) degrees of freedom and so there is no reason to believe that an abnormally late first failure has occurred. It is also larger than the F** value of 0.282 which does not give us enough evidence to accept the hypothesis that τ_1 is an abnormally early failure. On the other hand consider Table 3.2.1b. On examination one is bound to notice the first failure time of 83.2651 and to feel that an abnormally long first failure has occurred. The calculated F_c value is 6.5952 which is far above the F* value of 3.55. The hypothesis that a late first failure has occurred is very teneble. The p value is 0.9929 and so F would be marginally significant at $(1-p)=(1-0.9929)=0.0071$ percent level. A visual examination of the data indicates that the first failure time of 83.2651 in comparison with the rest of the data is rather large and even a relatively inexperienced observer could conclude that an abnormally late failure has occurred without much statistical reasoning. It seems obvious that the acceptance of an abnormally long first failure can only be brought about by an extremely large value for the first failure. Such large values can easily be detected by visual examination and the F test of the kind used by Epstein may be superflous. Tables and graphs contained in the rest of this section corroborate this fact.

Table 3.1.1a Data Set 1 (sample size = 10)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
1.2899	0.0953
3.4364	0.2007
10.9396	0.3185
16.8621	0.4520
26.6231	0.6061
50.8577	0.7885
56.3283	1.0116
96.5339	1.2993
134.4343	1.7047
311.7527	2.3979

FIRST PARTIAL TOTAL LIFE : 12.8988

TOTAL REMAINING LIFE : 696.1587

VALUE OF F : 0.1668

P VALUE : 0.1523

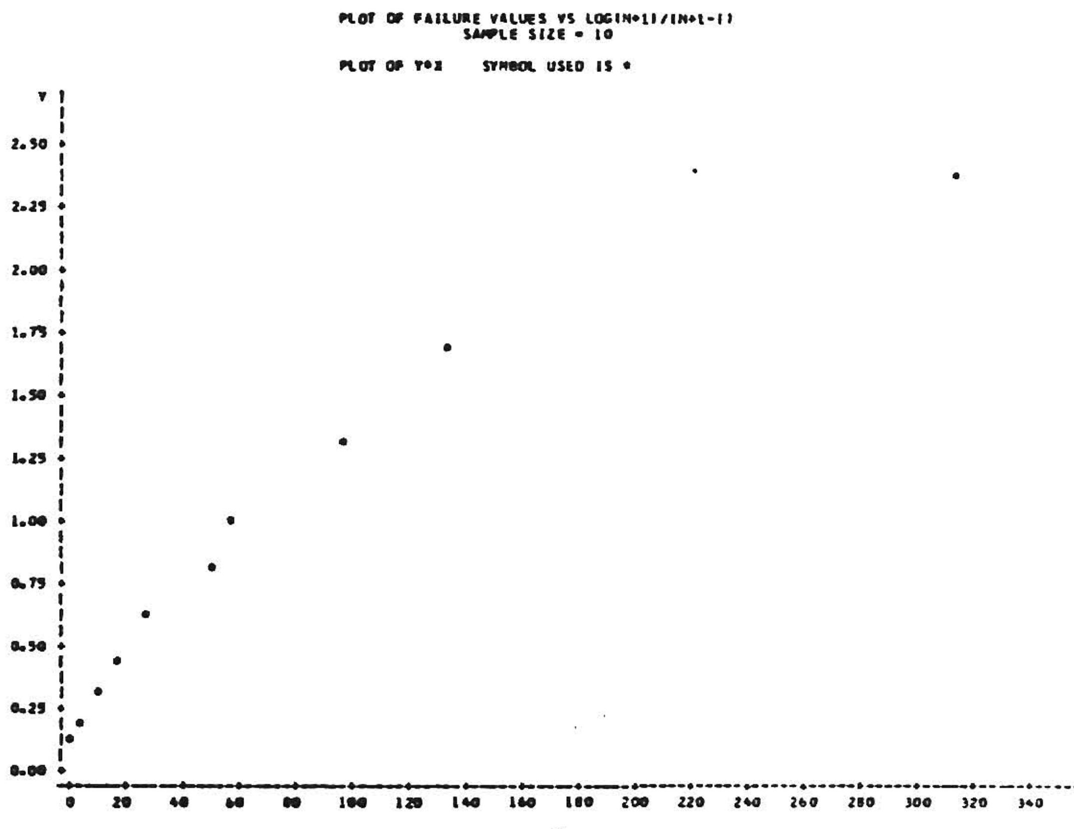


Figure 3.1.1a Plot of dataset 1 (sample size = 10)

Table 3.2.1a Data Set 2 (sample size = 10)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
53.2651	0.0953
71.5935	0.2007
92.8933	0.3185
106.1330	0.4520
154.6098	0.6061
162.5727	0.7885
177.7564	1.0116
195.1102	1.2993
241.3404	1.7047
413.6375	2.3979

FIRST PARTIAL TOTAL LIFE : 532.6514

TOTAL REMAINING LIFE : 1136.2598

VALUE OF F : 4.2190

P VALUE : 0.9686

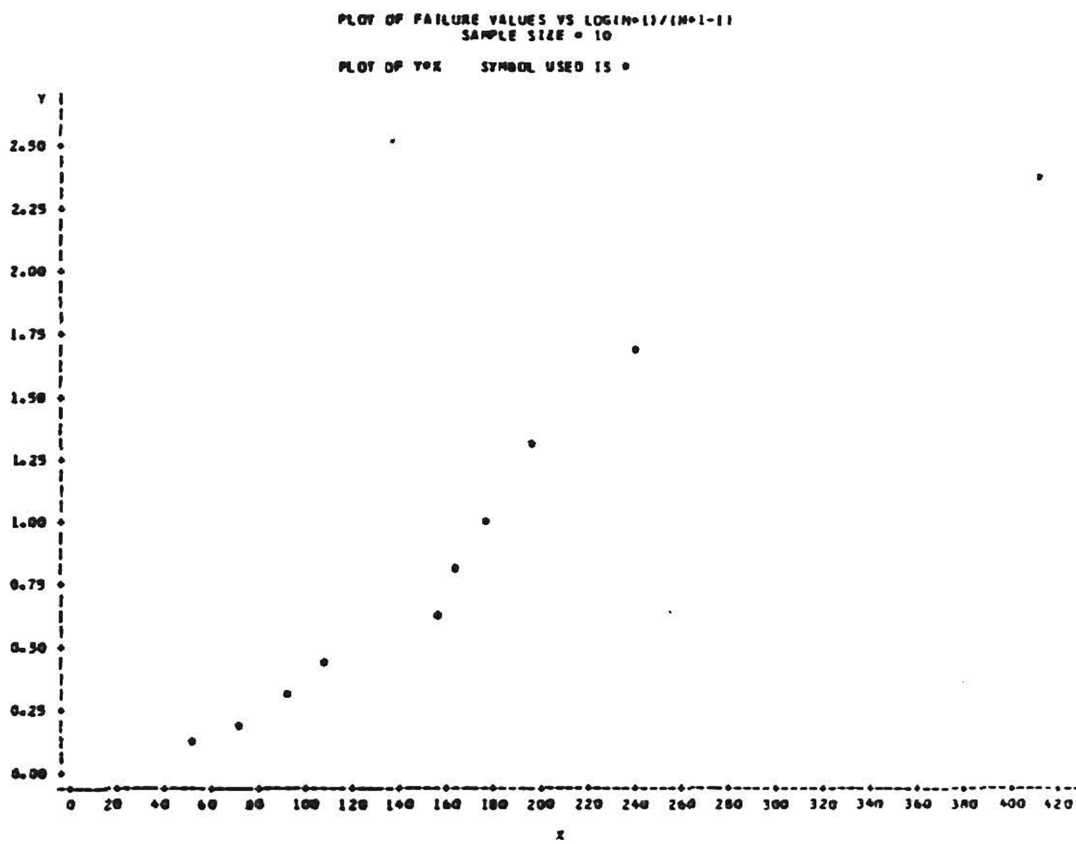


Figure 3.2.1a Plot of dataset 2 (sample size = 10)

Table 3.3.1a Data Set 3 (sample size = 10)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
10.6750	0.0953
10.7218	0.2007
11.9664	0.3185
13.5444	0.4520
34.3169	0.6061
56.9252	0.7885
123.9955	1.0116
148.8099	1.2993
307.7061	1.7047
342.3479	2.3979

FIRST PARTIAL TOTAL LIFE : 106.7505

TOTAL REMAINING LIFE : 954.2581

VALUE OF F : 1.0068

P VALUE : 0.6149

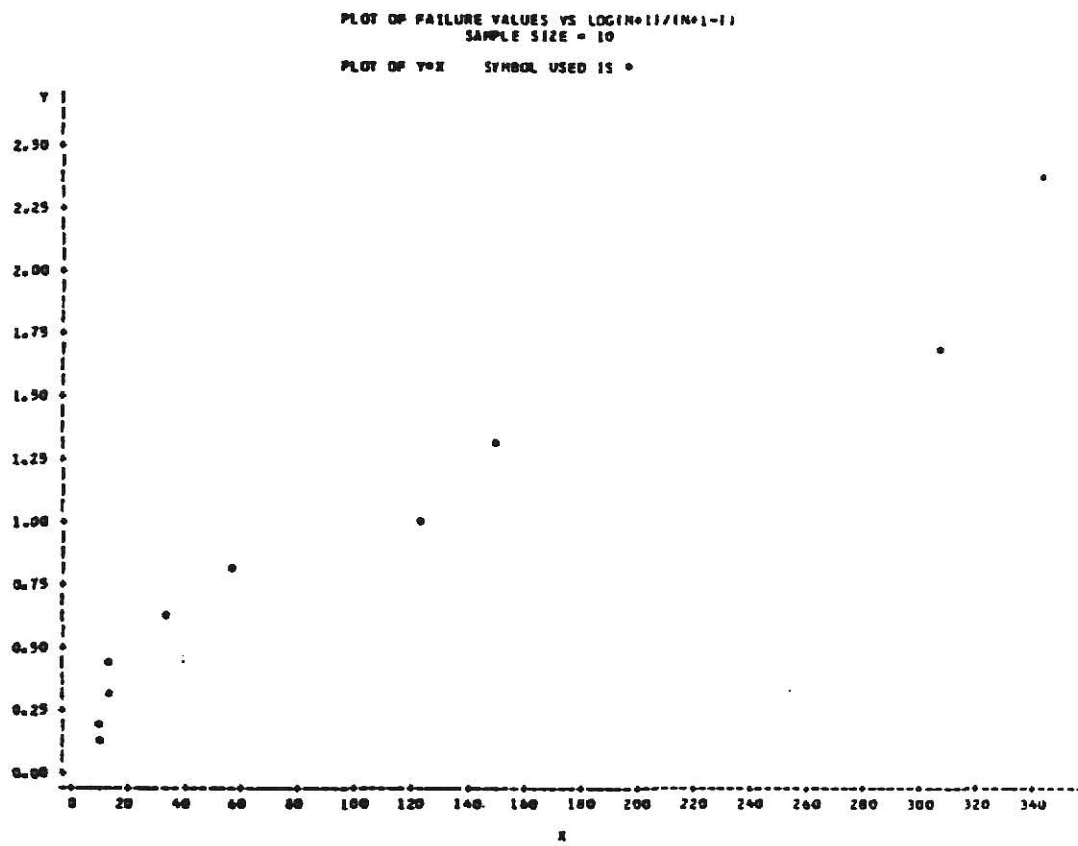


Figure 3.3.1a Plot of dataset 3 (sample size = 10)

Table 3.4.1a Data Set 1 (sample size = 20)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
1.2899	0.0488
3.4364	0.1001
10.9396	0.1542
16.8621	0.2113
26.6231	0.2719
50.8577	0.3365
53.2651	0.4055
56.3283	0.4796
71.5935	0.5596
92.8933	0.6466
96.5339	0.7419
106.1330	0.8473
134.4343	0.9651
154.6098	1.0986
162.5727	1.2528
177.7564	1.4351
195.1102	1.6582
241.3404	1.9459
311.7527	2.3514
413.6375	3.0445

FIRST PARTIAL TOTAL LIFE : 25.7977

TOTAL REMAINING LIFE : 2352.1707

VALUE OF F : 0.2084

P VALUE : 0.1872

Figure 3.4.1a Plot of dataset 1 (sample size = 20)

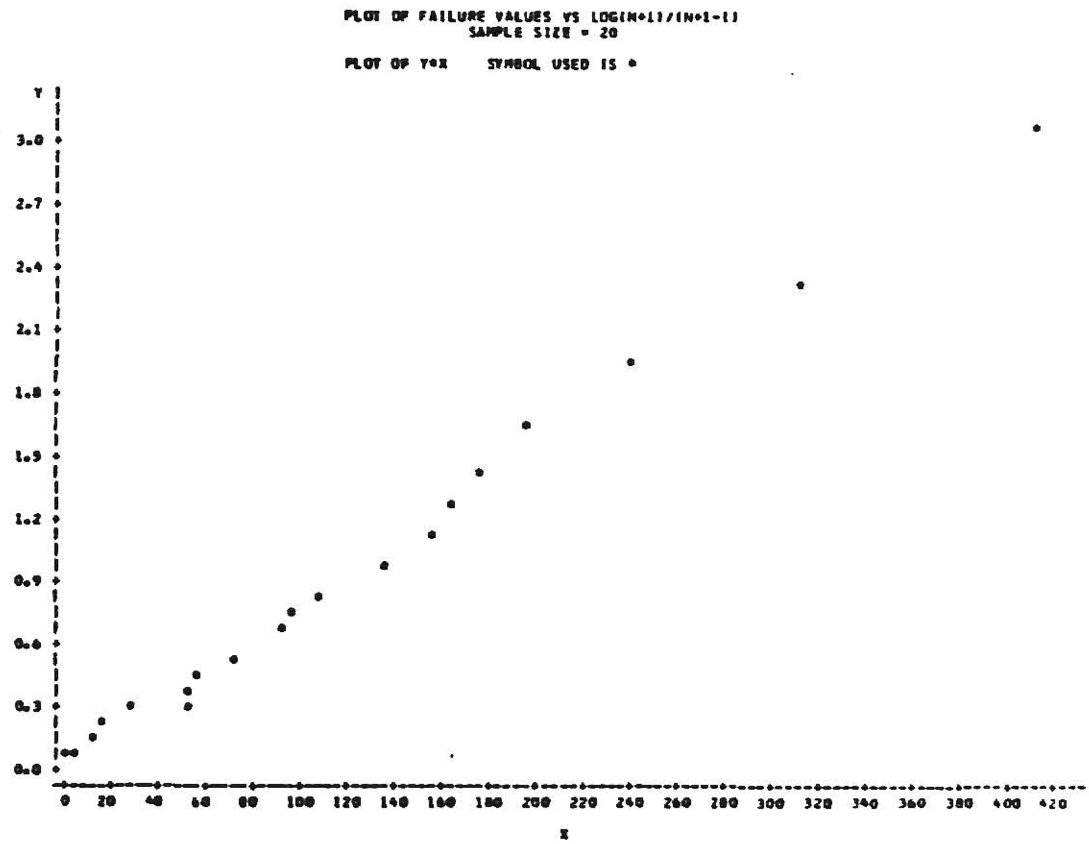


Table 3.5.1a Data Set 2 (sample size = 20)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
10.6750	0.0488
10.7218	0.1001
11.9664	0.1542
13.5444	0.2113
19.0593	0.2719
20.7202	0.3365
34.3169	0.4055
38.0696	0.4796
41.6007	0.5596
41.6669	0.6466
48.4067	0.7419
56.9252	0.8473
68.8148	0.9651
73.6535	1.0986
77.3869	1.2528
123.9955	1.4351
148.8099	1.6582
177.5704	1.9459
307.7061	2.3514
342.3479	3.0445

FIRST PARTIAL TOTAL LIFE : 213.5010

TOTAL REMAINING LIFE : 1454.4561

VALUE OF F : 2.7890

P VALUE : 0.9259

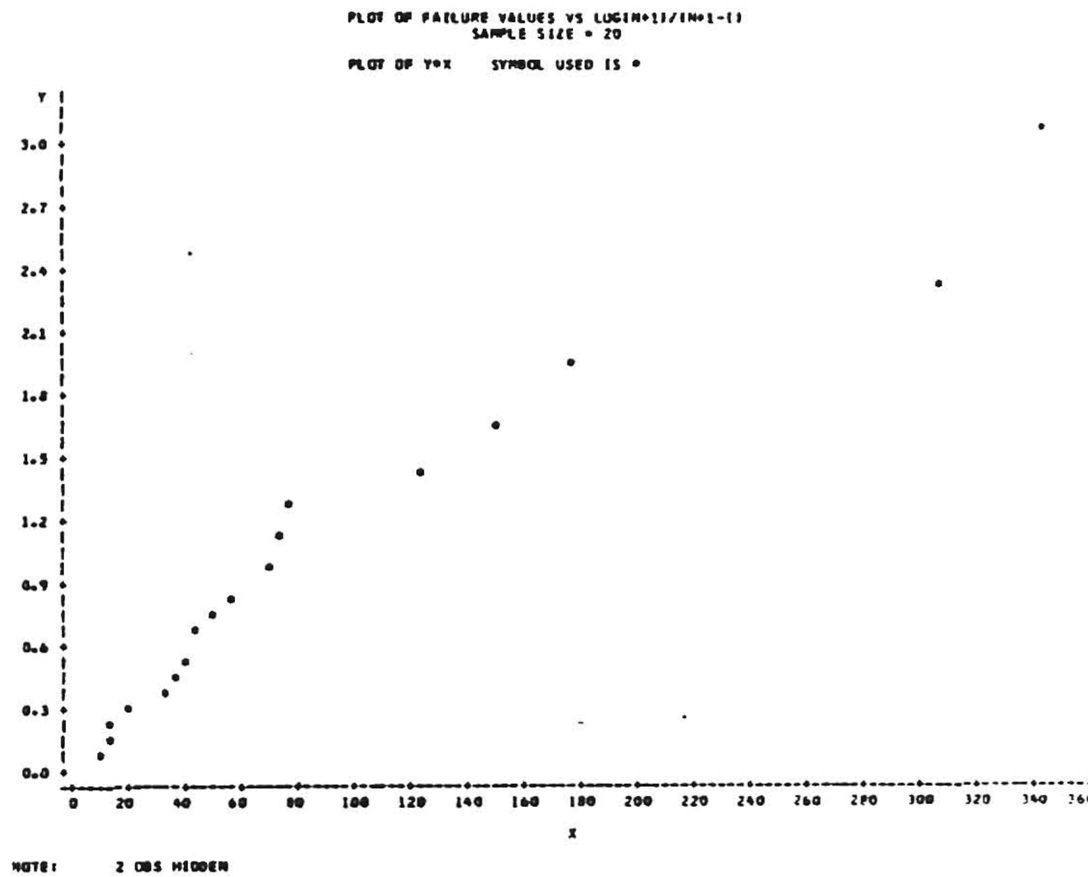


Figure 3.5.1a Plot of dataset 2 (sample size = 20)

Table 3.6.1a Data Set 3 (sample size = 20)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
2.3725	0.0488
4.6707	0.1001
4.9013	0.1542
8.5276	0.2113
25.2894	0.2719
31.8289	0.3365
34.1526	0.4055
38.6104	0.4796
54.2916	0.5596
63.7366	0.6466
72.3403	0.7419
73.5133	0.8473
110.7111	0.9651
124.4890	1.0986
174.9946	1.2528
192.5572	1.4351
225.7567	1.6582
240.5881	1.9459
296.9680	2.3514
314.8708	3.0445

FIRST PARTIAL TOTAL LIFE : 47.4494

TOTAL REMAINING LIFE : 2047.7200

VALUE OF F : 0.4403

P VALUE : 0.3529

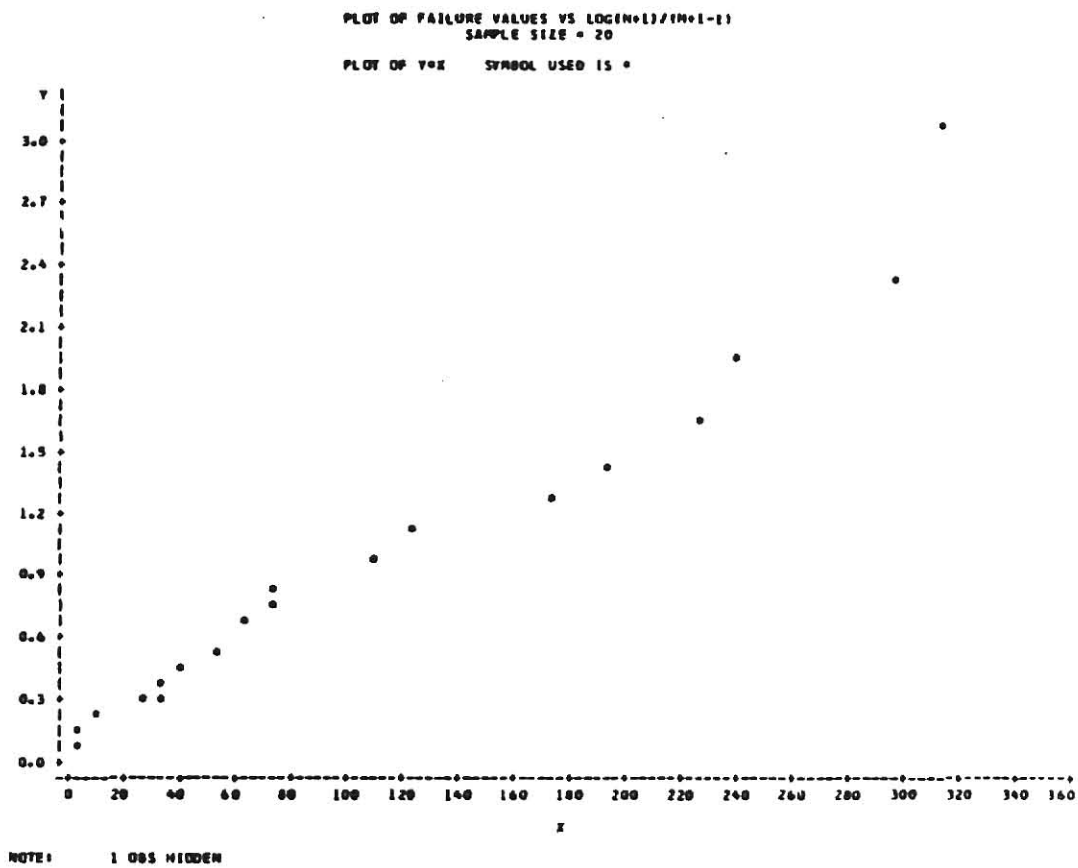


Figure 3.6.1a Plot of dataset 3 (sample size = 20)

