NOTICE

NASA

Technical Memorandum 82052

Estimation Procedures to Measure and Monitor Failure Rates of Components During Thermal-Vacuum Testing

(NASA-TM-82052) ESTIMATION PROCEDURES TO MEASURE AND MONITOR FAILURE RATES OF COMPONENTS DURING THERMAL-VACUUM TESTING (NASA) 30 p HC A03/MF A01 CSCL 14D

N81-15383

Unclas

G3/38 12803

R. E. Williams and Raymond Kruger

NOVEMBER 1980

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

TM 82052

ESTIMATION PROCEDURES TO MEASURE AND MONITOR FAILURE RATES OF COMPONENTS DURING THERMAL-VACUUM TESTING

R. E. Williams*

and

Raymond Kruger

November 1980

*Dr. Williams took part in this study during his association with the NASA/Goddard Space Flight Center through the Summer Faculty Fellowship Program.

GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland

ESTIMATION PROCEDURES TO MEASURE AND MONITOR FAILURE RATES OF COMPONENTS DURING THERMAL-VACUUM TESTING

R. E. Williams
and
Raymond Kruger

ABSTRACT

This report describes estimation procedures for measuring component failure rates, for comparing the failure rates of two different groups of components, and for formulating confidence intervals for testing hypotheses (based on failure rates) that the two groups perform similarly or differently. Appendix A contains an example of an analysis in which these methods are applied to investigate the characteristics of two groups of spacecraft components.

The estimation procedures are adaptable to system level testing and to monitoring failure characteristics in orbit.

CONTENTS

ESTIMATION PROCEDURES TO MEASURE AND MONITOR FAILURE RATES OF COMPONENTS DURING THERMAL-VACUUM TESTING

INTRODUCTION

The performance of spacecraft components in test affects the expenditure of resources. This performance may also be indicative of later performance. If one could define the failure characteristics of this performance and could compare the performance between two groups, one a test group and the other a control or norm, then one could better evaluate whether to expend more or less resources in developing these components and also predict the performance of these components during later phases of the program.

It was the purpose of this study to develop methods for both defining the performance of a group of components and for quantifying the similarity of this performance to that of another group of components.

SECTION 1. ESTIMATION PROCEDURE TO MEASURE FAILURE RATES

The thermal-vacuum test process subjects individual components or groups of components to temperature and vacuum conditions similar to what they would encounter in space. (A component, as referred to in this report, is an assemblage of parts making up a basic unit or "black box" such as a transmitter or a power supply that performs a specific function in a system.) The duration of the test may vary.

During a test, a failure may occur that completely disables the unit. In this case the test is generally stopped, and repairs are made before the test is continued. Sometimes a failure occurs that does not totally disable the component. One may discover this failure at any time between its occurrence and the end of the test. The failure is then repaired, and the test manager may decide to retest. In some cases, a test is terminated with no failures. Therefore, we are looking at a time process that has failures (deaths), terminations (losses), and late entries (where repairs are made and the item is retested). (It should be noted that failure discovery rates rather than failure rates are actually monitored.)

To model the process, we use the Kaplan-Meier Product Limit (PL) estimation procedure as defined and described in (1), (2), and (3) and apply it to data as collected in Kruger-Norris (4).

The procedure, used in the Kruger-Norris thermal-vacuum test optimization model (4), normalizes the data and uses certain criteria for goodness of fit to estimate the parameters of a modified Duane model (5). From these parameters, one can project the failure rate into system level testing and then on to orbital operations. These projections are used to calculate a parameter called "availability" in reference (4) which, in turn, is used in the determination of optimal test conditions.

The PL estimation technique provides a basis for defining failure rates and also allows for statistical comparisons as shown later. In order to develop PL estimates, we define the probability of survival, P(t), as

$$P(t) = P(T > t). \quad (1)$$

T denotes the event that the component or collection of components has survived, and $P(T > t)$ means the probability that the component has survived beyond time t.