Table 3.7.1a Data Set 1 (sample size = 50)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
2.1432	0.0198
2.3725	0.0400
4.9013	0.0606
8.5276	0.0817
12.4412	0.1032
12.8290	0.1252
13.3276	0.1476
15.4369	0.1706
16.2734	0.1942
17.3833	0.2183
17.4855	0.2429
22.8405	0.2683
25.2894	0.2942
27.8629	0.3209
37.4231	0.3483
46.8214	0.3765
49.2653	0.4055
55.4600	0.4353
57.5969	0.4661
61.2964	0.4978
61.5988	0.5306
73.5133	0.5645
77.7005	0.5996
91.0238	0.6360
97.0180	0.6737
98.6029	0.7129
106.2859	0.7538
110.7111	0.7963
123.9681	0.8408
132.3256	0.8873
136.3790	0.9361
145.5964	0.9874
146.3358	1.0415
152.1531	1.0986
166.5521	1.1592
168.1229	1.2238
174.9946	1.2928
183.2989	1.3669
192.5572	1.4469
194.6810	1.5339
203.2502	1.6292
221.1299	1.7346
225.7567	1.8524
227.8092	1.9859
240.5881	2.1401
242.8024	2.3224
263.9922	2.5455
266.1216	2.8332
329.2297	3.2387
373.0393	3.9318

FIRST PARTIAL TOTAL LIFE : 107.1599

TOTAL REMAINING LIFE : 5626.9180 P VALUE : 0.6032

VALUE OF F : 0.9332

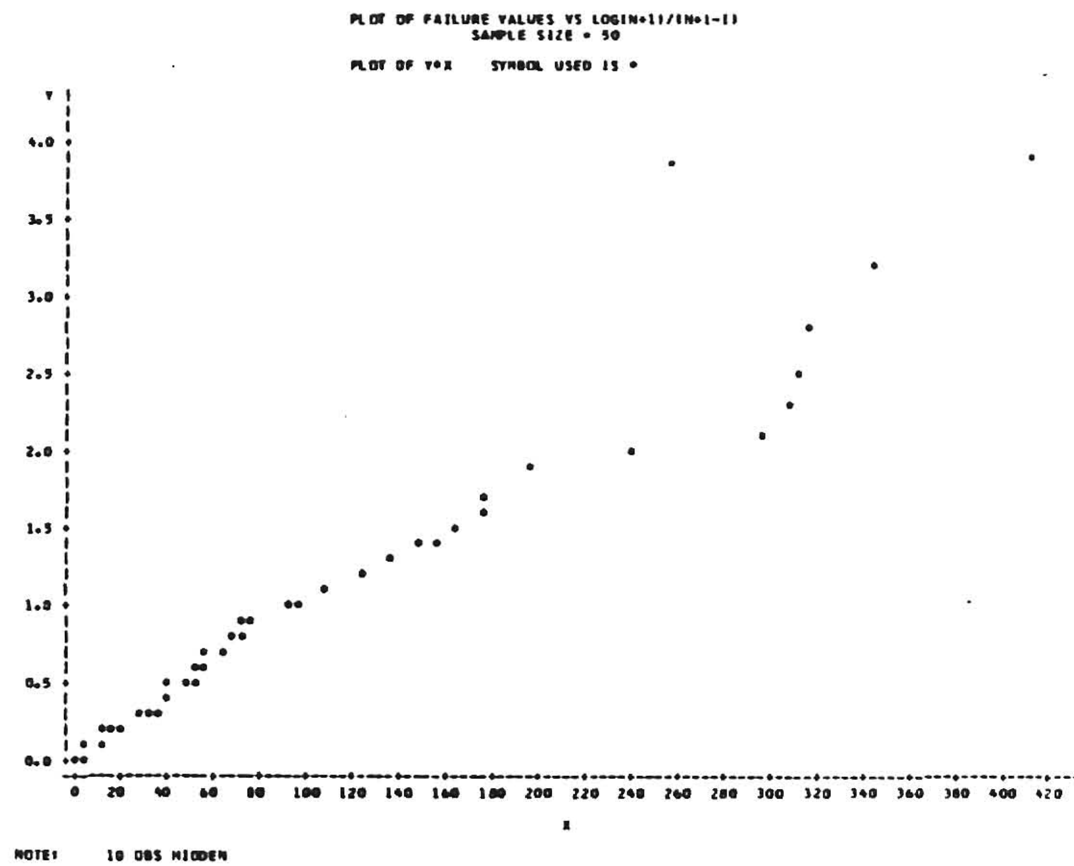


Figure 3.7.1a Plot of dataset 1 (sample size = 50)

Table 3.8.1a Data Set 2 (sample size = 50)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
1.2899	0.0198
3.4364	0.0400
4.6707	0.0606
10.6750	0.0817
10.7218	0.1032
10.9396	0.1252
11.9664	0.1476
13.5444	0.1706
16.8621	0.1942
19.0593	0.2183
20.7202	0.2429
26.6231	0.2683
31.8289	0.2942
34.1526	0.3209
34.3169	0.3483
38.0696	0.3765
38.6104	0.4055
41.6007	0.4353
41.6669	0.4661
48.4067	0.4978
50.8577	0.5306
53.2651	0.5645
54.2916	0.5996
56.3283	0.6360
56.9252	0.6737
63.7366	0.7129
68.8148	0.7538
71.5935	0.7963
72.3403	0.8408
73.6535	0.8873
77.3869	0.9361
92.8933	0.9874
96.5339	1.0415
106.1330	1.0986
123.9955	1.1592
124.4890	1.2238
134.4343	1.2928
148.8099	1.3669
154.6098	1.4469
162.5727	1.5339
177.5704	1.6292
177.7564	1.7346
195.1102	1.8524
241.3404	1.9859
296.9680	2.1401
307.7061	2.3224
311.7527	2.5455
314.8708	2.8332
342.3479	3.2387
413.6375	3.9318

FIRST PARTIAL TOTAL LIFE : 64.4942

TOTAL REMAINING LIFE : 5017.3711

P VALUE : 0.4652

VALUE OF F : 0.6299

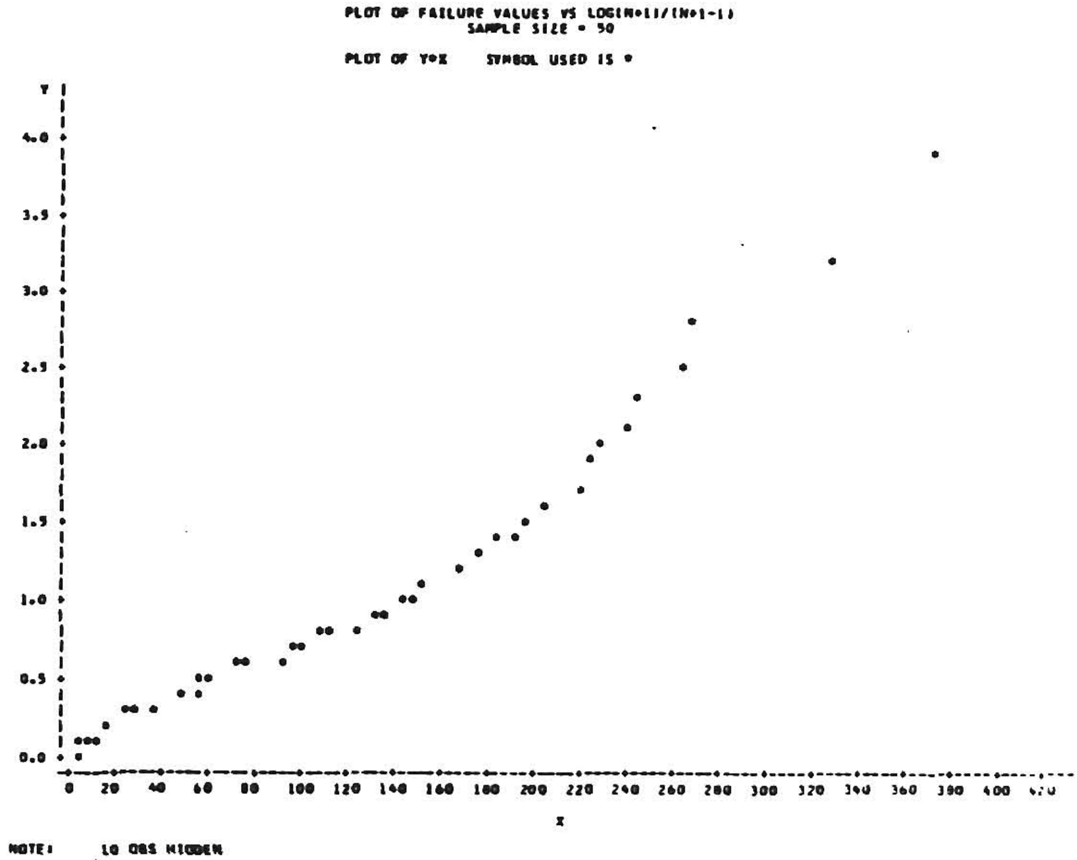


Figure 3.8.1a Plot of dataset 2 (sample size = 50)

Table 3.9.1a Data Set 3 (sample size = 50)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
1.3196	0.0198
1.8175	0.0400
2.7019	0.0606
3.4795	0.0817
4.1330	0.1032
5.2891	0.1252
6.5166	0.1476
8.8340	0.1706
9.7076	0.1942
10.8286	0.2183
14.3622	0.2429
15.0757	0.2683
18.7584	0.2942
21.4171	0.3209
21.8972	0.3483
23.8281	0.3765
25.2010	0.4055
26.1662	0.4353
33.2185	0.4661
36.1799	0.4978
37.9151	0.5306
39.3690	0.5645
40.3130	0.5996
40.9823	0.6360
44.0797	0.6737
50.4448	0.7129
55.5059	0.7538
61.4861	0.7963
63.7433	0.8408
64.2689	0.8873
73.6968	0.9361
86.9809	0.9874
89.9077	1.0415
90.7809	1.0986
93.1424	1.1592
97.3529	1.2238
105.7253	1.2928
113.4026	1.3669
120.7522	1.4469
135.4202	1.5339
145.1620	1.6292
156.4495	1.7346
157.9022	1.8524
162.2135	1.9859
184.2194	2.1401
201.1155	2.3224
221.0051	2.5455
271.4558	2.8332
325.2720	3.2387
396.5349	3.9318

FIRST PARTIAL TOTAL LIFE : 65.9812

TOTAL REMAINING LIFE : 3951.3464 P VALUE : 0.5558

VALUE OF F : 0.8182

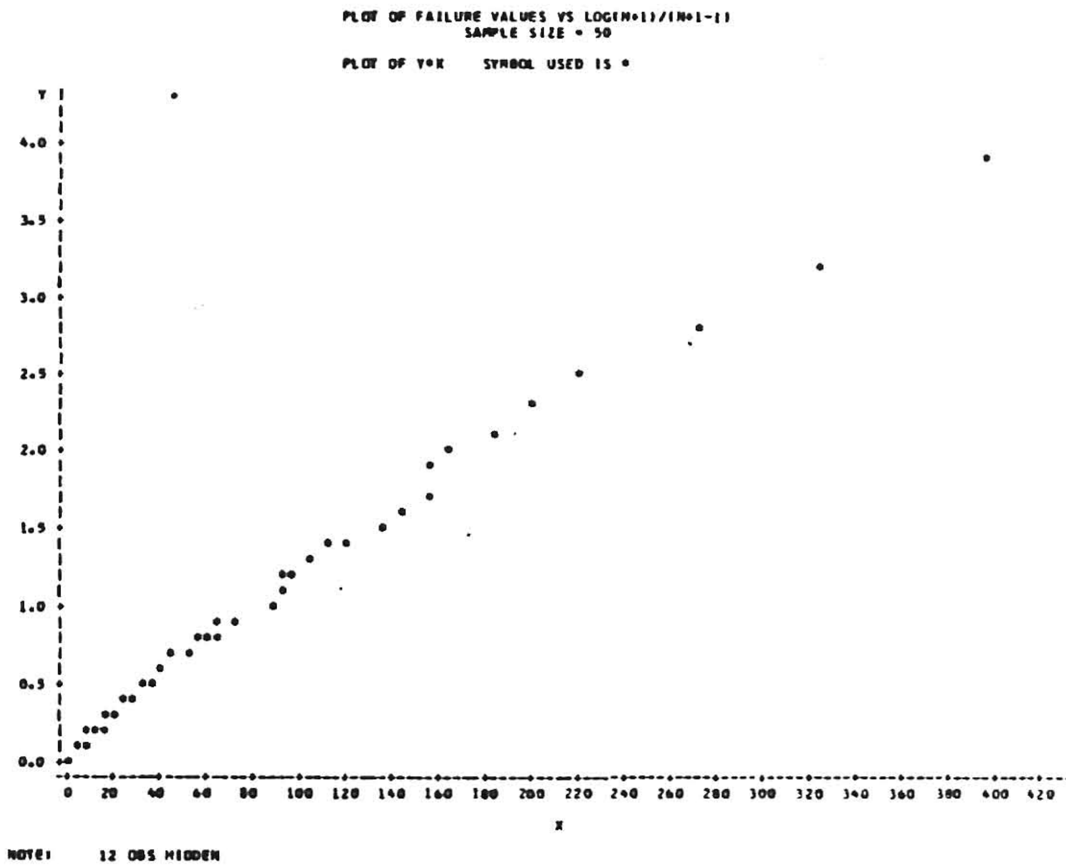


Figure 3.9.1a Plot of dataset 3 (sample size = 50)

Table 3.1.1b Data Set 1 (sample size = 10)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
21.2899	0.0953
23.4364	0.2007
30.9396	0.3185
36.8621	0.4520
46.6231	0.6061
70.8577	0.7885
76.3283	1.0116
116.5339	1.2993
154.4343	1.7047
331.7527	2.3979

FIRST PARTIAL TOTAL LIFE : 212.8987

TOTAL REMAINING LIFE : 696.1587

VALUE OF F : 2.7524

P VALUE : 0.9094

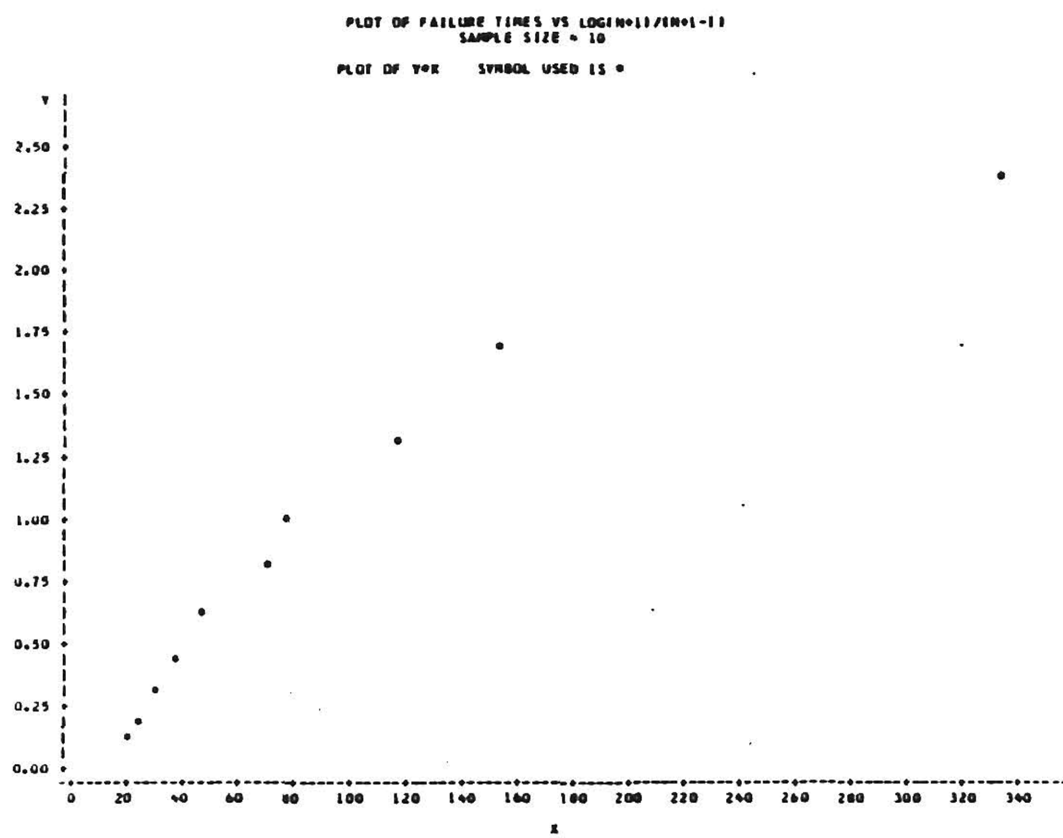


Figure 3.1.1b Plot of dataset 1 (sample size = 10)

Table 3.2.1b Data Set 2 (sample size = 10)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
83.2051	0.0953
101.5925	0.2007
122.8933	0.3185
136.1330	0.4520
184.6098	0.6061
192.5727	0.7885
207.7564	1.0116
225.1102	1.2993
271.3403	1.7047
443.6375	2.3979

FIRST PARTIAL TOTAL LIFE : 832.6514

TOTAL REMAINING LIFE : 1136.2598

VALUE OF F : 6.5952

P VALUE : 0.9929

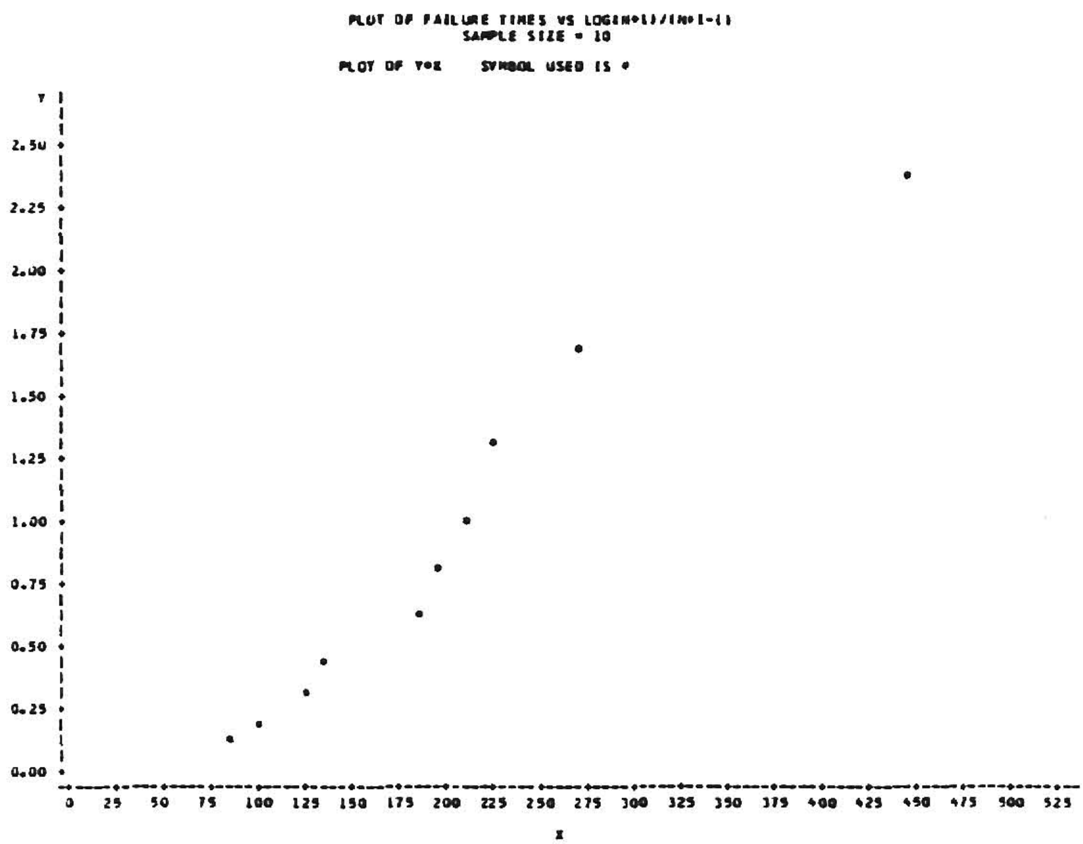


Figure 3.2.1b Plot of dataset 2 (sample size = 10)

Table 3.3.1b Data Set 3 (sample size = 10)

THE SORTED FAILURE TIMES	* LOG(N+1)/(N+1-I)
50.6750	0.0953
50.7218	0.2007
51.9664	0.3185
53.5443	0.4520
74.3169	0.6061
96.9252	0.7885
163.9955	1.0116
188.8099	1.2993
347.7061	1.7047
382.3479	2.3979

FIRST PARTIAL TOTAL LIFE : 506.7505

TOTAL REMAINING LIFE : 954.2581

VALUE OF F : 4.7794

P VALUE : 0.9784

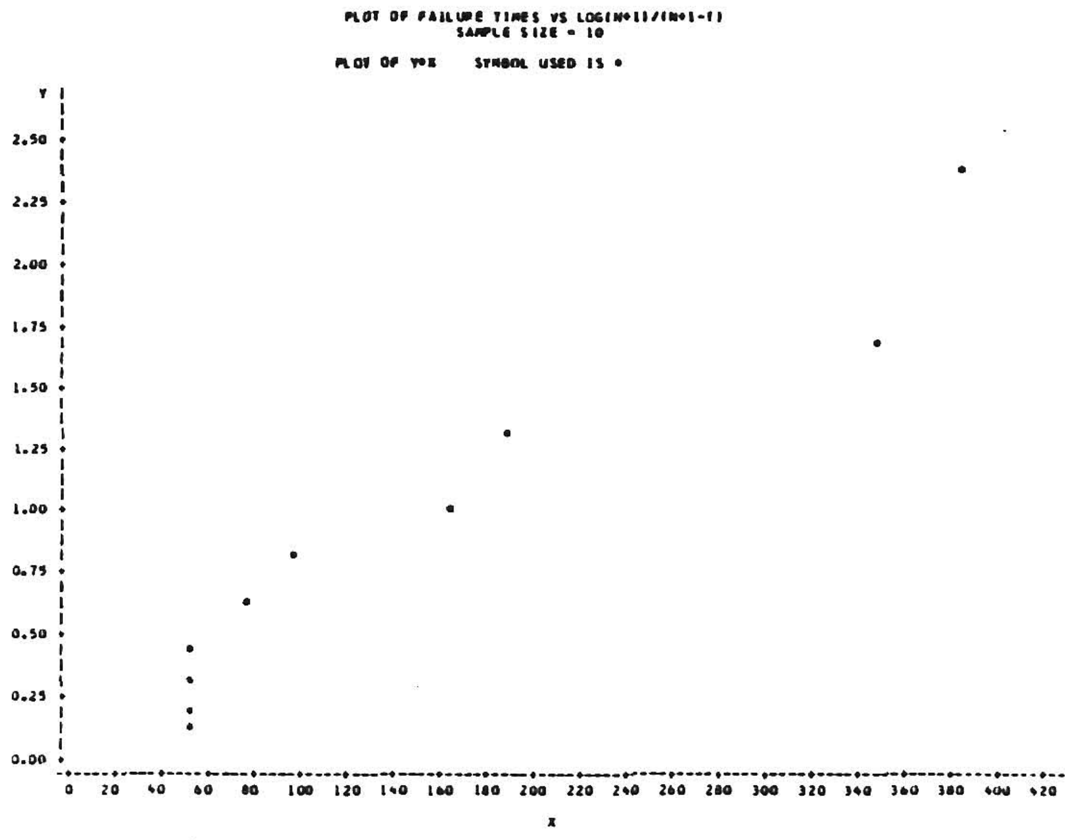


Figure 3.3.1b Plot of dataset 3 (sample size = 10)

Table 3.4.1b Data Set 1 (sample size = 20)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
22.3725	0.0488
24.9013	0.1001
28.5276	0.1542
32.4411	0.2113
32.8290	0.2719
33.3276	0.3365
35.4369	0.4055
36.2734	0.4796
37.4855	0.5596
45.2894	0.6466
47.8629	0.7419
69.2653	0.8473
93.5133	0.9651
111.0238	1.0986
130.7111	1.2528
194.9946	1.4351
212.5572	1.6582
223.2502	1.9459
245.7567	2.3514
260.5879	3.0445

FIRST PARTIAL TOTAL LIFE : 447.4492

TOTAL REMAINING LIFE : 1470.9570

VALUE OF F : 5.7796

P VALUE : 0.9936

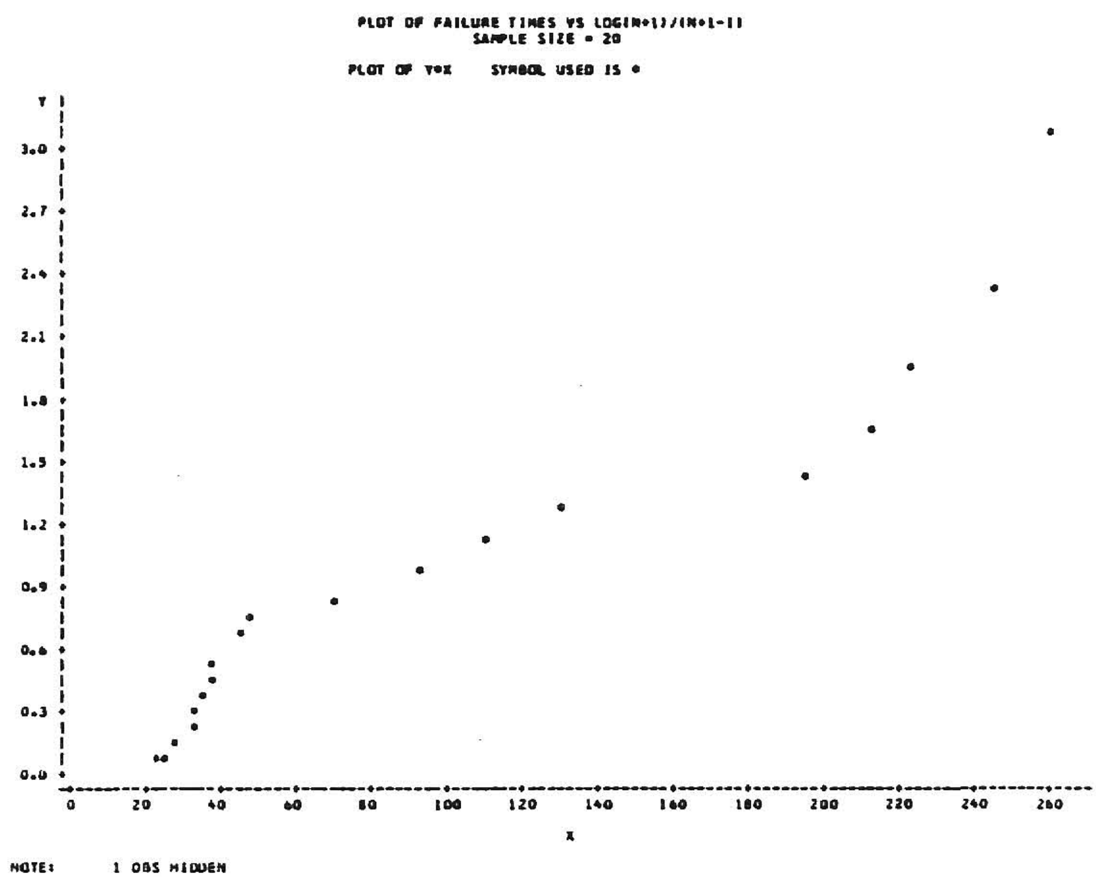


Figure 3.4.1b Plot of dataset 1 (sample size = 20)

Table 3.5. 1b Data Set 2 (sample size = 20)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
32.1432	0.0488
47.3833	0.1001
52.8405	0.1542
67.4231	0.2113
76.8214	0.2719
87.5969	0.3365
91.2964	0.4055
107.7005	0.4796
162.3256	0.5596
175.5964	0.6466
182.1531	0.7419
196.5521	0.8473
198.1229	0.9651
213.2989	1.0986
224.6810	1.2528
272.8022	1.4351
293.9922	1.6582
296.1216	1.9459
359.2297	2.3514
403.0393	3.0445

FIRST PARTIAL TOTAL LIFE : 642.8638

TOTAL REMAINING LIFE : 2898.2549

VALUE OF F : 4.2144

P VALUE : 0.9778

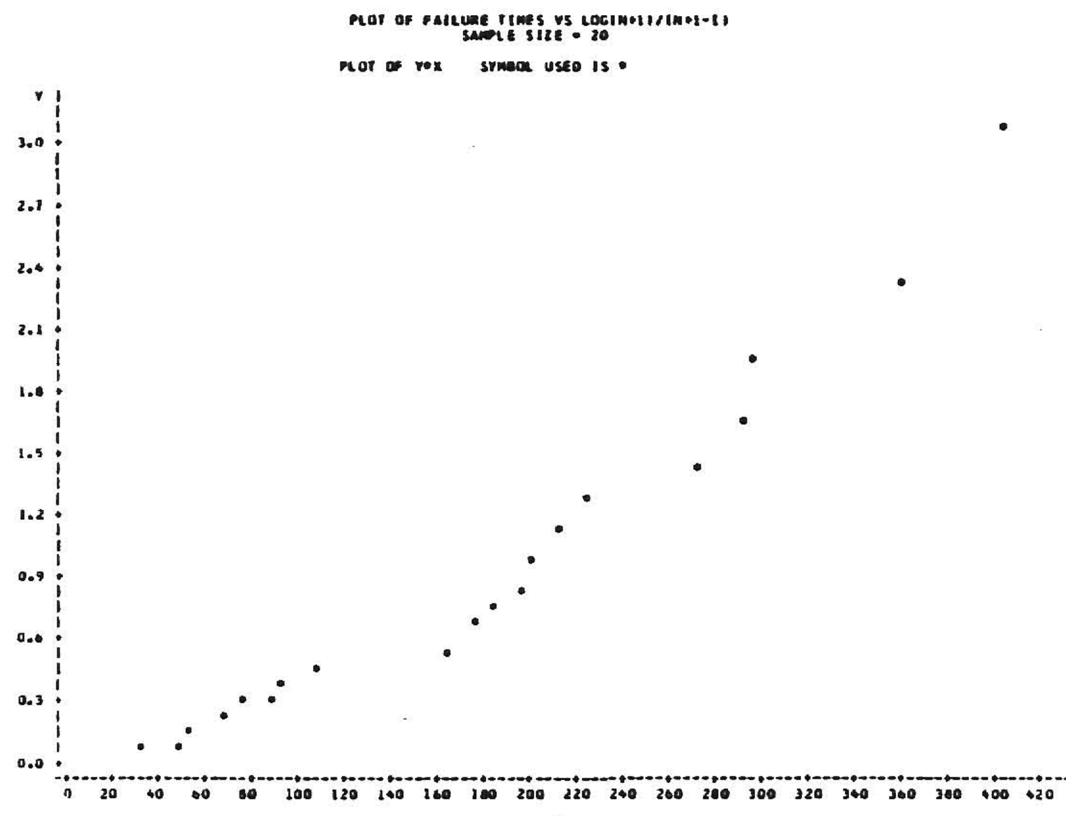


Figure 3.5.1b Plot of dataset 2 (sample size = 20)

Table 3.6.1b Data Set 3 (sample size = 20)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
46.5166	0.0488
55.0757	0.1001
61.4171	0.1542
61.8972	0.2113
63.8281	0.2719
73.2185	0.3365
79.3690	0.4055
90.4448	0.4796
95.4600	0.5596
95.5059	0.6466
101.5988	0.7419
137.0180	0.8473
137.3529	0.9651
138.6029	1.0986
146.2859	1.2528
163.9681	1.4351
176.3790	1.6582
186.3358	1.9459
261.1299	2.3514
267.8091	3.0445

FIRST PARTIAL TOTAL LIFE : 930.3323

TOTAL REMAINING LIFE : 1508.8792

VALUE OF F : 11.7149

P VALUE : 0.9999

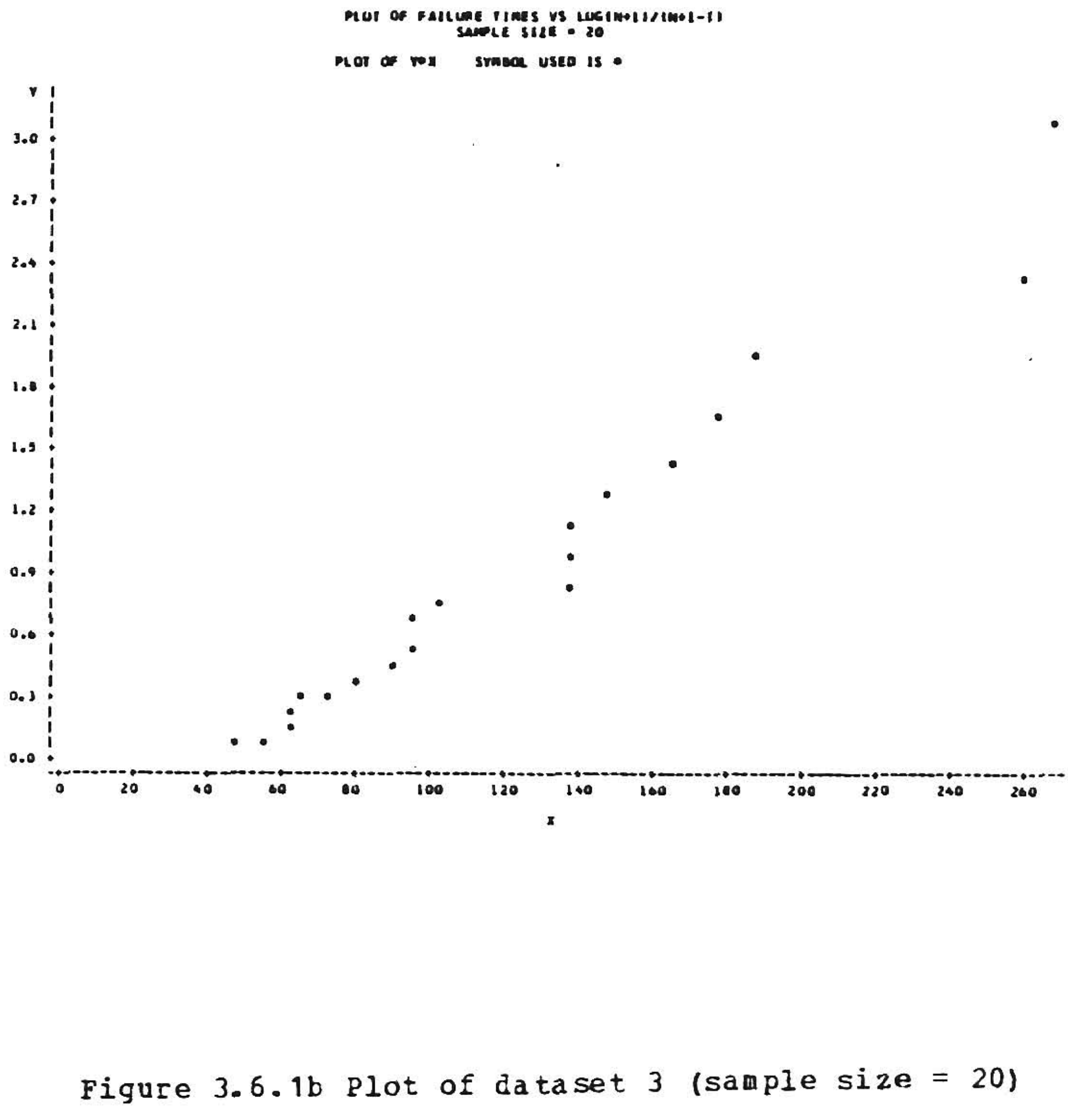


Figure 3.6.1b Plot of dataset 3 (sample size = 20)

Table 3.7.1b Data Set 1 (sample size = 50)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
21.7296	0.0198
22.9096	0.0400
28.0350	0.0606
28.3641	0.0817
29.9639	0.1032
33.3063	0.1252
33.7101	0.1476
33.7683	0.1706
37.3031	0.1942
38.2519	0.2183
40.2145	0.2429
41.1013	0.2683
42.3200	0.2942
43.0069	0.3209
46.2419	0.3483
48.0416	0.3765
50.0085	0.4055
51.9600	0.4353
55.2493	0.4661
57.3970	0.4978
58.2365	0.5306
59.0719	0.5645
60.7533	0.5996
61.4787	0.6360
65.6818	0.6737
67.4823	0.7129
69.9243	0.7538
78.0801	0.7963
80.3974	0.8408
88.7176	0.8873
91.1403	0.9361
91.2204	0.9874
93.4318	1.0415
103.2584	1.0986
105.6698	1.1592
110.7868	1.2238
115.3936	1.2928
117.1772	1.3669
142.2274	1.4469
152.0406	1.5339
179.7837	1.6292
182.3114	1.7346
194.3257	1.8524
195.5833	1.9859
214.8207	2.1401
236.4843	2.3224
259.5554	2.5455
283.2014	2.8332
403.1289	3.2387
430.8633	3.9318

FIRST PARTIAL TOTAL LIFE : 1086.4814

TOTAL REMAINING LIFE : 4088.6250

P VALUE : 1.0000

VALUE OF F : 13.0209

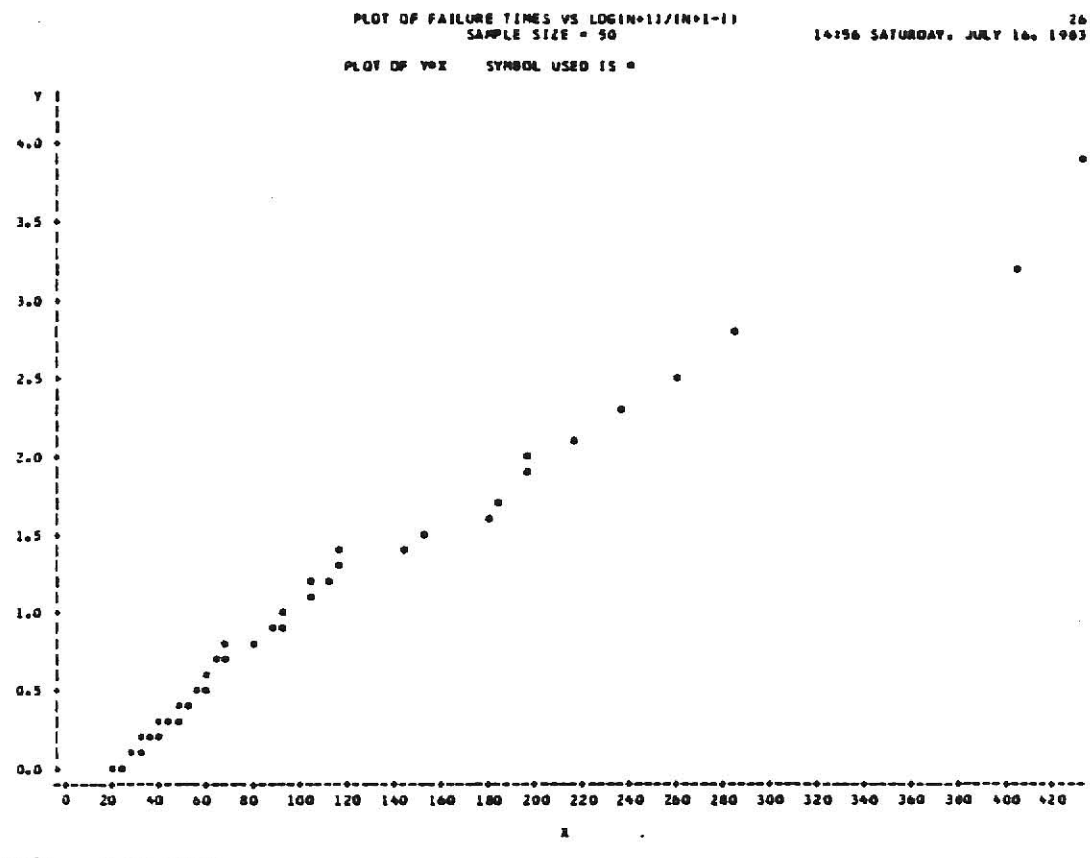


Figure 3.7.1b Plot of dataset 1 (sample size = 50)

Table 3.8.1b Data Set 2 (sample size = 50)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
32.0774	0.0198
33.0719	0.0400
35.6826	0.0606
37.0512	0.0817
43.8292	0.1032
45.7823	0.1252
45.9269	0.1476
56.5153	0.1706
58.7621	0.1942
59.6197	0.2183
60.3581	0.2429
62.0037	0.2683
63.6247	0.2942
63.6607	0.3209
64.7851	0.3483
66.6313	0.3765
72.4106	0.4055
74.2461	0.4353
74.6334	0.4661
81.5862	0.4978
84.6754	0.5306
87.8786	0.5645
90.1051	0.5996
90.9822	0.6360
93.5997	0.6737
96.0506	0.7129
100.1339	0.7538
107.7891	0.7963
108.6090	0.8408
113.2875	0.8873
114.3021	0.9361
127.6920	0.9874
129.3956	1.0415
142.0139	1.0986
144.3672	1.1592
146.2643	1.2238
158.4921	1.2928
178.6637	1.3669
190.6230	1.4469
191.2206	1.5339
191.5503	1.6292
195.9030	1.7346
204.8893	1.8524
250.2849	1.9859
269.8003	2.1401
294.3164	2.3224
318.7368	2.5455
362.0647	2.8332
411.8210	3.2387
423.3215	3.9318

FIRST PARTIAL TOTAL LIFE : 1603.8637

TOTAL REMAINING LIFE : 4947.2070

P VALUE : 1.0000

VALUE OF F : 15.3856

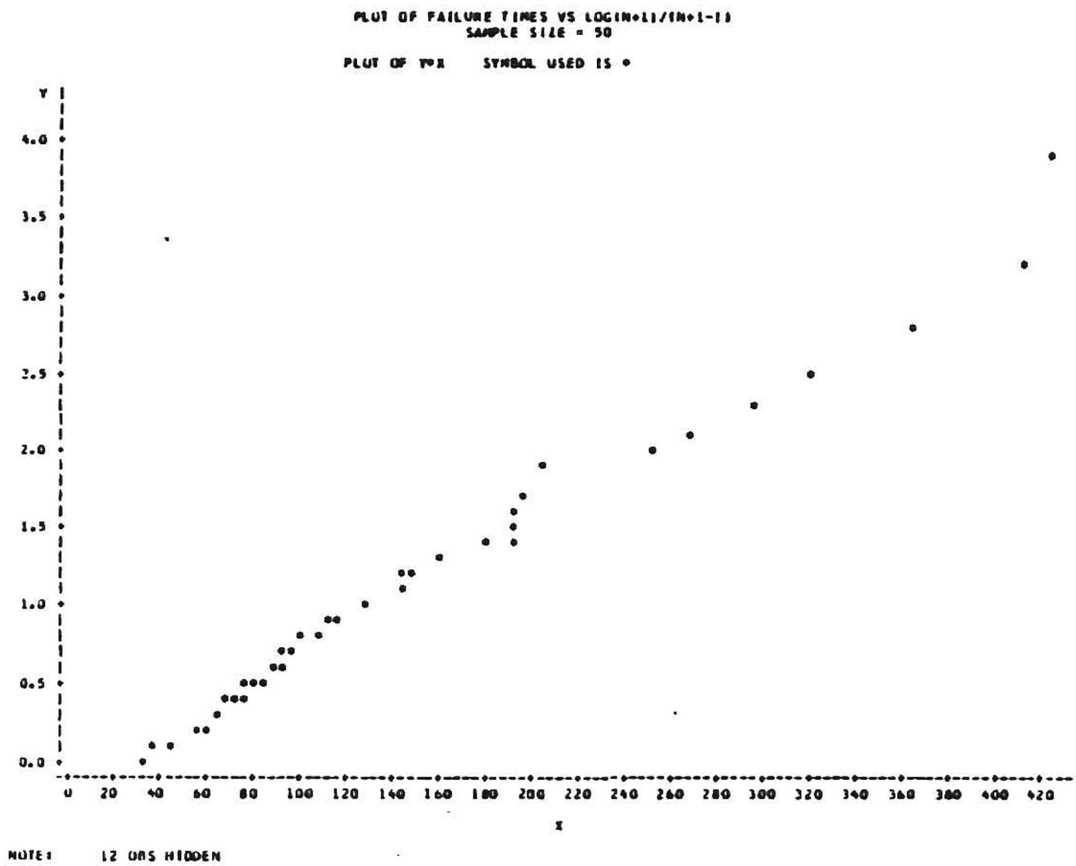


Figure 3.8.1b Plot of dataset 2 (sample size = 50)

Table 3.9.1b Data Set 3 (sample size = 50)

THE SORTED FAILURE TIMES	LOG(N+1)/(N+1-I)
40.4512	0.0198
42.6535	0.0400
45.0784	0.0606
47.8680	0.0817
48.6864	0.1032
49.9966	0.1252
50.8953	0.1476
51.7538	0.1706
56.5774	0.1942
56.7513	0.2183
62.6115	0.2429
64.8396	0.2683
65.3385	0.2942
69.1434	0.3209
70.3432	0.3483
72.8706	0.3765
74.0221	0.4055
74.6021	0.4353
77.0434	0.4661
83.6468	0.4978
84.2706	0.5306
86.9002	0.5645
87.5278	0.5996
91.7741	0.6360
100.0979	0.6737
104.6743	0.7129
104.8414	0.7538
105.6438	0.7963
117.1845	0.8408
118.6195	0.8873
121.1020	0.9361
147.3255	0.9874
154.7951	1.0415
156.8154	1.0986
170.6568	1.1592
195.1810	1.2238
198.4023	1.2928
206.9637	1.3669
226.9966	1.4469
236.0165	1.5339
252.3070	1.6292
252.7328	1.7346
257.5464	1.8524
264.7366	1.9859
292.4844	2.1401
304.9395	2.3224
309.6929	2.5455
339.1934	2.8332
440.3555	3.2387
453.8977	3.9318

FIRST PARTIAL TOTAL LIFE : 2022.5586

TOTAL REMAINING LIFE : 5166.2617

P VALUE : 1.0000

VALUE OF F : 19.1832

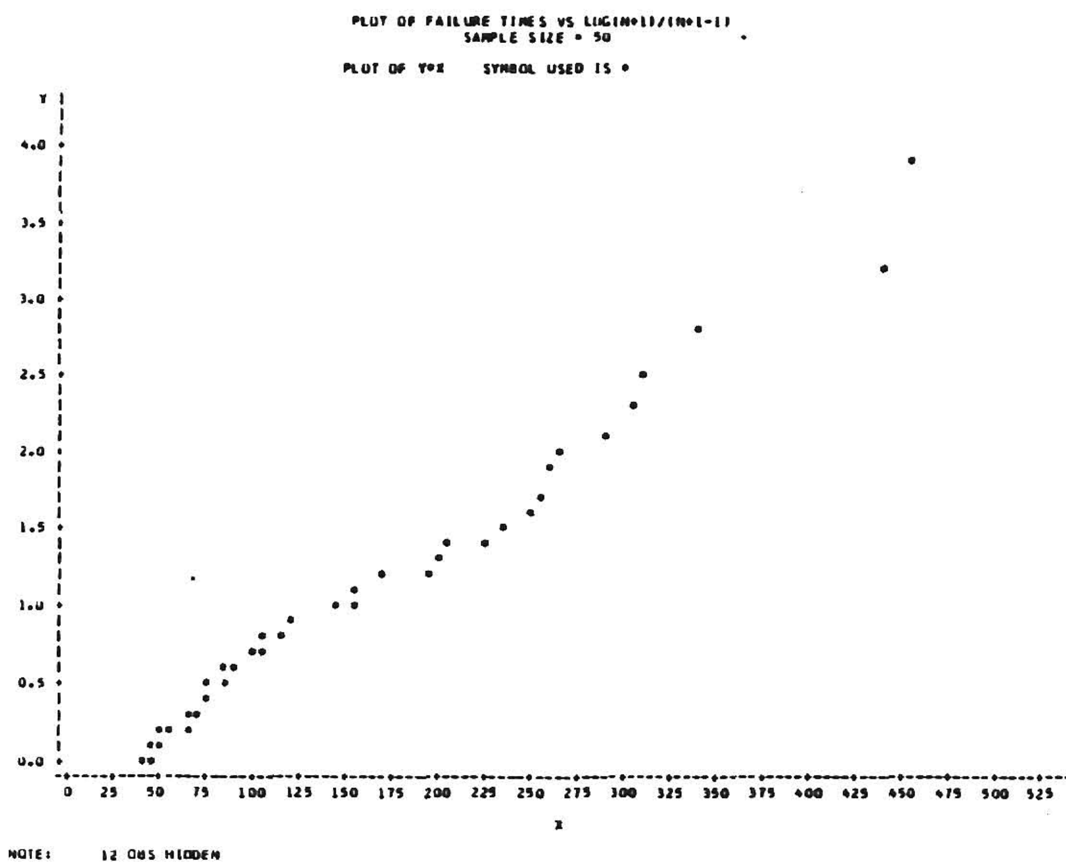


Figure 3.9.1b Plot of dataset 3 (sample size = 50)

SECTION 4.