The PL estimate is based on procedures given in Kaplan-Meier (1) and estimates $P(t)$. The relationship between the PL estimate and the failure distribution for component life is given by

$$F(t) = 1 - P(t). \quad (2)$$

In component level testing, we have a distribution based on component count which is defined by dividing the number of failures that have occurred up until time t by the total number of components, N, remaining on test. This ratio is the cumulative frequency of component failures over time and will be called $F^*(t)$. Since the PL estimation procedure incorporates the same type of life information about component level testing as does $F^*(t)$,

$$F^*(t) = F(t) = 1 - P(t). \quad (3)$$

To estimate $F^*(t)$, we first estimate $P(t)$. The estimate of $F^*(t)$, $\hat{F}^*(t)$, is related to the estimate of $P(t)$, $\hat{P}(t)$, by

$$\hat{F}^*(t) = 1 - \hat{P}(t). \quad (4)$$

The PL estimate is a consistent estimate of $\hat{P}(t)$ and thus a consistent estimate of $\hat{F}^*(t)$.

In order to calculate the PL estimate, we do the following:

(a) The age (or time) scale is divided into suitably chosen intervals, $(0, \mu_1)$, (μ_1, μ_2), (μ_2, μ_3), . . . The intervals may be chosen as time intervals where the μ_j's correspond to the failure times or may be intervals of constant length. For notation in this report, the time from μ_{j-1} to μ_j is designated as $\mu_{j-1,j}$.

A different estimation procedure exists for each case. By choosing time intervals of constant length, one may graphically monitor and compare two different processes based on two different methods of normalizing the data.

(b) For each interval $(\mu_{j-1,j})$, one estimates the proportion of items alive just after μ_{j-1} and that survive beyond μ_j.

(c) If t is a division point (it may be introduced specifically, if necessary) the proportion, P(t), of the population surviving beyond t is estimated by the product of the estimated proportion surviving each interval, $p_{j-1,j}$, for all intervals prior to t.

(d) If there are ties between the failure time and time that a test ends, the time that a test ends is taken to occur immediately after the corresponding failure time. If a failure and test time fall on the endpoint of an interval as defined in (a), then the test time goes into the following interval (to the right) as a loss, and the failure time stays in the interval in which it occurred.

(e) If a repair occurs and the test then continues, the loss enters the process as a negative quantity. Let y be defined as the loss measurement. This loss occurs when items go off test and when items come on test. Suppose one item appeared on test during a time interval and two items went off test during the corresponding time interval. The overall loss is $y = 2 - 1 = 1$. If no items went off test during this interval, then $y = -1$.

Under the restrictions given in (a) through (e), the PL estimate is

$$\hat{P}(t) = \prod_{j=1}^{t} \hat{p}_{j-1,j}(t). \tag{5}$$

To account for the variations previously outlined and to formulate the above process in terms of an iterative time process, we let

n_{j-1} ≡ Number of items on test at time μ_{j-1}.

$d_{j-1,j}$ ≡ Number of items failed during the interval $\mu_{j-1,j}$.

$y_{j-1,j}$ ≡ Number of items whose test terminates (truncates) or are put back on test (late entries)—entered as a positive or negative number respectively—during the interval $\mu_{j-1,j}$.

From the above definitions, we have

$$\beta_{j-1,j} = \frac{n_{j-1} - y_{j-1,j} - d_{j-1,j}}{n_{j-1} - y_{j-1,j}} . \qquad (6)$$

To illustrate the use of the above formulas, we consider the following examples:

<u>Example I</u>: Intervals Established by Times of Failures

Example Parameters:

Eight items on test
Failure times: 8, 31, 51, 92
Truncation times: 10, 27, 70, 121
Entry time: 80

The entry time corresponds to one of the items being repaired and then retested.

We illustrate the time line schematically in Figure 1 with a failure time indicated by F, a truncation time indicated by T, and an entry time indicated by E.