THE EFFECT OF CONTAMINATION.

It is of frequent interest to study the situation that arises when samples from a particular exponential distribution are doctored, contaminated or seeded by using either superior or inferior values generated from another exponential distribution. The result is a mixed distribution the result of which this section wishes to study by simulation.

Consider the following distributions where exponential populations 1 and 2 are mixed together in the ratio c1 and c2 (c1+c2=1). The resulting population density is

$$f(t) = c_1 f_1(t) + c_2 f_2(t)$$

where $f_1(t) = \frac{1}{\theta_1} e^{-(t/\theta_1)}$ and $f_2(t) = \frac{1}{\theta_2} e^{-(t/\theta_2)}$

then $$F(t) = \int_0^t f(x)dx = C_1 F_1(t) + C_2 F_2(t)$$

$$R(t) = C_1 R_1(t) + C_2 R_2(t)$$

i.e. $$1 - F(t) = C_1 e^{-(t/\theta_1)} + C_2 e^{-(t/\theta_2)}$$

$$= e^{-t/\theta_2}(C_2 + C_1 e^{-t(1/\theta_1 - 1/\theta_2)})$$

$$\frac{1}{1 - F(t)} = \frac{e^{-t/\theta_2}}{C_2 + C_1 e^{-t(1/\theta_1 - 1/\theta_2)}} = \frac{e^{-t/\theta_2}}{D}$$

$$\ln \frac{1}{1 - F(t)} = t/\theta_2 - \ln(D)$$

Where $$D = C_2 + C_1 e^{-t(1/\theta_1 - 1/\theta_2)}$$

Now assume that $\theta_2 >> \theta_1$ and $C_2 >> C_1$. This means that f(1) is the 'dud' population and f(2) is the good population. If $\theta_2 > \theta_1$ then :

$$\frac{1}{\theta_1} - \frac{1}{\theta_2} > 0$$

This means that $e^{-t(1/\theta_1 - 1/\theta_2)}$ is a small positive number(<1), Hence we expect the plot of the pseudo-sample of failure times vs. the value of y obtained from the log function to be displaced by d units along the y axis.

This section is discussed under two cases. Case 1 deals with the actual contamination where the assumption is that a few values of the original superior sample are replaced by members(duds) from another inferior exponential distribution. The contamination level is less than half the original sample, i.e. the factor of contamination is < 0.5.

Case 2 deals with the case of pseudo contamination where members of an inferior sample ($\theta=50$)are sprinkled or seeded with a few members arising out of a superior sample($\theta=100$). This is accomplished by using the same program with G greater than 0.5 instead of less than 0.5. This is what we call pseudo-contamination or reverse contamination.

DISCUSSION.

Several samples of size 10, 20 and 50 were generated. The original sample had a mean of 100 while the contamination or seeding was brought about by overlaying the original values by values from an exponential distribution with mean 50. The level of contamination was varied from 10% to 90% of the original sample and those under 50% and those over 50% are distinguished as the cases mentioned earlier. Appendix 5 is the program that performs the overlay of values creating the mixed sample. The values are passed on to the SAS program outlined in Appendix 5 and the corresponding plots are generated. We consider only two situations under each case. 20% and 40% contamination situations are considered under the first case while 60% and 90% seeding conditions are considered under the second case.

CASE 1.

Table 4.1.1a displays a sample size of 10 contaminated by two values. The F and the p values are determined for comparisons. Figure 4.1.1a plots the failure values against y, the logarithm of the reciprocal reliability function. The plot does indicate a linear trend although a cluster of the last few x values tends to curve the graph upwards in the upper tail. The randomness of the x values gives the graph this unpredictable characteristic. Table 4.2.1a depicts a second sample. Figure 4.2.1a, which plots these values shows a linear fit. As theorized earlier, it was expected that the plots would be shifted by an increment along the y axis. Such an increment is noticed in Figure 4.2.1a with the first few values tending to align themselves a certain increment above the zero point. The effect of the offset is rather clear in Figure 4.3.1a where a sample of 20 with contamination level of 4 is considered. The first few values seem to be setting a linear trend which is offset from the rest of the values which have a different linear trend. Next, samples of 50 are considered where 10 and 20 values were replaced to create the level of contamination. A definite linear trend is noticed in the case where sample size is 50. The values tend to be bunched together giving the notion that they have come from the same distribution. In Figure 4.6.1a it is again fairly evident that the first 20 values that serve as the contamination count are setting a different linear trend from the rest. The fact that these 20 values lie on a line above the rest indicate the possibility of intersecting the y axis at a point well above that which the line connecting values from an uncontaminated sample would have. The theory is again borne out by this example.

CASE 2.

It was also neccessary to consider the situation that may arise when a population containing so called bad items are sprinkled with a few good values. In light of this situation the samples generated by the exponential distribution with $\theta = 50$ was considered the 'defective' sample while the sample with $\theta = 100$ was considered 'good'. In Table 4.7.1b a sample of 10 with contamination level of 6 was considered. Figure 4.7.1b indicates a straightforward linear trend. The same is evident in Figure 4.8.1b. There is also a definite offset with the imaginary line joining the points tending to intersect the y axis with a positive intercept. Examples with sample size 20 and 50 give very similar results (Tables 4.3.1b to 4.6.2b).

Table 4.1.1a Data Set 1 (sample size = 10)

NUMBER OF VALUES OVERLAYED : 2

THE SAMPLE SIZE : 10

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
8.8814	0.0953
14.6716	0.2007
16.2337	0.3185
35.2208	0.4520
40.8486	0.6061
56.9211	0.7885
57.7773	1.0116
61.0787	1.2993
62.7404	1.7047
67.5314	2.3979

FIRST PARTIAL TOTAL LIFE : 88.8144

TOTAL REMAINING LIFE : 333.0903

VALUE OF F : 2.3997

P VALUE : 0.8808

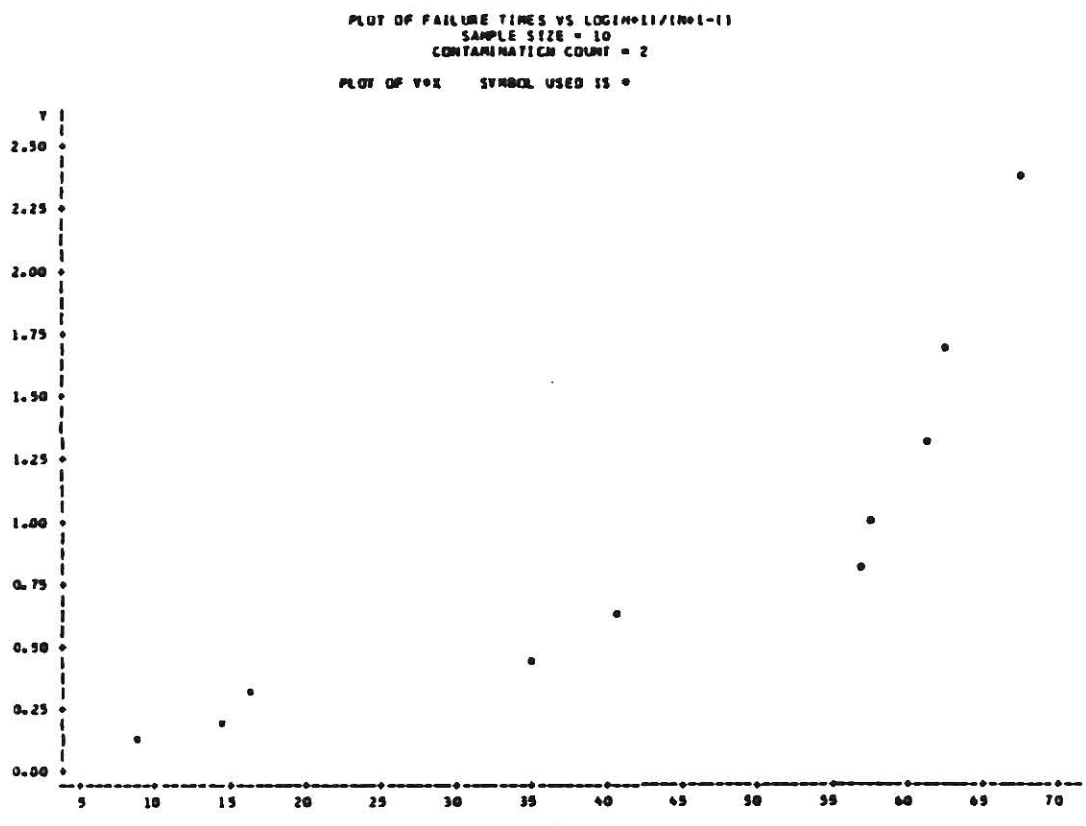


Figure 4.1.1a Plot of dataset 1 (sample size = 10)

Table 4.2.1a Data Set 2 (sample size = 10)

NUMBER OF VALUES OVERLAYED : 4

THE SAMPLE SIZE : 10

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
2.7776	0.0953
6.0569	0.2007
11.5541	0.3185
20.1893	0.4520
32.9350	0.6061
51.9329	0.7885
60.1901	1.0116
96.9115	1.2993
193.9680	1.7047
236.8900	2.3979

FIRST PARTIAL TOTAL LIFE : 27.7759

TOTAL REMAINING LIFE : 685.6289

VALUE OF F : 0.3646

P VALUE : 0.3005

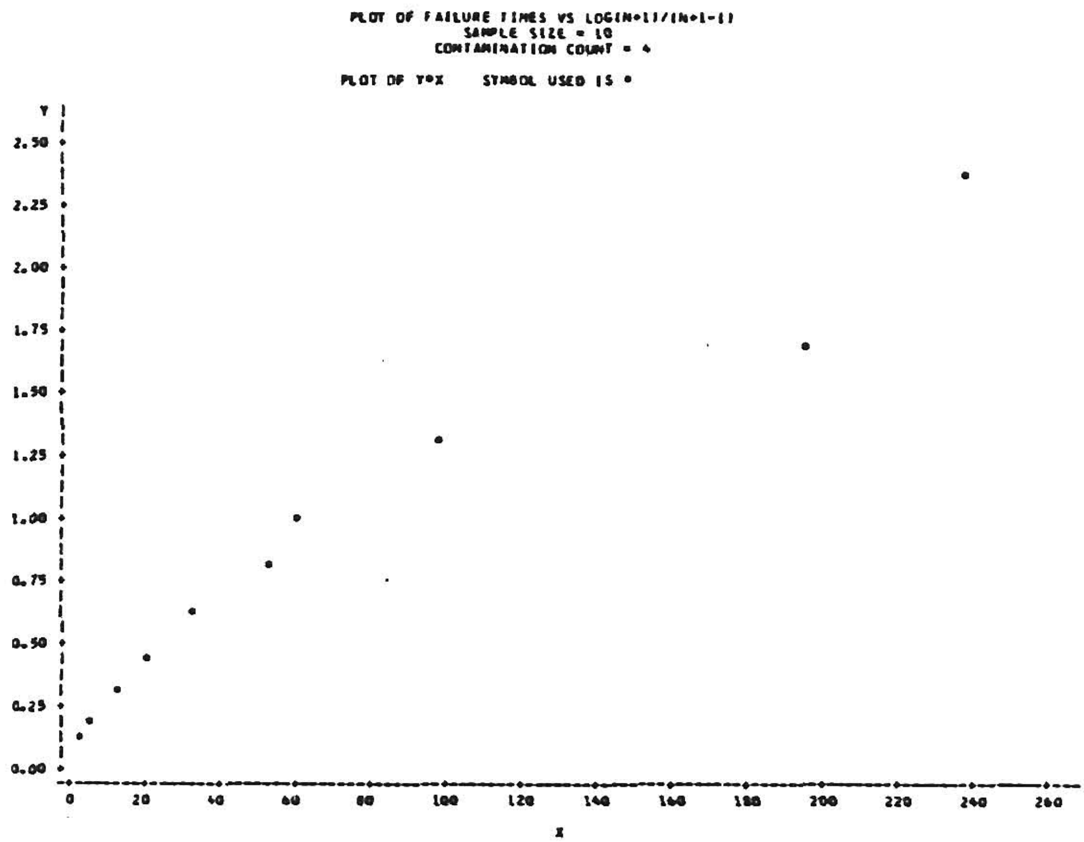


Figure 4.2.1a Plot of dataset 2 (sample size = 10)

Table 4.3.1a Data Set 1 (sample size = 20)

NUMBER OF VALUES OVERLAYED : 4

THE SAMPLE SIZE : 20

THETA1 VALUE : 100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
0.0793	0.0488
0.7289	0.1001
4.8815	0.1542
5.5841	0.2113
11.4941	0.2719
22.1081	0.3365
44.8755	0.4055
51.9010	0.4796
61.6016	0.5596
61.9660	0.6466
63.1937	0.7419
82.1780	0.8473
103.4732	0.9651
113.7945	1.0986
134.0978	1.2528
136.2845	1.4351
162.9979	1.6582
178.0940	1.9459
249.1390	2.3514
287.7517	3.0445

FIRST PARTIAL TOTAL LIFE : 1.5867

TOTAL REMAINING LIFE : 1774.6365

VALUE OF F : 0.0170

P VALUE : 0.0168

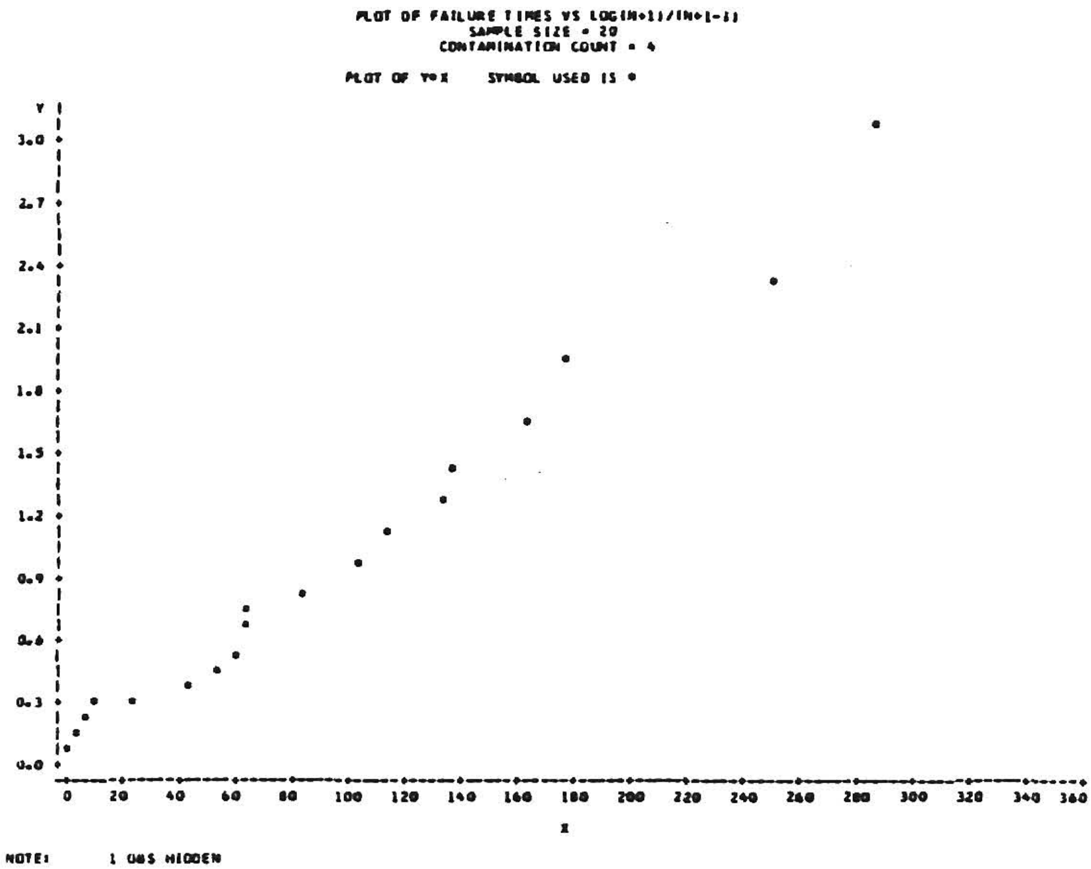


Figure 4.3.1a Plot of dataset 1 (sample size = 20)

Table 4.4.1a Data Set 2 (sample size = 20)

NUMBER OF VALUES OVERLAYED : 8

THE SAMPLE SIZE : 20

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
11.0872	0.0488
16.7595	0.1001
17.0030	0.1542
27.1503	0.2113
32.4364	0.2719
42.0389	0.3365
44.1955	0.4055
46.6443	0.4796
47.6656	0.5596
60.1025	0.6466
72.2987	0.7419
75.9129	0.8473
80.9809	0.9651
93.8709	1.0986
109.3161	1.2528
139.7870	1.4351
144.7603	1.6582
186.3355	1.9459
212.1520	2.3514
313.2422	3.0445

FIRST PARTIAL TOTAL LIFE : 221.7440

TOTAL REMAINING LIFE : 1551.9944

VALUE OF F : 2.7147

P VALUE : 0.9209

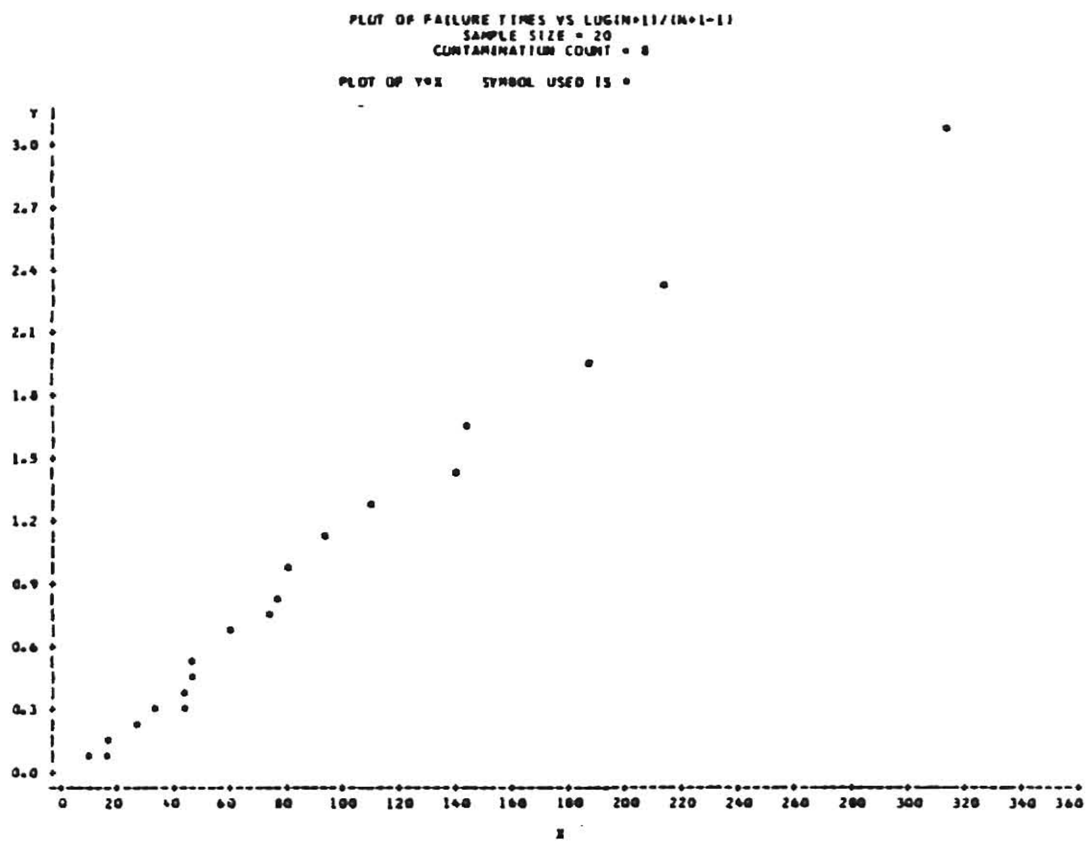


Figure 4.4.1a Plot of dataset 2 (sample size = 20)

Table 4.5.1a Data Set 1 (sample size = 50)

NUMBER OF VALUES OVERLAYED : 10

THE SAMPLE SIZE : 50

THETA1 VALUE : 100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
0.5700	0.0198
2.2135	0.0400
4.0544	0.0606
4.1447	0.0817
7.8109	0.1032
8.3861	0.1252
9.5992	0.1476
10.4842	0.1706
10.5033	0.1942
12.1726	0.2183
15.1661	0.2429
17.1240	0.2683
22.7629	0.2942
25.0957	0.3209
25.2879	0.3483
26.6558	0.3765
36.0455	0.4055
39.3718	0.4353
41.3207	0.4661
43.0425	0.4978
45.0497	0.5306
46.0678	0.5645
60.6718	0.5996
68.1448	0.6360
74.0796	0.6737
74.3605	0.7129
74.6390	0.7538
74.6982	0.7963
84.2120	0.8408
84.5305	0.8873
86.6575	0.9361
94.4023	0.9874
99.5795	1.0415
100.1956	1.0986
102.4823	1.1592
103.1877	1.2238
103.4405	1.2928
110.0394	1.3669
119.5966	1.4469
125.0491	1.5339
137.8167	1.6292
138.6122	1.7346
156.9678	1.8524
167.9348	1.9859
266.4990	2.1401
278.3672	2.3224
280.0330	2.5455
328.8071	2.9332
332.5474	3.2387
516.4559	3.9318

FIRST PARTIAL TOTAL LIFE : 28.4997

TOTAL REMAINING LIFE : 4668.4219

VALUE OF F : 0.2991

P VALUE : 0.2579

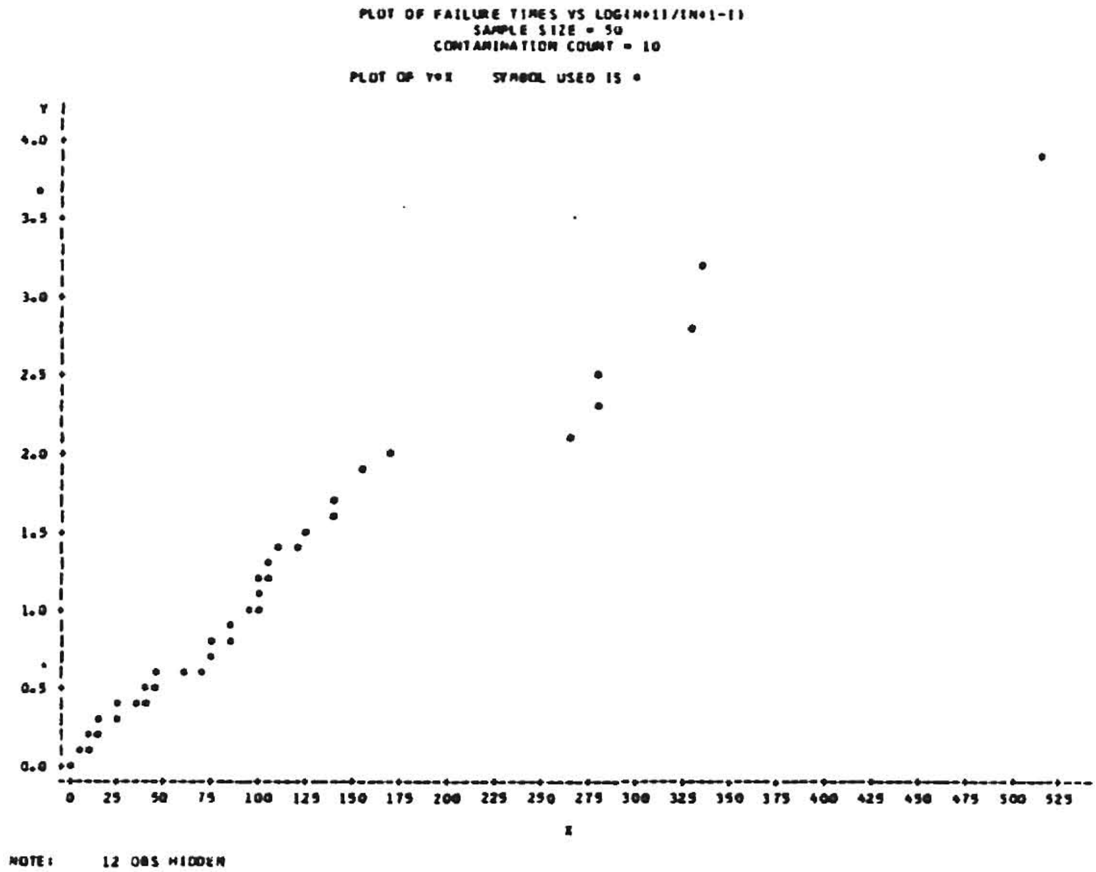


Figure 4.5.1a Plot of dataset 1 (sample size = 50)

Table 4.6.1a Data Set 2 (sample size = 50)

NUMBER OF VALUES OVERLAYED : 20

THE SAMPLE SIZE : 50

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
2.2741	0.0198
3.1103	0.0400
3.4849	0.0606
3.8620	0.0817
4.1245	0.1032
6.4663	0.1252
7.7005	0.1476
8.3114	0.1706
10.1222	0.1942
14.9911	0.2183
16.1337	0.2429
16.3869	0.2683
17.1139	0.2942
17.5618	0.3209
17.6554	0.3483
17.9345	0.3765
25.1785	0.4055
32.5736	0.4353
33.7341	0.4661
34.2237	0.4978
37.5421	0.5306
38.6411	0.5645
42.1867	0.5996
47.3379	0.6360
49.7746	0.6737
52.2986	0.7129
52.9976	0.7538
53.5114	0.7963
56.3198	0.8408
70.3465	0.8873
95.5845	0.9361
98.1009	0.9874
99.0140	1.0415
110.1698	1.0986
114.2343	1.1592
119.0824	1.2238
123.2688	1.2928
132.7077	1.3669
141.0222	1.4469
144.3156	1.5339
148.0611	1.6292
148.4611	1.7346
151.3232	1.8524
155.3198	1.9859
160.4433	2.1401
161.2984	2.3224
218.5370	2.5455
254.3703	2.8332
255.7660	3.2387
353.4915	3.9318

FIRST PARTIAL TOTAL LIFE : 113.7046

TOTAL REMAINING LIFE : 3864.7617

VALUE OF F : 1.4416

P VALUE : 0.7585

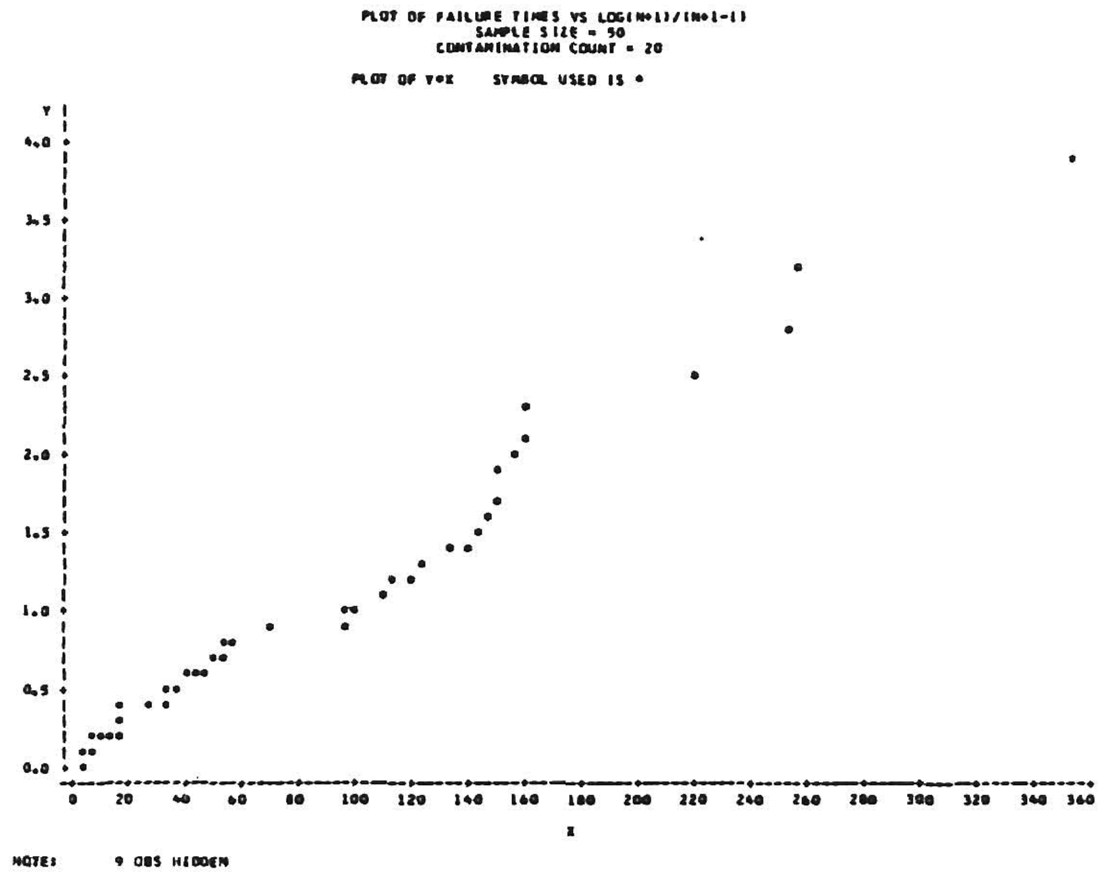


Figure 4.6.1a Plot of dataset 2 (sample size = 50)

Table 4.1.1b Data Set 1 (sample size = 10)

NUMBER OF VALUES OVERLAYED : 6

THE SAMPLE SIZE : 10

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
18.2714	0.0953
18.3261	0.2007
23.6487	0.3185
40.6747	0.4520
53.7338	0.6061
65.0953	0.7885
76.3532	1.0116
107.1635	1.2993
152.1412	1.7047
209.9460	2.3979

FIRST PARTIAL TOTAL LIFE : 182.7136

TOTAL REMAINING LIFE : 582.6399

VALUE OF F : 2.8224

P VALUE : 0.9141

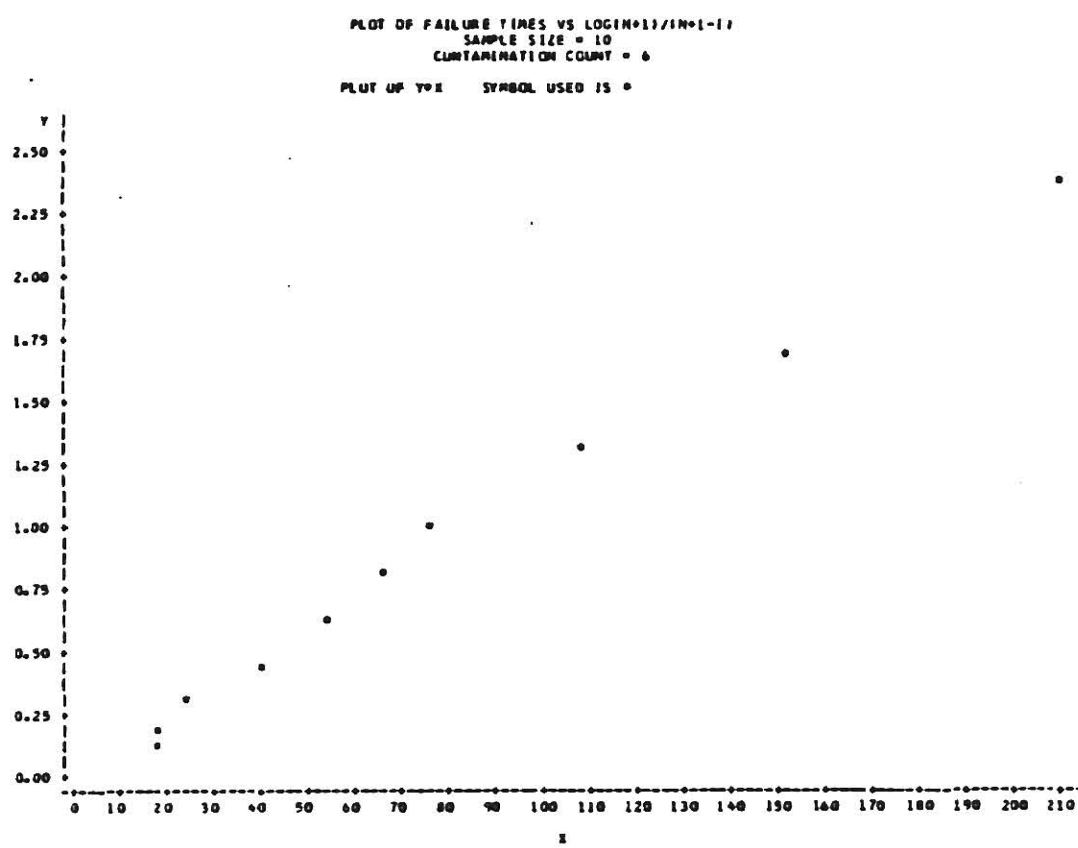


Figure 4.1.1b Plot of dataset 1 (sample size = 10)

Table 4.2.1b Data Set 2 (sample size = 10)

NUMBER OF VALUES OVERLAYED : 9

THE SAMPLE SIZE : 10

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
4.7821	0.0953
7.9898	0.2007
14.1267	0.3185
58.4250	0.4520
66.2398	0.6061
67.7996	0.7885
90.2284	1.0116
166.1073	1.2993
169.4813	1.7047
188.9646	2.3979

FIRST PARTIAL TOTAL LIFE : 47.8206

TOTAL REMAINING LIFE : 786.3228

VALUE OF F : 0.5473

P VALUE : 0.4122

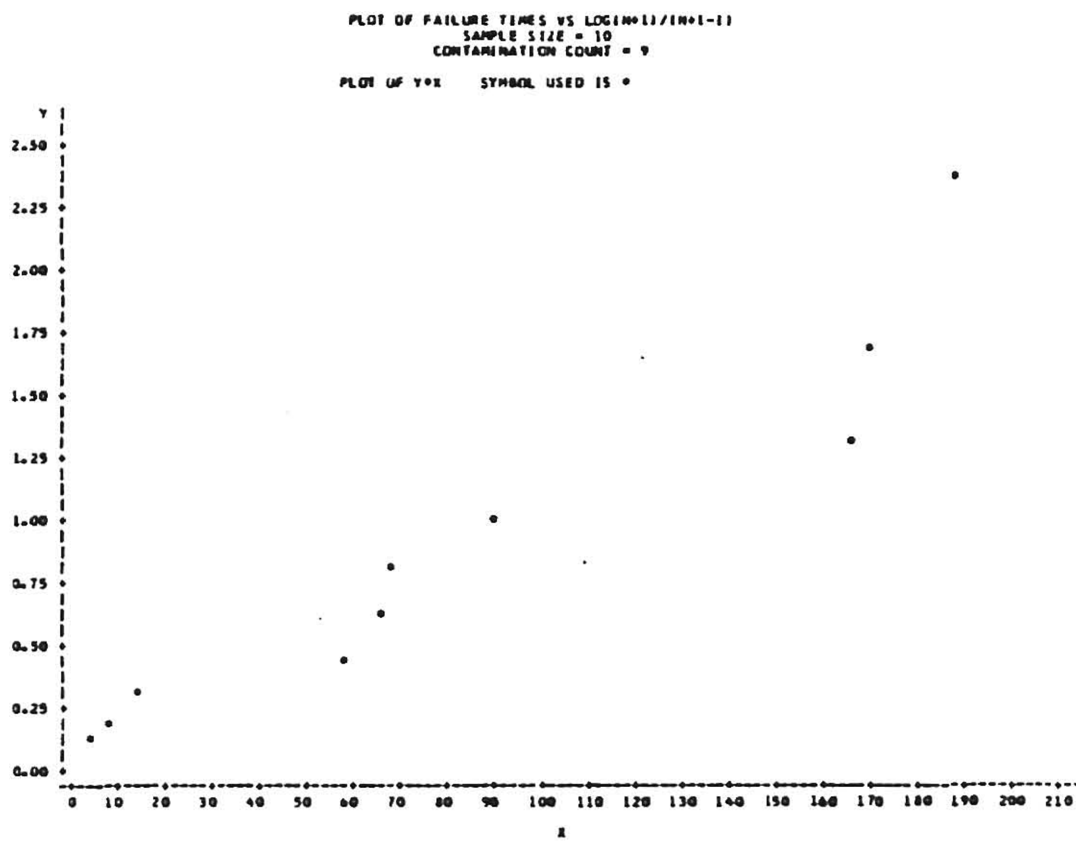


Figure 4.2.1b Plot of dataset 2 (sample size = 10)

Table 4.3.1b Data Set 1 (sample size = 20)

NUMBER OF VALUES OVERLAYED : 12

THE SAMPLE SIZE : 20

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
5.7087	0.0488
6.5010	0.1001
6.5168	0.1542
7.1920	0.2113
25.5154	0.2719
29.2225	0.3365
38.8055	0.4055
38.8721	0.4796
50.7994	0.5596
54.8679	0.6466
56.3284	0.7419
59.7311	0.8473
89.9870	0.9651
92.2048	1.0986
119.8158	1.2528
137.2579	1.4351
150.1005	1.6582
215.3425	1.9459
227.4684	2.3514
289.4424	3.0445

FIRST PARTIAL TOTAL LIFE : 114.1746

TOTAL REMAINING LIFE : 1587.5039

VALUE OF F : 1.3665

P VALUE : 0.7328

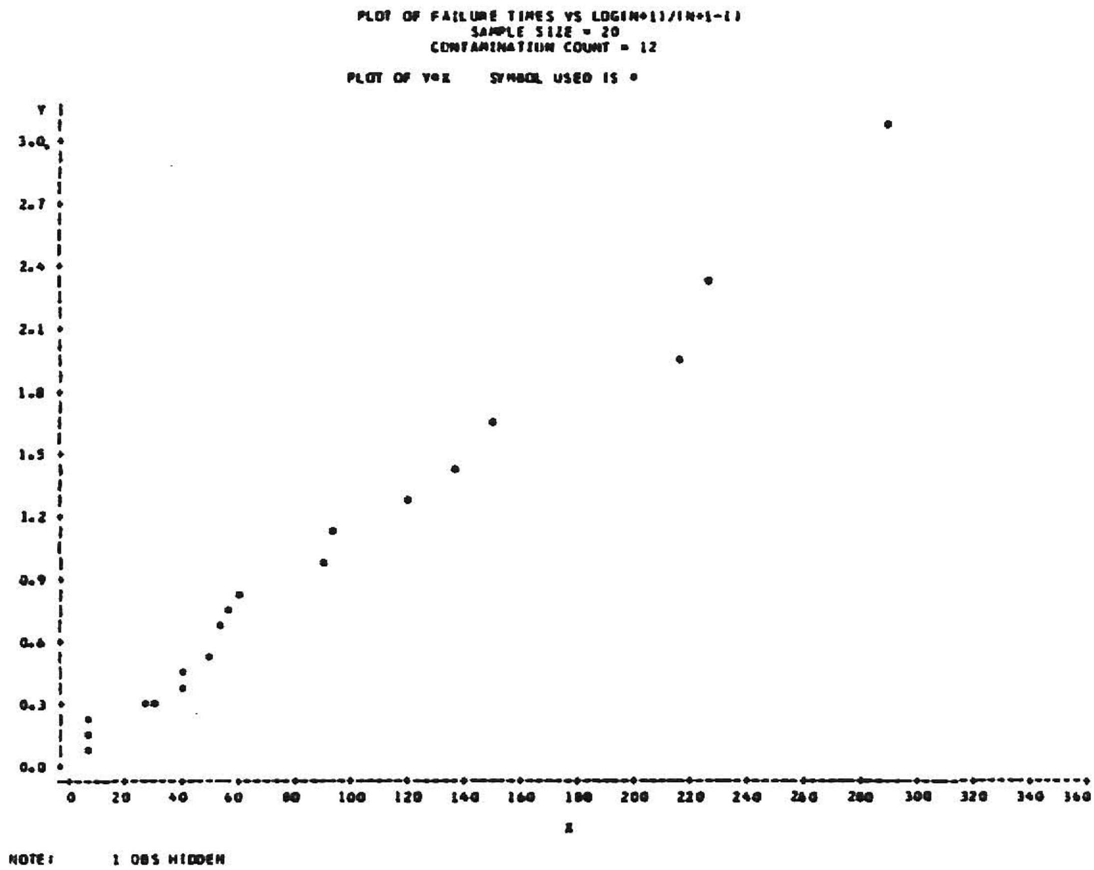


Figure 4.3.1b Plot of dataset 1 (sample size = 20)

Table 4.4.1b Data Set 2 (sample size = 20)

NUMBER OF VALUES OVERLAYED : 18

THE SAMPLE SIZE : 20

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
1.4229	0.0488
2.8006	0.1001
3.6159	0.1542
7.6896	0.2113
11.3403	0.2719
33.4037	0.3365
35.2077	0.4055
36.6419	0.4796
38.3683	0.5596
39.2477	0.6466
40.2394	0.7419
48.3435	0.8473
51.0444	0.9651
57.4014	1.0986
60.6486	1.2528
65.3205	1.4351
74.9447	1.6582
81.9165	1.9459
89.3893	2.3514
99.1853	3.0445