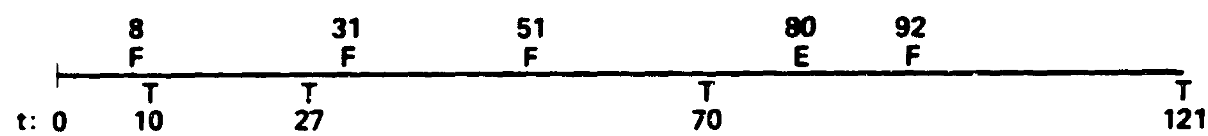
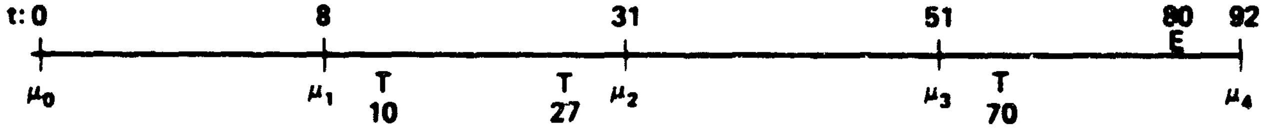


Figure 1. Time Line of Events for Examples I and II

The first calculation scheme uses intervals determined by the times of failures; thus, $\mu_0 = 0$, $\mu_1 = 8$, $\mu_2 = 31$, $\mu_3 = 51$, and $\mu_4 = 92$.

The calculation scheme is given below in Figure 2.



$n_0 = 8$ $d_{0,1} = 1$ $y_{0,1} = 0$	$n_1 = n_0 - y_{0,1} - d_{0,1}$ $= 8 - 0 - 1$ $= 7$ $d_{1,2} = 1$ $y_{1,2} = 2$	$n_2 = n_1 - y_{1,2} - d_{1,2}$ $= 7 - 2 - 1$ $= 4$ $d_{2,3} = 1$ $y_{2,3} = 0$	$n_3 = n_2 - y_{2,3} - d_{2,3}$ $= 4 - 0 - 1$ $= 3$ $d_{3,4} = 1$ $y_{3,4} = 1 - 1 = 0$
$\hat{p}_{0,1} = \dfrac{n_0 - y_{0,1} - d_{0,1}}{n_0 - y_{0,1}}$ $= \dfrac{8-0-1}{8-0}$ $= \dfrac{7}{8}$ $= 0.875$	$\hat{p}_{1,2} = \dfrac{n_1 - y_{1,2} - d_{1,2}}{n_1 - y_{1,2}}$ $= \dfrac{7-2-1}{7-2}$ $= \dfrac{4}{5}$ $= 0.800$	$\hat{p}_{2,3} = \dfrac{n_2 - y_{2,3} - d_{2,3}}{n_2 - y_{2,3}}$ $= \dfrac{4-0-1}{4-0}$ $= \dfrac{3}{4}$ $= 0.750$	$\hat{p}_{3,4} = \dfrac{n_3 - y_{3,4} - d_{3,4}}{n_3 - y_{3,4}}$ $= \dfrac{3-0-1}{3-0}$ $= \dfrac{2}{3}$ $= 0.667$

Figure 2. Calculation Scheme for Example I

Using equation (5), one can find the following PL's:

$$\hat{P}(8) = \frac{7}{8} = 0.875$$

$$\hat{P}(31) = \left(\frac{7}{8}\right) \cdot \left(\frac{4}{5}\right) = 0.700$$

$$\hat{P}(51) = \left(\frac{7}{8}\right) \cdot \left(\frac{4}{5}\right) \cdot \left(\frac{3}{4}\right) = 0.525$$

$$\hat{P}(92) = \left(\frac{7}{8}\right) \cdot \left(\frac{4}{5}\right) \cdot \left(\frac{3}{4}\right) \cdot \left(\frac{2}{3}\right) = 0.350$$

Using equation (4) we may then graph the probability of failure, $\hat{F}^*(t)$, as a function of time as shown below in Figure 3.