FIRST PARTIAL TOTAL LIFE : 28.4570

TOTAL REMAINING LIFE : 849.7134

VALUE OF F : 0.6363

P VALUE : 0.4652

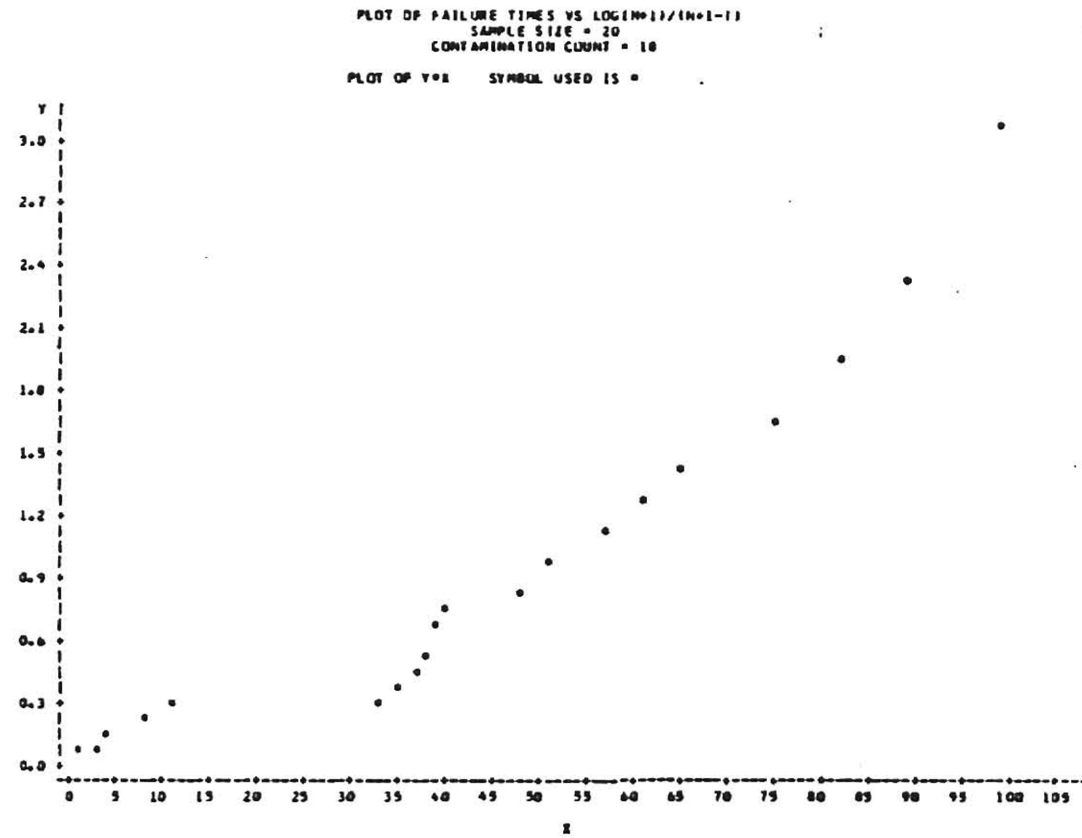


Figure 4.4.1b Plot of dataset 2 (sample size = 20)

Table 4.5.1b Data Set 1 (sample size = 50)

NUMBER OF VALUES OVERLAYED : 30

THE SAMPLE SIZE : 50

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
2.1506	0.0198
2.2929	0.0400
2.8617	0.0606
3.9362	0.0817
5.6889	0.1032
5.8525	0.1252
8.0843	0.1476
8.6660	0.1706
11.9973	0.1942
14.7759	0.2183
16.2506	0.2429
17.7935	0.2683
19.1284	0.2942
19.3577	0.3209
21.7553	0.3483
23.0185	0.3765
23.8176	0.4055
25.3602	0.4353
28.2803	0.4661
32.9239	0.4978
33.5629	0.5306
33.7298	0.5645
34.6773	0.5996
35.6040	0.6360
35.6280	0.6737
38.2151	0.7129
38.7758	0.7538
46.7540	0.7963
46.9218	0.8408
49.2411	0.8873
55.8519	0.9361
70.5379	0.9874
78.3243	1.0415
83.0166	1.0986
88.7932	1.1592
91.4881	1.2238
96.4554	1.2928
97.3703	1.3669
98.4129	1.4469
99.2719	1.5339
106.5970	1.6292
112.3747	1.7346
125.2829	1.8524
126.9490	1.9859
148.0810	2.1401
151.4960	2.3224
171.9731	2.5455
178.0238	2.8332
202.9797	3.2387
203.7360	3.9318

FIRST PARTIAL TOTAL LIFE : 107.5289

TOTAL REMAINING LIFE : 2966.5850

VALUE OF F : 1.7761

P VALUE : 0.8253

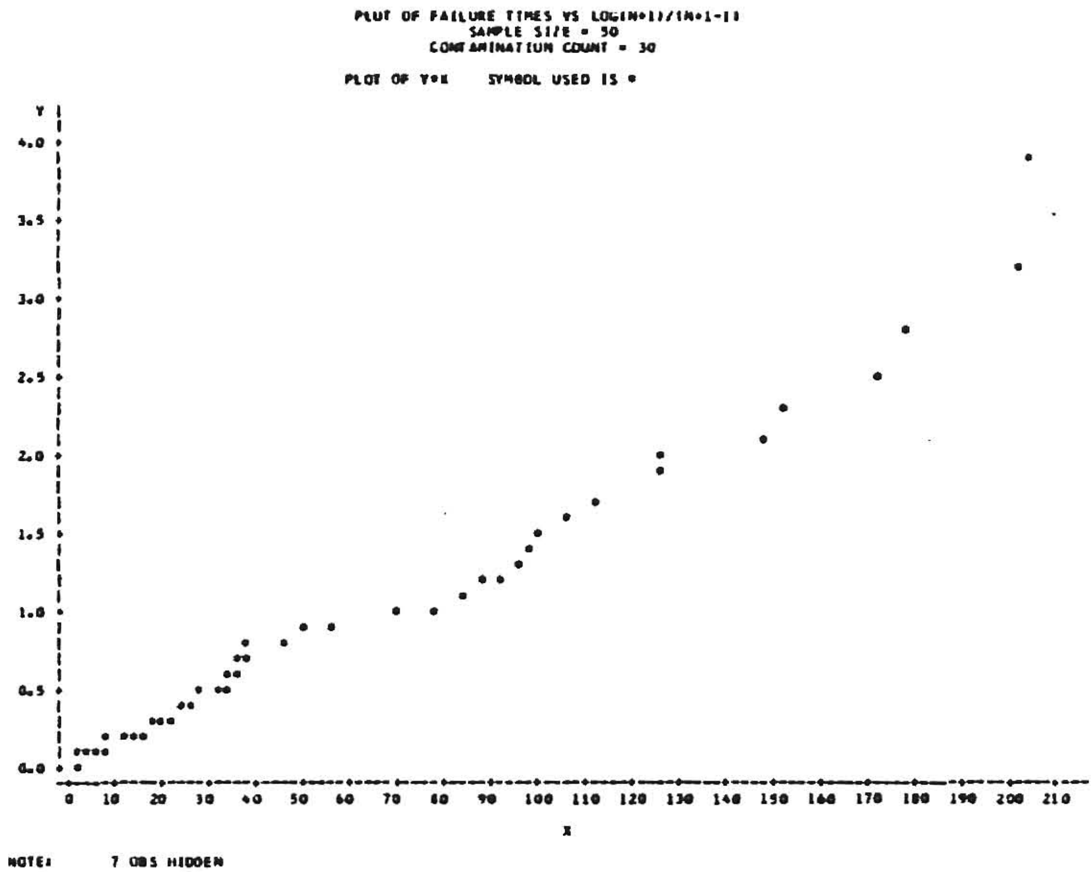


Figure 4.5.1b Plot of dataset 1 (sample size = 50)

Table 4.6.1b Data Set 2 (sample size = 50)

NUMBER OF VALUES OVERLAYED : 45

THE SAMPLE SIZE : 50

THETA1 VALUE :100.00

THETA2 VALUE : 50.00

FAILURE VALUES	Y VALUES
0.7623	0.0198
1.3958	0.0400
1.8252	0.0606
2.0755	0.0817
4.0571	0.1032
5.1049	0.1252
6.6102	0.1476
10.3578	0.1706
13.2920	0.1942
14.7632	0.2183
14.9236	0.2429
16.1201	0.2683
17.4100	0.2942
17.9323	0.3209
20.8343	0.3483
20.8404	0.3765
21.4093	0.4055
21.9363	0.4353
22.7222	0.4661
24.6107	0.4978
26.6132	0.5306
28.1177	0.5645
31.4135	0.5996
36.4061	0.6360
44.4073	0.6737
44.6077	0.7129
45.1189	0.7538
45.2890	0.7963
45.8211	0.8408
46.6563	0.8873
51.0918	0.9361
54.0914	0.9874
55.6717	1.0415
57.1099	1.0986
58.5573	1.1592
60.8899	1.2238
62.6764	1.2928
64.9430	1.3669
91.3608	1.4469
93.5316	1.5339
98.7514	1.6292
100.0353	1.7346
103.0399	1.8524
107.8437	1.9859
109.5312	2.1401
113.2365	2.3224
123.2966	2.5455
158.3794	2.8332
173.9402	3.2387
207.0641	3.9318

FIRST PARTIAL TOTAL LIFE : 38.1158

TOTAL REMAINING LIFE : 2560.3552

VALUE OF F : 0.7295

P VALUE : 0.5152

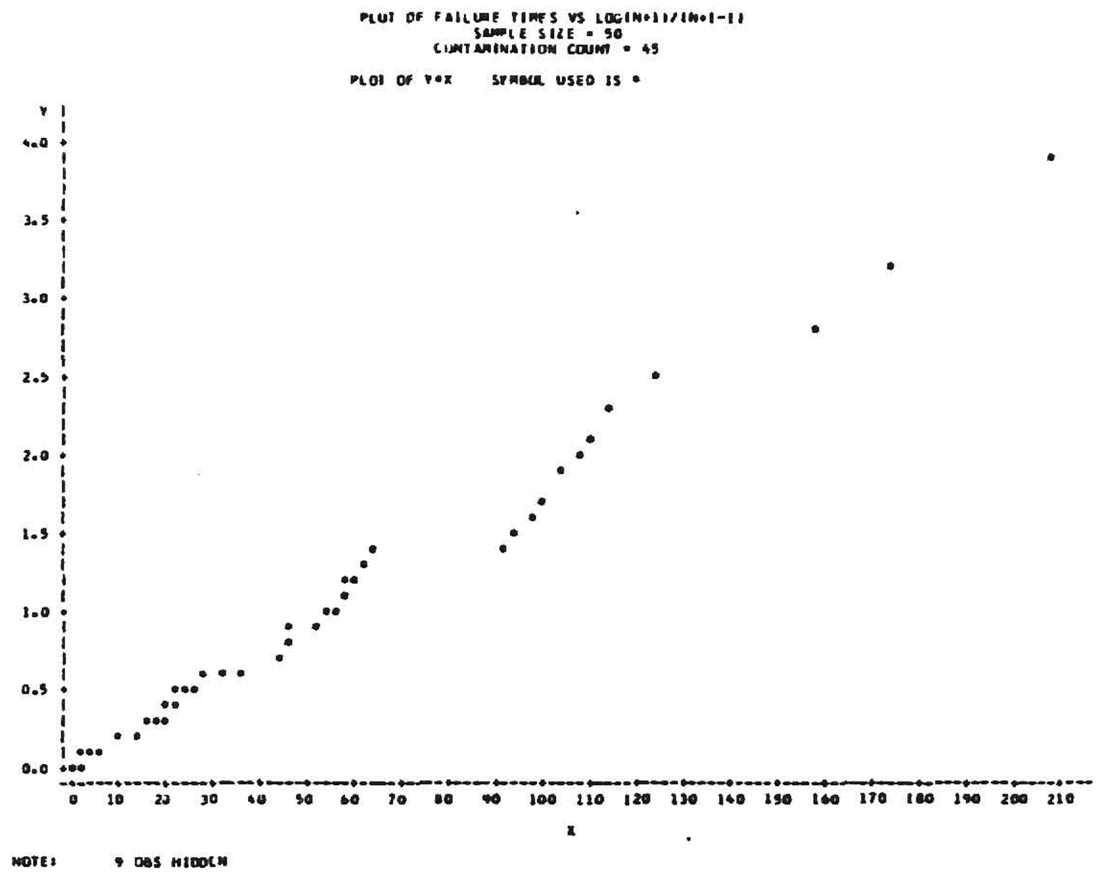


Figure 4.6.1b Plot of dataset 2 (sample size = 50)

CONCLUSIONS.

From the various studies conducted the following conclusions were drawn :

(1) Graphical procedures indicate a strong linear trend in the plot of failure times vs the logarithm of the reciprocal reliability function.

(2) It is fairly evident from the various examples cited that abnormally early or late failures in the case of samples arising from exponential distributions can be recognized by fairly inexperienced observers rather than by having to be subjected to statistical acceptance or rejection by the test proposed by Epstein.

(3) Samples often indicate that the shifted exponential can be recognized by the naked eye and it seems rather evident that rejection by Epstein's test can be brought about only in the case of samples that clearly appear pathological.

Appendix 1

SIMULATION PROGRAM - PLOTTING EXPONENTIAL DEVIATES

This explanation precedes the SAS program that simulates the generation of random exponential deviates. The step DATA RANDOM creates the data set of the same name. The cumulative distribution of an exponential distribution with intensity $\lambda=1$ is given by

$$F(t) = 1 - e^{-t} \qquad t>0$$

$$\therefore \quad 1-F(t) = e^{-t}$$

$$\therefore \quad t = -\log[1 - F(t)]$$

The reason for generating five sets is to obtain several samples in a single run without having to compile the program repeatedly. Line 2 is the DO loop that generates a sample of 10 (5, 10, 20 &50 were also used) exponential deviates using the principle above. The values assumed by the DO loop index also controls the sample size.

PROC SORT OUT=FAILTIM1; BY X; gives the name failtim1 to the data set rearranged by the sort routine. Values are sorted by x(failure times). This is seen in line 16.

DATA PLOTSET;

SET FAILTIM1;

Y=LOG(11/11-N);

Lines 21, 22 & 23 create a dataset named plotset containing variables and values previously created by the OUT=FAILTIM1 statement. The SET statement brings these observations and places them within a new dataset with the name of plotset. The statement Y=LOG(N+1)/(N+1-I) creates another variable Y in the dataset. In line 24 the PROC PRINT statement prints out the data and the VAR statement that follows controls the printout of the variables x and y. Line 25 PROC PLOT; PLOT Y * X ='*'; plots a graph with values along the y-axis and values of x along the x axis.

Simulation Program - Plotting Exponential deviates

```
1        S A S   L O G    OS SAS 82.2B    OS/360 MVT JOB VM104325 STEP SAS      PROC

NOTE: THE JOB VM104325 HAS BEEN RUN UNDER RELEASE 82.2B OF SAS AT KANSAS STATE UNIVERSITY

NOTE: SAS OPTIONS SPECIFIED ARE:
      NOINCLUDE NUGRAPHICS  SORT=4

         DATA RANDOM;
         DO I=1 TO 5;
         F=UNIFORM(81392);
         F1=UNIFORM(24839);
         F2=UNIFORM(67345);
         F3=UNIFORM(41756);
         F4=UNIFORM(46574);
         X=-LOG(1-F);
         X1=-LOG(1-F1);
         X2=-LOG(1-F2);
         X3=-LOG(1-F3);
         X4=-LOG(1-F4);
         OUTPUT;
         END;
         DROP I F F1 F2 F3 F4;

NOTE: DATA SET WORK.RANDOM HAS 5 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.24 SECONDS AND 202K.

         PROC SORT OUT=FAILTIM1;BY X;

NOTE: DATA SET WORK.FAILTIM1 HAS 5 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: THE PROCEDURE SORT USED 0.74 SECONDS AND 264K.

         PROC SORT OUT=FAILTIM2;BY X1;

NOTE: DATA SET WORK.FAILTIM2 HAS 5 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: THE PROCEDURE SORT USED 0.68 SECONDS AND 264K.

         PROC SORT OUT=FAILTIM3;BY X2;

NOTE: DATA SET WORK.FAILTIM3 HAS 5 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: THE PROCEDURE SORT USED 0.72 SECONDS AND 264K.

         PROC SORT OUT=FAILTIM4;BY X3;

NOTE: DATA SET WORK.FAILTIM4 HAS 5 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: THE PROCEDURE SORT USED 0.78 SECONDS AND 264K.

         PROC SORT OUT=FAILTIM5;BY X4;

NOTE: DATA SET WORK.FAILTIM5 HAS 5 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: THE PROCEDURE SORT USED 0.73 SECONDS AND 264K.

         DATA PLOTSET;
         SET FAILTIM1;
         Y=LOG(6/(6-_N_));

NOTE: DATA SET WORK.PLOTSET HAS 5 OBSERVATIONS AND 6 VARIABLES. 366 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.15 SECONDS AND 202K.

         PROC PRINT ;VAR X Y;
```

```
2        S A S   L O G    OS SAS 82.2B    OS/360 MVT JOB VM104325 STEP SAS       PROC

NOTE: THE PROCEDURE PRINT USED 0.23 SECONDS AND 202K AND PRINTED PAGE 1.

25          PROC PLOT;PLOT Y*X= '*'; TITLE LOG((N+1)/(N+1-I)) VS FAILURE TIMES;

NOTE: THE PROCEDURE PLOT USED 0.30 SECONDS AND 208K AND PRINTED PAGE 2.

26          DATA PLOTSET2;
27          SET FAILTIM2;
28          Y=LOG(6/(6-_N_));

NOTE: DATA SET WORK.PLOTSET2 HAS 5 OBSERVATIONS AND 6 VARIABLES. 366 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.14 SECONDS AND 202K.

29          PROC PRINT ; VAR X1 Y;

NOTE: THE PROCEDURE PRINT USED 0.24 SECONDS AND 202K AND PRINTED PAGE 3.

30          PROC PLOT;PLOT Y*X1= '*'; TITLE LOG((N+1)/(N+1-I)) VS FAILURE TIMES;

NOTE: THE PROCEDURE PLOT USED 0.34 SECONDS AND 208K AND PRINTED PAGE 4.

31          DATA PLOTSET3;
32          SET FAILTIM3;
33          Y=LOG(6/(6-_N_));

NOTE: DATA SET WORK.PLOTSET3 HAS 5 OBSERVATIONS AND 6 VARIABLES. 366 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.12 SECONDS AND 202K.

34          PROC PRINT;VAR X2 Y;

NOTE: THE PROCEDURE PRINT USED 0.24 SECONDS AND 202K AND PRINTED PAGE 5.

35          PROC PLOT;PLOT Y*X2= '*'; TITLE LOG((N+1)/(N+1-I)) VS FAILURE TIMES;

NOTE: THE PROCEDURE PLOT USED 0.32 SECONDS AND 208K AND PRINTED PAGE 6.

36          DATA PLOTSET4;
37          SET FAILTIM4;
38          Y=LOG(6/(6-_N_));

NOTE: DATA SET WORK.PLOTSET4 HAS 5 OBSERVATIONS AND 6 VARIABLES. 366 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.13 SECONDS AND 202K.

39          PROC PRINT; VAR X3 Y;

NOTE: THE PROCEDURE PRINT USED 0.23 SECONDS AND 202K AND PRINTED PAGE 7.

40          PROC PLOT; PLOT Y*X3= '*'; TITLE LOG((N+1)/(N+1-I)) VS FAILURE TIMES;

NOTE: THE PROCEDURE PLOT USED 0.32 SECONDS AND 208K AND PRINTED PAGE 8.

41          DATA PLOTSET5;
42          SET FAILTIM5;
43          Y=LOG(6/(6-_N_));

NOTE: DATA SET WORK.PLOTSET5 HAS 5 OBSERVATIONS AND 6 VARIABLES. 366 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.13 SECONDS AND 202K.
```

```
S A S   L O G    OS SAS 82.2B    OS/360 MVT JOB VM104325 STEP SAS      PROC

4          PROC PRINT; VAR X4 Y;

NOTE: THE PROCEDURE PRINT USED 0.21 SECONDS AND 202K AND PRINTED PAGE 9.

5          PROC PLOT; PLOT Y*X4= '*'; TITLE LOG((N+1)/(N+1-I)) VS FAILURE TIMES;

NOTE: THE PROCEDURE PLOT USED 0.31 SECONDS AND 208K AND PRINTED PAGE 10.
NOTE: SAS USED 264K MEMORY.

NOTE: SAS INSTITUTE INC.
      SAS CIRCLE
      PO BOX 8000
      CARY, N.C. 27511-8000
```

Appendix 2

SIMULATION PROGRAM

TEST FOR ABNORMALLY EARLY OR LATE FAILURES

This explanation precedes the FORTRAN/SAS program that simulates the test for abnormally early or late failures. Line 1 (denoted 0001 in the program) and 2 make the appropriate declarations for the variables used in the program. In addition line 2 also declares a few arrays and they are appropriately dimensioned. Line 5 begins a DO loop the index values of which controls the number of samples of a particular size generated. The value of XM in line 6 corresponds to the mean of the exponential distribution from which the deviates are generated.The value of NR in line 6 indicates the sample size.

Line 13 commences a DO loop which generates the y values. Line 18 within the loop writes the values of the failure times and y values to disk to enable it to be passed into a SAS program which plots these values later. Line 23 commences a DO loop that accumulates the (NR-1) failure times suitably weighted and writes them into TOTBAL which gives the total life for the periods remaining after the first failure has occured. Line 31 calculates the F value which is to be compared to an F* value with 2 and 2r-2 degrees of freedom.

Line 1 of the SAS LOG creates a data set by the name of COMPLEX. Line 2 indicates that a file stored on disk has to be read in from the appropriate columns. Lines 4 and 5 create 10 different data set samples each of size 10 having the names sample1, sample2, sample3 etc.. Line 17 PROC PRINT DATA = SAMPLE1; prints only the data contained in data set sample1. The PROC PLOT feature in line 27 plots the values of x vs y using values from sample1.

Simulation Program - Test for abnormally early or late failures

```
FORTRAN IV G LEVEL  21                MAIN              DATE = 83196         00/53/21

      C
      C     THIS PROGRAM TESTS FOR ABNORMALLY EARLY OR LATE FIRST FAILURE.
      C
            INTEGER NR,NL,N1,N2,IER
            REAL TBAL(50),R(50),XM,TOTBAL,F,P,Y(50)
            DOUBLE PRECISION DSEED
            DSEED=123457.D00
            DO 100 K=1,10
            XM=100.
            NR=10
            CALL GGEXN (DSEED,XM,NR,R)
      C
      C     GGEXN IS AN IMSL SUBROUTINE THAT GENERATES RANDOM EXPONENTIAL
      C     DEVIATES. A CERTAIN SEED IS SPECIFIED AND 'NR' SPECIFIES THE
      C     NUMBER OF DEVIATES GENERATED. 'XR' IS THE VALUE OF THETA, THE
      C     MEAN LIFE IN THE EXPONENTIAL DISTRIBUTION. 'R' IS THE VECTOR
      C     OF VALUES GENERATED CORRESPONDING TO 'NR'.
      C
            CALL VSRTA (R,NR)
      C
      C     VSRTA IS AN IMSL SUBROUTINE THAT SORTS THE DATA.'R' CORRESPONDS
      C     TO THE INPUT VECTOR OF VALUES TO BE SORTED WHILE 'NR' CORRESPONDS
      C     TO THE NUMBER OF VALUES GENERATED.
      C
            WRITE(6,6)
      6     FORMAT('1',10X,'THE SORTED FAILURE TIMES',10X,'LOG(N+1)/(N+1-I)'
           $//)
            X=1
            DO 95 I=1,NR
            Y(I)=ALOG((NR+1)/(NR+1-X))
            X=X+1
            WRITE(6,7)R(I),Y(I)
       7    FORMAT(18X,F10.4,20X,F10.4)
            WRITE(10,8)R(I),Y(I),XM
       8    FORMAT(5X,F10.4,5X,F10.4,5X,F10.4)
      95    CONTINUE
            TBAL(1)=R(1)*NR
      C
      C     TBAL(1) GIVES THE TOTAL LIFE TILL THE FIRST FAILURE.
      C
            TOTBAL=0.
            DO 10 I=2,NR
            TBAL(I)=(R(I)-R(I-1))*(NR+1-I)
```

```
FORTRAN IV G LEVEL  21                MAIN                DATE = 83196          00/53/21

              C
              C     TBAL(I) GIVES THE TOTAL LIFE FOR INDIVIDUAL FAILURES.
              C
 0025               TOTBAL=TOTBAL + TBAL(I)
              C
              C     TOTBAL GIVES THE TOTAL LIFE FOR THE PERIODS REMAINING AFTEP
              C     THE FIRST FAILURE HAS OCCURED.
              C
 0026         10    CONTINUE
 0027               WRITE(6,11) TBAL(1)
 0028         11    FORMAT ('0',10X,'FIRST PARTIAL TOTAL LIFE :',F10.4)
 0029               WRITE(6,12) TOTBAL
 0030         12    FORMAT('0',10X,'TOTAL REMAINING LIFE :',F10.4)
 0031               F=TBAL(1)/TOTBAL*(NR-1)
 0032               N1=2
 0033               N2=2*(NR-1)
 0034               CALL MDFD (F,N1,N2,P,IER)
              C
              C     MDFD IS AN IMSL SUBROUTINE THAT GIVES THE RIGHT HAND AREA
              C     'P' CORRESPONDING TO A GIVEN COMPUTED F WITH N1 AND N2, THE
              C     NUMERATOR AND DENOMINATOR DEGREES OF FREEDOM RESPECTIVELY.
              C     'IER' IS THE ERROR VALUE
              C
 0035               WRITE(6,13) F
 0036         13    FORMAT('0',10X,'VALUE OF F :',F10.4)
 0037               WRITE(6,14) P,IER
 0038         14    FORMAT('0',10X,'P VALUE :',F10.4//11X,'ERROR VALUE :',I5)
 0039         100   CONTINUE
 0040               STOP
 0041               END
```

```
1        S A S   L O G    OS SAS 82.2B    OS/360 MVT JOB VM005314 STEP SAS        PROC

NOTE: THE JOB VM005314 HAS BEEN RUN UNDER RELEASE 82.2B OF SAS
      AT KANSAS STATE UNIVERSITY (03010001).

NOTE: SAS OPTIONS SPECIFIED ARE:
      NOINCLUDE NOGRAPHICS LINESIZE=115,PAGESIZE=55 SORT=4

         DATA COMPLEX;
         INFILE DATASET;
         INPUT X 6-15 Y 21-30 THETA 36-45;


NOTE: INFILE DATASET IS:
      DSNAME=SYS83196.T004849.RV000.VM005314.TEMP,
      UNIT=SYSDA,VOL=SER=KSCC5A,DISP=OLD,
      DCB=(BLKSIZE=6160,LRECL=80,RECFM=FB)

NOTE: 100 LINES WERE READ FROM INFILE DATASET.
NOTE: DATA SET WORK.COMPLEX HAS 100 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.28 SECONDS AND 202K.

         DATA SAMPLE1 SAMPLE2 SAMPLE3 SAMPLE4 SAMPLE5 SAMPLE6 SAMPLE7
         SAMPLE8 SAMPLE9 SAMPLE10;
         SET COMPLEX;
         IF             _N_<=10 THEN OUTPUT SAMPLE1;
         IF _N_ >10 AND _N_ <=20 THEN OUTPUT SAMPLE2;
         IF _N_ >20 AND _N_ <=30 THEN OUTPUT SAMPLE3;
         IF _N_ >30 AND _N_ <=40 THEN OUTPUT SAMPLE4;
         IF _N_ >40 AND _N_ <=50 THEN OUTPUT SAMPLE5;
         IF _N_ >50 AND _N_ <=60 THEN OUTPUT SAMPLE6;
         IF _N_ >60 AND _N_ <=70 THEN OUTPUT SAMPLE7;
         IF _N_ >70 AND _N_ <=80 THEN OUTPUT SAMPLE8;
         IF _N_ >80 AND _N_ <=90 THEN OUTPUT SAMPLE9 ;
         IF _N_ >90 AND _N_ <=100 THEN OUTPUT SAMPLE10;

NOTE: DATA SET WORK.SAMPLE1 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE2 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE3 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE4 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE5 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE6 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE7 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE8 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE9 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE10 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.42 SECONDS AND 324K.

         PROC PRINT DATA = SAMPLE1;TITLE SAMPLE OF SIZE 10;

NOTE: THE PROCEDURE PRINT USED 0.22 SECONDS AND 202K AND PRINTED PAGE 1.

         PROC PRINT DATA = SAMPLE2;

NOTE: THE PROCEDURE PRINT USED 0.25 SECONDS AND 202K AND PRINTED PAGE 2.
```

```
2      S A S   L O G    OS SAS 82.2B    OS/360 MVT JOB VM005314 STEP SAS      PROC

19          PROC PRINT DATA = SAMPLE3;

NOTE: THE PROCEDURE PRINT USED 0.23 SECONDS AND 202K AND PRINTED PAGE 3.

20          PROC PRINT DATA = SAMPLE4;

NOTE: THE PROCEDURE PRINT USED 0.23 SECONDS AND 202K AND PRINTED PAGE 4.

21          PROC PRINT DATA = SAMPLE5;

NOTE: THE PROCEDURE PRINT USED 0.22 SECONDS AND 202K AND PRINTED PAGE 5.

22          PROC PRINT DATA = SAMPLE6;

NOTE: THE PROCEDURE PRINT USED 0.23 SECONDS AND 202K AND PRINTED PAGE 6.

23          PROC PRINT DATA = SAMPLE7;

NOTE: THE PROCEDURE PRINT USED 0.24 SECONDS AND 202K AND PRINTED PAGE 7.

24          PROC PRINT DATA = SAMPLE8;

NOTE: THE PROCEDURE PRINT USED 0.25 SECONDS AND 202K AND PRINTED PAGE 8.

25          PROC PRINT DATA = SAMPLE9;

NOTE: THE PROCEDURE PRINT USED 0.21 SECONDS AND 202K AND PRINTED PAGE 9.

26          PROC PRINT DATA = SAMPLE10;

NOTE: THE PROCEDURE PRINT USED 0.25 SECONDS AND 202K AND PRINTED PAGE 10.

27          PROC PLOT DATA = SAMPLE1; PLOT Y*X='*';
28          TITLE PLOT OF FAILURE VALUES VS LOG(N+1)/(N+1-I);
29          TITLE2 SAMPLE SIZE = 10;

NOTE: THE PROCEDURE PLOT USED 0.34 SECONDS AND 208K AND PRINTED PAGE 11.

30          PROC PLOT DATA = SAMPLE2; PLOT Y*X='*';

NOTE: THE PROCEDURE PLOT USED 0.29 SECONDS AND 208K AND PRINTED PAGE 12.

31          PROC PLOT DATA = SAMPLE3; PLOT Y*X='*';

NOTE: THE PROCEDURE PLOT USED 0.34 SECONDS AND 208K AND PRINTED PAGE 13.

32          PROC PLOT DATA = SAMPLE4; PLOT Y*X='*';

NOTE: THE PROCEDURE PLOT USED 0.34 SECONDS AND 208K AND PRINTED PAGE 14.

33          PROC PLOT DATA = SAMPLE5; PLOT Y*X='*';
```

```
3       S A S   L O G    OS SAS 82.2B    OS/360 MVT JOB VM005314 STEP SAS       PROC
NOTE: THE PROCEDURE PLOT USED 0.33 SECONDS AND 208K AND PRINTED PAGE 15.
34          PROC PLOT DATA = SAMPLE6; PLOT Y*X='*';
NOTE: THE PROCEDURE PLOT USED 0.31 SECONDS AND 208K AND PRINTED PAGE 16.
35          PROC PLOT DATA = SAMPLE7; PLOT Y*X='*';
NOTE: THE PROCEDURE PLOT USED 0.33 SECONDS AND 208K AND PRINTED PAGE 17.
36          PROC PLOT DATA = SAMPLE8; PLOT Y*X='*';
NOTE: THE PROCEDURE PLOT USED 0.32 SECONDS AND 208K AND PRINTED PAGE 18.
37          PROC PLOT DATA = SAMPLE9; PLOT Y*X='*';
NOTE: THE PROCEDURE PLOT USED 0.29 SECONDS AND 208K AND PRINTED PAGE 19.
38          PROC PLOT DATA = SAMPLE10; PLOT Y*X='*';
NOTE: THE PROCEDURE PLOT USED 0.33 SECONDS AND 208K AND PRINTED PAGE 20.
NOTE: SAS USED 324K MEMORY.

NOTE: SAS INSTITUTE INC.
      SAS CIRCLE
      PO BOX 8000
      CARY, N.C. 27511-8000
```

Appendix 3

SIMULATION PROGRAM

TEST FOR ABNORMALLY LATE FIRST FAILURES

This explanation precedes the FORTRAN/SAS program that simulates the test for abnormally early first failures. This program is very similar to the program shown in Appendix 2. However, this program generates 5 samples of size 10, 20 and 50 each in one single run. Line 5 initializes the value of the intercept A to 20, the value by which each of the failure times are incremented. Line 8 features the only other difference where the DO loop generates samples of varying size. The rest of the program is almost identical to the program shown in Appendix 2.

```
FORTRAN IV G LEVEL  21                    MAIN              DATE = 83197          14/55/40

              C
              C       THIS PROGRAM TESTS FOR ABNORMALLY LATE FIRST FAILURE.
              C
 0001                 INTEGER NR,NL,N1,N2,IER
 0002                 REAL TBAL(50),R(50),XM,TOTBAL,F,P,Y(50)
 0003                 DOUBLE PRECISION DSEED
 0004                 DSEED=123457.0D0
 0005                 A=20.
 0006                 NR=10
 0007                 XM=100.
 0008                 DO 110 L=1,3
 0009                 IF(L.EQ.3) NR=50
 0010                 DO 100 K=1,5
 0011                 CALL GGEXN (DSEED,XM,NR,R)
              C
              C       GGEXN IS AN IMSL SUBROUTINE THAT GENERATES RANDOM EXPONENTIAL
              C       DEVIATES. A CERTAIN SEED IS SPECIFIED AND 'NR' SPECIFIES THE
              C       NUMBER OF DEVIATES GENERATED. 'XR' IS THE VALUE OF THETA, THE
              C       MEAN LIFE IN THE EXPONENTIAL DISTRIBUTION. 'R' IS THE VECTOR
              C       OF VALUES GENERATED CORRESPONDING TO 'NR'.
              C
 0012                 CALL VSRTA (R,NR)
              C
              C       VSRTA IS AN IMSL SUBROUTINE THAT SORTS THE DATA.'R' CORRESPONDS
              C       TO THE INPUT VECTOR OF VALUES TO BE SORTED WHILE 'NR' CORRESPONDS
              C       TO THE NUMBER OF VALUES GENERATED.
              C
 0013                 WRITE(6,6)
 0014         6       FORMAT('1',10X,'THE SORTED FAILURE TIMES',10X,'LOG(N+1)/(N+1-I)'
                     $//)
 0015                 X=1
 0016                 DO 95 I=1,NR
 0017                 R(I)=R(I)+A
 0018                 Y(I)=ALOG((NR+1)/(NR+1-X))
 0019                 X=X+1
 0020                 WRITE(6,7)R(I),Y(I)
 0021          7      FORMAT(18X,F10.4,20X,F10.4)
 0022                 WRITE(10,8)R(I),Y(I),A
 0023          8      FORMAT(5X,F10.4,5X,F10.4,5X,F10.4)
 0024         95      CONTINUE
 0025                 TBAL(1)=R(1)*NR
              C
              C       TBAL(1) GIVES THE TOTAL LIFE TILL THE FIRST FAILURE.
```

```
FORTRAN IV G LEVEL  21                MAIN              DATE = 83197          14/55/40

      C
                TOTBAL=0.
                DO 10 I=2,NR
                TBAL(I)=(R(I)-R(I-1))*(NR+1-I)
      C
      C         TBAL(I) GIVES THE TOTAL LIFE FOR INDIVIDUAL FAILURES.
      C
                TOTBAL=TOTBAL + TBAL(I)
      C
      C         TOTBAL GIVES THE TOTAL LIFE FOR THE PERIODS REMAINING AFTER
      C         THE FIRST FAILURE HAS OCCURED.
      C
      10        CONTINUE
                WRITE(6,11) TBAL(1)
      11        FORMAT ('0',10X,'FIRST PARTIAL TOTAL LIFE :',F10.4)
                WRITE(6,12) TOTBAL
      12        FORMAT('0',10X,'TOTAL REMAINING LIFE :',F10.4)
                F=TBAL(1)/TOTBAL*(NR-1)
                N1=2
                N2=2*(NR-1)
                CALL MDFD (F,N1,N2,P,IER)
      C
      C         MDFD IS AN IMSL SUBROUTINE THAT GIVES THE RIGHT HAND AREA
      C         'P' CORRESPONDING TO A GIVEN COMPUTED F WITH N1 AND N2, THE
      C         NUMERATOR AND DEMONINATOR DEGREES OF FREEDOM RESPECTIVELY.
      C         'IER' IS THE ERROR VALUE
      C
                WRITE(6,13) F
      13        FORMAT('0',10X,'VALUE OF F :',F10.4)
                WRITE(6,14) P,IER
      14        FORMAT('0',10X,'P VALUE :',F10.4//11X,'ERROR VALUE :',I5)
                A=A+10.
      100       CONTINUE
                A=20.
                NR=NR+10
      110       CONTINUE
                STOP
                END
```

```
1        S A S   L O G    OS SAS 82.28    OS/360 MVT JOB VM145534 STEP SAS       PROC

NOTE: THE JOB VM145534 HAS BEEN RUN UNDER RELEASE 82.28 OF SAS
      AT KANSAS STATE UNIVERSITY (03010001).

NOTE: SAS OPTIONS SPECIFIED ARE:
      NOINCLUDE NOGRAPHICS LINESIZE=115,PAGESIZE=55 SORT=4

          DATA COMPLEX;
          INFILE DATASET;
          INPUT X 6-15 Y 21-30 INT_CEPT 3o-45;


NOTE: INFILE DATASET IS:
      DSNAME=SYS83197.T145248.RV000.VM145534.TEMP,
      UNIT=SYSDA,VOL=SER=KSCC5A,DISP=OLD,
      DCB=(BLKSIZE=6160,LRECL=80,RECFM=FB)

NOTE: 400 LINES WERE READ FROM INFILE DATASET.
NOTE: DATA SET WORK.COMPLEX HAS 400 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.51 SECONDS AND 202K.

          DATA SAMPLE1 SAMPLE2 SAMPLE3 SAMPLE4 SAMPLE5 SAMPLE6 SAMPLE7 SAMPLE8
          SAMPLE9 SAMPLE10 SAMPLE11 SAMPLE12 SAMPLE13 SAMPLE14 SAMPLE15;
          SET COMPLEX;
          IF             _N_<=10 THEN OUTPUT SAMPLE1;
          IF _N_ >10 AND _N_ <=20 THEN OUTPUT SAMPLE2;
          IF _N_ >20 AND _N_ <=30 THEN OUTPUT SAMPLE3;
          IF _N_ >30 AND _N_ <=40 THEN OUTPUT SAMPLE4;
          IF _N_ >40 AND _N_ <=50 THEN OUTPUT SAMPLE5;
          IF _N_ >50 AND _N_ <=70 THEN OUTPUT SAMPLE6;
          IF _N_ >70 AND _N_ <=90 THEN OUTPUT SAMPLE7;
          IF _N_ >90 AND _N_ <=110 THEN OUTPUT SAMPLE8;
          IF _N_ >110 AND _N_ <=130 THEN OUTPUT SAMPLE9;
          IF _N_ >130 AND _N_ <=150 THEN OUTPUT SAMPLE10;
          IF _N_ >150 AND _N_ <=200 THEN OUTPUT SAMPLE11;
          IF _N_ >200 AND _N_ <=250 THEN OUTPUT SAMPLE12;
          IF _N_ >250 AND _N_ <=300 THEN OUTPUT SAMPLE13;
          IF _N_ >300 AND _N_ <=350 THEN OUTPUT SAMPLE14;
          IF _N_ >350 AND _N_ <=400 THEN OUTPUT SAMPLE15;

NOTE: DATA SET WORK.SAMPLE1 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE2 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE3 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE4 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE5 HAS 10 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE6 HAS 20 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE7 HAS 20 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE8 HAS 20 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE9 HAS 20 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE10 HAS 20 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE11 HAS 50 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE12 HAS 50 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE13 HAS 50 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
```

```
2        S A S   L O G    OS SAS 82.28    OS/360 MVT JOB VM145534 STEP SAS       PROC

NOTE: DATA SET WORK.SAMPLE14 HAS 50 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE15 HAS 50 OBSERVATIONS AND 3 VARIABLES. 680 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.82 SECONDS AND 432K.

22          PROC PRINT DATA = SAMPLE1;TITLE1 THETA VALUE = 100;
23          TITLE2 SAMPLE OF SIZE =10;

NOTE: THE PROCEDURE PRINT USED 0.31 SECONDS AND 202K AND PRINTED PAGE 1.

24          PROC PRINT DATA = SAMPLE2;

NOTE: THE PROCEDURE PRINT USED 0.24 SECONDS AND 202K AND PRINTED PAGE 2.

25          PROC PRINT DATA = SAMPLE3;

NOTE: THE PROCEDURE PRINT USED 0.28 SECONDS AND 202K AND PRINTED PAGE 3.

26          PROC PRINT DATA = SAMPLE4;

NOTE: THE PROCEDURE PRINT USED 0.26 SECONDS AND 202K AND PRINTED PAGE 4.

27          PROC PRINT DATA = SAMPLE5;

NOTE: THE PROCEDURE PRINT USED 0.27 SECONDS AND 202K AND PRINTED PAGE 5.

28          PROC PRINT DATA = SAMPLE6;TITLE2 SAMPLE OF SIZE 20;

NOTE: THE PROCEDURE PRINT USED 0.29 SECONDS AND 202K AND PRINTED PAGE 6.

29          PROC PRINT DATA = SAMPLE7;

NOTE: THE PROCEDURE PRINT USED 0.30 SECONDS AND 202K AND PRINTED PAGE 7.

30          PROC PRINT DATA = SAMPLE8;

NOTE: THE PROCEDURE PRINT USED 0.29 SECONDS AND 202K AND PRINTED PAGE 8.

31          PROC PRINT DATA = SAMPLE9;

NOTE: THE PROCEDURE PRINT USED 0.31 SECONDS AND 202K AND PRINTED PAGE 9.

32          PROC PRINT DATA = SAMPLE10;

NOTE: THE PROCEDURE PRINT USED 0.27 SECONDS AND 202K AND PRINTED PAGE 10.

33          PROC PRINT DATA = SAMPLE11;TITLE2 SAMPLE OF SIZE 50;

NOTE: THE PROCEDURE PRINT USED 0.35 SECONDS AND 202K AND PRINTED PAGE 11.

34          PROC PRINT DATA = SAMPLE12;

NOTE: THE PROCEDURE PRINT USED 0.32 SECONDS AND 202K AND PRINTED PAGE 12.
```

```
35        PROC PRINT DATA = SAMPLE13;

NOTE: THE PROCEDURE PRINT USED 0.32 SECONDS AND 202K AND PRINTED PAGE 13.

36        PROC PRINT DATA = SAMPLE14;

NOTE: THE PROCEDURE PRINT USED 0.32 SECONDS AND 202K AND PRINTED PAGE 14.

37        PROC PRINT DATA = SAMPLE15;

NOTE. THE PROCEDURE PRINT USED 0.35 SECONDS AND 202K AND PRINTED PAGE 15.

38        PROC PLOT DATA = SAMPLE1;PLOT Y*X='*'/HZERO;
39        TITLE1 PLOT OF FAILURE TIMES VS LOG(N+1)/(N+1-I);
40        TITLE2 SAMPLE SIZE = 10;

NOTE: THE PROCEDURE PLOT USED 0.39 SECONDS AND 208K AND PRINTED PAGE 16.

41        PROC PLOT DATA = SAMPLE2;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.37 SECONDS AND 208K AND PRINTED PAGE 17.

42        PROC PLOT DATA = SAMPLE3;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.38 SECONDS AND 208K AND PRINTED PAGE 18.

43        PROC PLOT DATA = SAMPLE4;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.35 SECONDS AND 208K AND PRINTED PAGE 19.

44        PROC PLOT DATA = SAMPLE5;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.33 SECONDS AND 208K AND PRINTED PAGE 20.

45        PROC PLOT DATA = SAMPLE6;PLOT Y*X='*'/HZERO;TITLE2 SAMPLE SIZE = 20;

NOTE: THE PROCEDURE PLOT USED 0.33 SECONDS AND 208K AND PRINTED PAGE 21.

46        PROC PLOT DATA = SAMPLE7;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.35 SECONDS AND 208K AND PRINTED PAGE 22.

47        PROC PLOT DATA = SAMPLE8;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.34 SECONDS AND 208K AND PRINTED PAGE 23.

48        PROC PLOT DATA = SAMPLE9;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.32 SECONDS AND 208K AND PRINTED PAGE 24.

49        PROC PLOT DATA = SAMPLE10;PLOT Y*X='*'/HZERO;
```

```
4       S A S   L O G    OS SAS 82.2B    OS/360 MVT JOB VM145534 STEP SAS       PROC

NOTE: THE PROCEDURE PLOT USED 0.37 SECONDS AND 208K AND PRINTED PAGE 25.

50          PROC PLOT DATA = SAMPLE11;PLOT Y*X='*'/HZERO;TITLE2 SAMPLE SIZE = 50;

NOTE: THE PROCEDURE PLOT USED 0.38 SECONDS AND 208K AND PRINTED PAGE 26.

51          PROC PLOT DATA = SAMPLE12;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.36 SECONDS AND 208K AND PRINTED PAGE 27.

52          PROC PLOT DATA = SAMPLE13;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.38 SECONDS AND 208K AND PRINTED PAGE 28.

53          PROC PLOT DATA = SAMPLE14;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.43 SECONDS AND 208K AND PRINTED PAGE 29.

54          PROC PLOT DATA = SAMPLE15;PLOT Y*X='*'/HZERO;

NOTE: THE PROCEDURE PLOT USED 0.42 SECONDS AND 208K AND PRINTED PAGE 30.
NOTE: SAS USED 432K MEMORY.

NOTE: SAS INSTITUTE INC.
      SAS CIRCLE
      PO BOX 8000
      CARY, N.C. 27511-8000
```

Appendix 4

SIMULATION PROGRAM

GENERATION OF CONTAMINATED SAMPLES

This explanation precedes the FORTRAN/SAS program that simulates the generation of contaminated samples. Line 5 of the program initializes the mean value of the parent exponential distribution to 100 while line 6 initializes the mean value of the distribution used to overlay the original values to 50. Line 7 initializes the value of the variable used to control the sample size.

In line 9 the DO loop generates samples of different sizes. Sizes of 10, 20 and 50 are generated. Line 14 is the outer DO loop that controls the level of contamination from 10% to 90%. Line 15 in the inner DO loop actually performs the overlay of values generated on the second passage through the loop.

Statement 10 initializes the value of G to 0.101. Since a value slightly greater than 0.1 is required for the accuracy of the INT function, this value was chosen. Line 20 stores the real value of G times the sample size in TEMP. N(2) in line 21 is the integer value of TEMP which on the first pass is seen to be 1 for a sample size of 10. The values obtained are then sorted in ascending order using line 27. A DO loop in line 33 generates the y values and these together with the sample of failure values X(I) are written to disk and then passed on to a SAS program which prints and plots these values. The SAS layout is very similar to the situations previously discussed in Appendix 3 and 4 and is rather self explanatory.

FORTRAN IV G LEVEL 21 MAIN DATE = 83195 23/01/04

```
C
C      THIS PROGRAM TESTS FOR THE EFFECTS OF CONTAMINATION BY DUDS.
C      A RANDOM SAMPLE OF EXPONENTIAL DEVIATES ARE GENERATED USING
C      A PARTICULAR SEED AND A CERTAIN VALUE OF THETA. THE MEAN LIFE
C      OF THE EXPONENTIAL DISTRIBUTION. A PROPORTION OF THESE VALUES
C      ARE OVERLAYED BY ANOTHER SET OF RANDOM EXPONENTIAL DEVIATES
C      GENERATED USING ANOTHER SEED AND A DIFFERENT VALUE OF THETA.
C      THESE VALUES ARE THEN INCORPORATED IN A SAS PROGRAM AND THEIR
C      EFFECT ON LOG (N+1)/(N+1-I) IS STUDIED USING PLOTS GENERATED.
C
       INTEGER NR,IER,N(2),N1,N2
       REAL R(150),TOTBAL,XM,F,P,Y(150),THETA(2),TBAL(150)
       DOUBLE PRECISION DSEED
       DSEED=123457.000
       THETA(1)=100.
       THETA(2)=50.
       N(1)=0
       XM=THETA(1)
C                     DO LOOP FOR DIFFERENT SAMPLE SIZES BEGINS HERE
       DO 120 M=1,3
C
       G=0.101
       N(1)=N(1)+10
       IF (M.EQ.3) N(1)=50
       NR=N(1)
C                     DO LOOP FOR VARIOUS CONTAMINATION LEVELS BEGINS HERE
       DO 115 L=1,9
C                     DO LOOP FOR CONTAMINATING A SINGLE SAMPLE BEGINS HERE
       DO 100 K=1,2
C
       CALL GGEXN(DSEED,XM,NR,R)
       DSEED=DSEED + 3457.000
       XM=THETA(2)
       IF(K.EQ.2) GO TO 100
       TEMP=G*N(1)
       N(2)=INT(TEMP)
       NR=N(2)
  100  CONTINUE
       NR=N(1)
       XM=THETA(1)
       J=N(2)
       CALL VSRTA(R,NR)
       WRITE(6,1)J,NR,THETA(1),THETA(2)
```

Simulation Program - Generation of Contaminated samples

```
FORTRAN IV G LEVEL  21                MAIN              DATE = 83195          23/01/04

0029         1      FORMAT('1','NUMBER OF VALUES OVERLAYED :',I5//1X,'THE SAMPLE
             *     $SIZE :',I5//1X,'THETA1 VALUE :',F6.2//1X,'THETA2 VALUE :',F6.2//)
0030                WRITE(6,2)
0031         2      FORMAT(' ','FAILURE VALUES',8X,'Y VALUES',3X,'DEFECTIVES'/)
0032                X=1.
             C                   DO LOOP FOR GENERATING THE LOG FUNCTION AND
             C                   PRINTING THEIR VALUES TOGETHER WITH FAILURE TIMES.
0033                DO 110 I=1,NR
0034                Y(I)=ALOG((NR+1)/(NR+1-X))
0035                X=X+1.
0036                WRITE(6,19)R(I),Y(I),J
0037        19      FORMAT(5X,F10.4,5X,F10.4,5X,I5)
0038                WRITE(10,20)R(I),Y(I),J ,THETA(1),THETA(2)
0039        20      FORMAT(5X,F10.4,5X,F10.4,5X,I5,5X,F10.4,5X,F10.4)
0040       110      CONTINUE
0041                TBAL(1)=R(1)*NR
0042                TOTBAL=0
0043                DO 111 I=2,NR
0044                TBAL(I)=(R(I)-R(I-1))*(NR+1-I)
0045                TOTBAL=TOTBAL + TBAL(I)
0046       111      CONTINUE
0047                WRITE(6,112) TBAL(1),TOTBAL
0048       112      FORMAT('0','FIRST PARTIAL TOTAL LIFE :',F10.4//1X,
                   $'TOTAL REMAINING LIFE :',F10.4)
0049                F=TBAL(1)/TOTBAL*(NR-1)
0050                N1=2
0051                N2=2*(NR-1)
0052                CALL MDFD(F,N1,N2,P,IER)
0053                WRITE(6,113)F,P,IER
0054       113      FORMAT('0','VALUE OF F :',F10.4//1X,'P VALUE :',F10.4//
                   $1X,'ERROR VALUE :',I5)
0055                G=G+0.1
0056       115      CONTINUE
0057       120      CONTINUE
0058                STOP
0059                END
```

```
1        S A S   L O G    OS SAS 82.28     OS/360 MVT JOB VM230058 STEP SAS       PROC

NOTE: THE JOB VM230058 HAS BEEN RUN UNDER RELEASE 82.28 OF SAS
      AT KANSAS STATE UNIVERSITY (03010001).

NOTE: SAS OPTIONS SPECIFIED ARE:
      NOINCLUDE NOGRAPHICS LINESIZE=115,PAGESIZE=55 SORT=4

1          DATA COMPLEX;
2          INFILE DATASET;
3          INPUT X 6-15 Y 21-30 NUM_DEF 36-40 THETA1 46-55 THETA2 61-70;


NOTE: INFILE DATASET IS:
      DSNAME=SYS83195.T225919.RV000.VM230058.TEMP,
      UNIT=SYSDA,VOL=SER=KSCC5A,DISP=OLD,
      DCB=(BLKSIZE=6160,LRECL=80,RECFM=FB)

NOTE: 720 LINES WERE READ FROM INFILE DATASET.
NOTE: DATA SET WORK.COMPLEX HAS 720 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: THE DATA STATEMENT USED 1.01 SECONDS AND 202K.

4          DATA SAMPLE1 SAMPLE2 SAMPLE3 SAMPLE4 SAMPLE5 SAMPLE6 SAMPLE7 SAMPLE8
5          SAMPLE9 SAMPLE10 SAMPLE11 SAMPLE12 SAMPLE13 SAMPLE14 SAMPLE15
6          SAMPLE16 SAMPLE17 SAMPLE18 SAMPLE19 SAMPLE20 SAMPLE21 SAMPLE22
7          SAMPLE23 SAMPLE24 SAMPLE25 SAMPLE26 SAMPLE27;
8          SET COMPLEX;
9          IF            _N_<=10 THEN OUTPUT SAMPLE1;
10         IF _N_ >10 AND _N_ <=20 THEN OUTPUT SAMPLE2;
11         IF _N_ >20 AND _N_ <=30 THEN OUTPUT SAMPLE3;
12         IF _N_ >30 AND _N_ <=40 THEN OUTPUT SAMPLE4;
13         IF _N_ >40 AND _N_ <=50 THEN OUTPUT SAMPLE5;
14         IF _N_ >50 AND _N_ <=60 THEN OUTPUT SAMPLE6;
15         IF _N_ >60 AND _N_ <=70 THEN OUTPUT SAMPLE7;
16         IF _N_ >70 AND _N_ <=80 THEN OUTPUT SAMPLE8;
17         IF _N_ >80 AND _N_ <=90 THEN OUTPUT SAMPLE9;
18         IF _N_ >90 AND _N_ <=110 THEN OUTPUT SAMPLE10;
19         IF _N_ >110 AND _N_ <=130 THEN OUTPUT SAMPLE11;
20         IF _N_ >130 AND _N_ <=150 THEN OUTPUT SAMPLE12;
21         IF _N_ >150 AND _N_ <=170 THEN OUTPUT SAMPLE13;
22         IF _N_ >170 AND _N_ <=190 THEN OUTPUT SAMPLE14;
23         IF _N_ >190 AND _N_ <=210 THEN OUTPUT SAMPLE15;
24         IF _N_ >210 AND _N_ <=230 THEN OUTPUT SAMPLE16;
25         IF _N_ >230 AND _N_ <=250 THEN OUTPUT SAMPLE17;
26         IF _N_ >250 AND _N_ <=270 THEN OUTPUT SAMPLE18;
27         IF _N_ >270 AND _N_ <=320 THEN OUTPUT SAMPLE19;
28         IF _N_ >320 AND _N_ <=370 THEN OUTPUT SAMPLE20;
29         IF _N_ >370 AND _N_ <=420 THEN OUTPUT SAMPLE21;
30         IF _N_ >420 AND _N_ <=470 THEN OUTPUT SAMPLE22;
31         IF _N_ >470 AND _N_ <=520 THEN OUTPUT SAMPLE23;
32         IF _N_ >520 AND _N_ <=570 THEN OUTPUT SAMPLE24;
33         IF _N_ >570 AND _N_ <=620 THEN OUTPUT SAMPLE25;
34         IF _N_ >620 AND _N_ <=670 THEN OUTPUT SAMPLE26;
35         IF _N_ >670 AND _N_ <=720 THEN OUTPUT SAMPLE27;
```

```
2         S A S   L O G    OS SAS 82.2B    OS/360 MVT JOB VM230058 STEP SAS         PROC

NOTE: DATA SET WORK.SAMPLE1 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE2 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE3 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE4 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE5 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE6 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE7 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE8 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE9 HAS 10 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE10 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE11 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE12 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE13 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE14 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE15 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE16 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE17 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE18 HAS 20 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE19 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE20 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE21 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE22 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE23 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE24 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE25 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE26 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: DATA SET WORK.SAMPLE27 HAS 50 OBSERVATIONS AND 5 VARIABLES. 433 OBS/TRK.
NOTE: THE DATA STATEMENT USED 1.70 SECONDS AND 692K.

36          PROC PRINT DATA = SAMPLE1;TITLE CONTAMINATED SAMPLE OF SIZE 10;

NOTE: THE PROCEDURE PRINT USED 0.29 SECONDS AND 202K AND PRINTED PAGE 1.

37          PROC PRINT DATA = SAMPLE2;

NOTE: THE PROCEDURE PRINT USED 0.28 SECONDS AND 202K AND PRINTED PAGE 2.

38          PROC PRINT DATA = SAMPLE3;

NOTE: THE PROCEDURE PRINT USED 0.31 SECONDS AND 202K AND PRINTED PAGE 3.

39          PROC PRINT DATA = SAMPLE4;

NOTE: THE PROCEDURE PRINT USED 0.28 SECONDS AND 202K AND PRINTED PAGE 4.

40          PROC PRINT DATA = SAMPLE5;

NOTE: THE PROCEDURE PRINT USED 0.29 SECONDS AND 202K AND PRINTED PAGE 5.

41          PROC PRINT DATA = SAMPLE6;
```

```
NOTE: THE PROCEDURE PRINT USED 0.27 SECONDS AND 202K AND PRINTED PAGE 6.
42          PROC PRINT DATA = SAMPLE7;
NOTE: THE PROCEDURE PRINT USED 0.28 SECONDS AND 202K AND PRINTED PAGE 7.
43          PROC PRINT DATA = SAMPLE8;
NOTE: THE PROCEDURE PRINT USED 0.25 SECONDS AND 202K AND PRINTED PAGE 8.
44          PROC PRINT DATA = SAMPLE9;
NOTE: THE PROCEDURE PRINT USED 0.25 SECONDS AND 202K AND PRINTED PAGE 9.
45          PROC PRINT DATA = SAMPLE10;TITLE CONTAMINATED SAMPLE OF SIZE 20;
NOTE: THE PROCEDURE PRINT USED 0.32 SECONDS AND 202K AND PRINTED PAGE 10.
46          PROC PRINT DATA = SAMPLE11;
NOTE: THE PROCEDURE PRINT USED 0.32 SECONDS AND 202K AND PRINTED PAGE 11.
47          PROC PRINT DATA = SAMPLE12;
NOTE: THE PROCEDURE PRINT USED 0.31 SECONDS AND 202K AND PRINTED PAGE 12.
48          PROC PRINT DATA = SAMPLE13;
NOTE: THE PROCEDURE PRINT USED 0.30 SECONDS AND 202K AND PRINTED PAGE 13.
49          PROC PRINT DATA = SAMPLE14;
NOTE: THE PROCEDURE PRINT USED 0.29 SECONDS AND 202K AND PRINTED PAGE 14.
50          PROC PRINT DATA = SAMPLE15;
NOTE: THE PROCEDURE PRINT USED 0.30 SECONDS AND 202K AND PRINTED PAGE 15.
51          PROC PRINT DATA = SAMPLE16;
NOTE: THE PROCEDURE PRINT USED 0.28 SECONDS AND 202K AND PRINTED PAGE 16.
52          PROC PRINT DATA = SAMPLE17;
NOTE: THE PROCEDURE PRINT USED 0.31 SECONDS AND 202K AND PRINTED PAGE 17.
53          PROC PRINT DATA = SAMPLE18;
NOTE: THE PROCEDURE PRINT USED 0.31 SECONDS AND 202K AND PRINTED PAGE 18.
54          PROC PRINT DATA = SAMPLE19;TITLE CONTAMINATED SAMPLE OF SIZE 50;
```

```
4        S A S   L O G    OS SAS 82.2D     OS/360 MVT JOB VM230058 STEP SAS        PROC

NOTE: THE PROCEDURE PRINT USED 0.40 SECONDS AND 202K AND PRINTED PAGE 19.

55          PROC PRINT DATA = SAMPLE20;

NOTE: THE PROCEDURE PRINT USED 0.37 SECONDS AND 202K AND PRINTED PAGE 20.

56          PROC PRINT DATA = SAMPLE21;

NOTE: THE PROCEDURE PRINT USED 0.40 SECONDS AND 202K AND PRINTED PAGE 21.

57          PROC PRINT DATA = SAMPLE22;

NOTE: THE PROCEDURE PRINT USED 0.41 SECONDS AND 202K AND PRINTED PAGE 22.

58          PROC PRINT DATA = SAMPLE23;

NOTE: THE PROCEDURE PRINT USED 0.40 SECONDS AND 202K AND PRINTED PAGE 23.

59          PROC PRINT DATA = SAMPLE24;

NOTE: THE PROCEDURE PRINT USED 0.38 SECONDS AND 202K AND PRINTED PAGE 24.

60          PROC PRINT DATA = SAMPLE25;

NOTE: THE PROCEDURE PRINT USED 0.39 SECONDS AND 202K AND PRINTED PAGE 25.

61          PROC PRINT DATA = SAMPLE26;

NOTE: THE PROCEDURE PRINT USED 0.39 SECONDS AND 202K AND PRINTED PAGE 26.

62          PROC PRINT DATA = SAMPLE27;

NOTE: THE PROCEDURE PRINT USED 0.40 SECONDS AND 202K AND PRINTED PAGE 27.

63          PROC PLOT DATA = SAMPLE1;PLOT Y*X='*';
64          TITLE1 PLOT OF FAILURE TIMES VS LOG(N+1)/(N+1-I);
65          TITLE2 SAMPLE SIZE = 10;
66          TITLE3 CONTAMINATION COUNT = 1;

NOTE: THE PROCEDURE PLOT USED 0.37 SECONDS AND 208K AND PRINTED PAGE 28.

67          PROC PLOT DATA = SAMPLE2;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 2;

NOTE: THE PROCEDURE PLOT USED 0.36 SECONDS AND 208K AND PRINTED PAGE 29.

68          PROC PLOT DATA = SAMPLE3;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 3;

NOTE: THE PROCEDURE PLOT USED 0.39 SECONDS AND 208K AND PRINTED PAGE 30.

69          PROC PLOT DATA = SAMPLE4;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 4;

NOTE: THE PROCEDURE PLOT USED 0.40 SECONDS AND 208K AND PRINTED PAGE 31.
```

```
70        PROC PLOT DATA = SAMPLE5;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 5;
NOTE: THE PROCEDURE PLOT USED 0.36 SECONDS AND 208K AND PRINTED PAGE 32.
71        PROC PLOT DATA = SAMPLE6;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 6;
NOTE: THE PROCEDURE PLOT USED 0.40 SECONDS AND 208K AND PRINTED PAGE 33.
72        PROC PLOT DATA = SAMPLE7;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 7;
NOTE: THE PROCEDURE PLOT USED 0.40 SECONDS AND 208K AND PRINTED PAGE 34.
73        PROC PLOT DATA = SAMPLE8;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 8;
NOTE: THE PROCEDURE PLOT USED 0.33 SECONDS AND 208K AND PRINTED PAGE 35.
74        PROC PLOT DATA = SAMPLE9;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 9;
NOTE: THE PROCEDURE PLOT USED 0.38 SECONDS AND 208K AND PRINTED PAGE 36.
75        PROC PLOT DATA = SAMPLE10;PLOT Y*X='*';TITLE2 SAMPLE SIZE = 20;
76        TITLE3 CONTAMINATION COUNT = 2;
NOTE: THE PROCEDURE PLOT USED 0.39 SECONDS AND 208K AND PRINTED PAGE 37.
77        PROC PLOT DATA = SAMPLE11;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 4;
NOTE: THE PROCEDURE PLOT USED 0.37 SECONDS AND 208K AND PRINTED PAGE 38.
78        PROC PLOT DATA = SAMPLE12;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 6;
NOTE: THE PROCEDURE PLOT USED 0.36 SECONDS AND 208K AND PRINTED PAGE 39.
79        PROC PLOT DATA = SAMPLE13;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 8;
NOTE: THE PROCEDURE PLOT USED 0.34 SECONDS AND 208K AND PRINTED PAGE 40.
80        PROC PLOT DATA = SAMPLE14;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 10;
NOTE: THE PROCEDURE PLOT USED 0.37 SECONDS AND 208K AND PRINTED PAGE 41.
81        PROC PLOT DATA = SAMPLE15;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 12;
NOTE: THE PROCEDURE PLOT USED 0.37 SECONDS AND 208K AND PRINTED PAGE 42.
82        PROC PLOT DATA = SAMPLE16;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 14;
NOTE: THE PROCEDURE PLOT USED 0.36 SECONDS AND 208K AND PRINTED PAGE 43.
83        PROC PLOT DATA = SAMPLE17;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 16;
```

```
NOTE: THE PROCEDURE PLOT USED 0.36 SECONDS AND 208K AND PRINTED PAGE 44.
84          PROC PLOT DATA = SAMPLE18;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 18;
NOTE: THE PROCEDURE PLOT USED 0.33 SECONDS AND 208K AND PRINTED PAGE 45.
85          PROC PLOT DATA = SAMPLE19;PLOT Y*X='*';TITLE2 SAMPLE SIZE = 50;
86          TITLE3 CONTAMINATION COUNT = 5;
NOTE: THE PROCEDURE PLOT USED 0.35 SECONDS AND 208K AND PRINTED PAGE 46.
87          PROC PLOT DATA = SAMPLE20;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 10;
NOTE: THE PROCEDURE PLOT USED 0.35 SECONDS AND 208K AND PRINTED PAGE 47.
88          PROC PLOT DATA = SAMPLE21;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 15;
NOTE: THE PROCEDURE PLOT USED 0.36 SECONDS AND 208K AND PRINTED PAGE 48.
89          PROC PLOT DATA = SAMPLE22;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 20;
NOTE: THE PROCEDURE PLOT USED 0.38 SECONDS AND 208K AND PRINTED PAGE 49.
90          PROC PLOT DATA = SAMPLE23;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 25;
NOTE: THE PROCEDURE PLOT USED 0.33 SECONDS AND 208K AND PRINTED PAGE 50.
91          PROC PLOT DATA = SAMPLE24;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 30;
NOTE: THE PROCEDURE PLOT USED 0.36 SECONDS AND 208K AND PRINTED PAGE 51.
92          PROC PLOT DATA = SAMPLE25;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 35;
NOTE: THE PROCEDURE PLOT USED 0.35 SECONDS AND 208K AND PRINTED PAGE 52.
93          PROC PLOT DATA = SAMPLE26;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 40;
NOTE: THE PROCEDURE PLOT USED 0.41 SECONDS AND 208K AND PRINTED PAGE 53.
94          PROC PLOT DATA = SAMPLE27;PLOT Y*X='*';TITLE3 CONTAMINATION COUNT = 45;
NOTE: THE PROCEDURE PLOT USED 0.35 SECONDS AND 208K AND PRINTED PAGE 54.
NOTE: SAS USED 692K MEMORY.

NOTE: SAS INSTITUTE INC.
      SAS CIRCLE
      PO BOX 8000
      CARY, N.C. 27511-8000
```

REFERENCES

1. Epstein B. and Sobel M. "Life testing". American Statistical Association Journal, Vol.48 pp 487-501 Sept. 1953.

2. Epstein B. and Sobel M. "Some theorems relevant to life testing from an exponential distribution". Annals of Mathematical Statistics, Vol.25 pp 373-381. 1954.

3(a) Epstein B. "Tests for the validity of the assumption that the underlying distribution of life is exponential". Part 1. Technometrics, Vol.2, Number 1. pp 83-101 Feb. 1960.

3(b) Epstein B. "Tests for the validity of the assumption that the underlying distribution of life is exponential". Part 2. Technometrics, Vol.2, Number 2. pp 167-183 Mar. 1960.

4. Epstein B. "Estimation of the parameters of two-parameter exponential distributions from censored samples". Technometrics, Vol.2, Number 3, pp 403-406 Aug. 1960

5. Epstein B. "Estimation from life test data". Technometrics, Vol.2, Number 4, pp 447-454 Nov. 1960

6. Statistical Analysis System(SAS) User's Guide. SAS Institute Inc. P.O. Box 10066, Raleigh, North Carolina 27605.

SIMULATION FOR TESTS ON THE VALIDITY OF THE ASSUMPTION THAT THE UNDERLYING DISTRIBUTION OF LIFE IS EXPONENTIAL

by

ANJO THOPPIL

B.E.(Production Engineering), University of Madras, Madras, India, 1980

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1984

he objective of this report is to demonstrate, with the aid f simulation, how the underlying distribution of component ife may be recognized as exponential, without resort to the tatistical maneuvers used by Benjamin Epstein in the series f articles he wrote in the 1960's. Specifically, the report eals with graphical procedures, abnormally early or late first ailures and the effects of contamination. Contamination involved overlaying values from a parent exponential distribution ith values arising out of a different exponential distribution. rograms are written in FORTRAN and SAS and are presented with xplanations in the appendices.