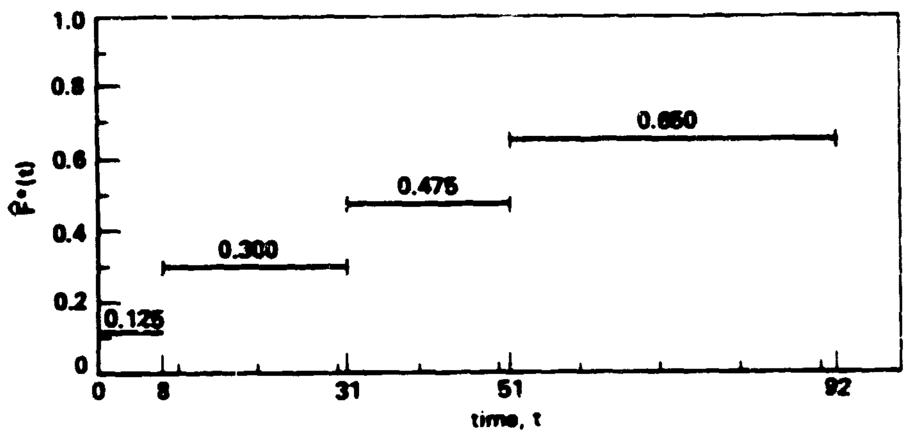


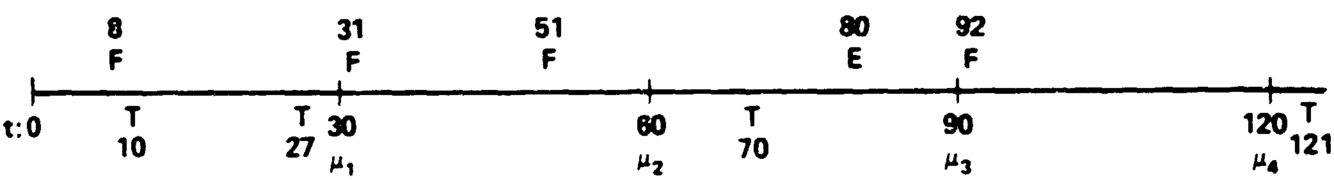
Figure 3. Probability of Failure vs. Time, Example I

To illustrate how the process changes when we change the time intervals to intervals of consistent length of thirty we have the following example:

<u>Example II.</u> Intervals Established at Fixed Times

$\mu_0 = 0,\ \mu_1 = 30,\ \mu_2 = 60,\ \mu_3 = 90,\ \mu_4 = 120$

The calculation scheme for these intervals is given in Figure 4.



$n_0 = 8$ $d_{0,1} = 1$ $y_{0,1} = 2$	$n_1 = n_0 - y_{0,1} - d_{0,1}$ $= 8 - 2 - 1$ $= 5$ $d_{1,2} = 2$ $y_{1,2} = 0$	$n_2 = n_1 - y_{1,2} - d_{1,2}$ $= 5 - 0 - 2$ $= 3$ $d_{2,3} = 0$ $y_{2,3} = 1 - 1$ $= 0$	$n_3 = n_2 - y_{2,3} - d_{2,3}$ $= 3 - 0 - 0$ $= 3$ $d_{3,4} = 1$ $y_{3,4} = 0$
$\hat{p}_{0,1} = \frac{n_0 - y_{0,1} - d_{0,1}}{n_0 - y_{0,1}}$	$\hat{p}_{1,2} = \frac{n_1 - y_{1,2} - d_{1,2}}{n_1 - y_{1,2}}$	$\hat{p}_{2,3} = \frac{n_2 - y_{2,3} - d_{2,3}}{n_2 - y_{2,3}}$	$\hat{p}_{3,4} = \frac{n_3 - y_{3,4} - d_{3,4}}{n_3 - y_{3,4}}$
$= \frac{8 - 2 - 1}{8 - 2}$	$= \frac{5 - 0 - 2}{5 - 0}$	$= \frac{3 - 0 - 0}{3 - 0}$	$= \frac{3 - 0 - 1}{3 - 0}$
$= \frac{5}{6}$	$= \frac{3}{5}$	$= \frac{3}{3}$	$= \frac{2}{3}$
$= 0.833$	$= 0.600$	$= 1.000$	$= 0.667$

Figure 4. Calculation Scheme for Example II

Again, $\hat{F}^*(t)$ may be calculated from equation (4); a plot of $\hat{F}^*(t)$ is shown in Figure 5.

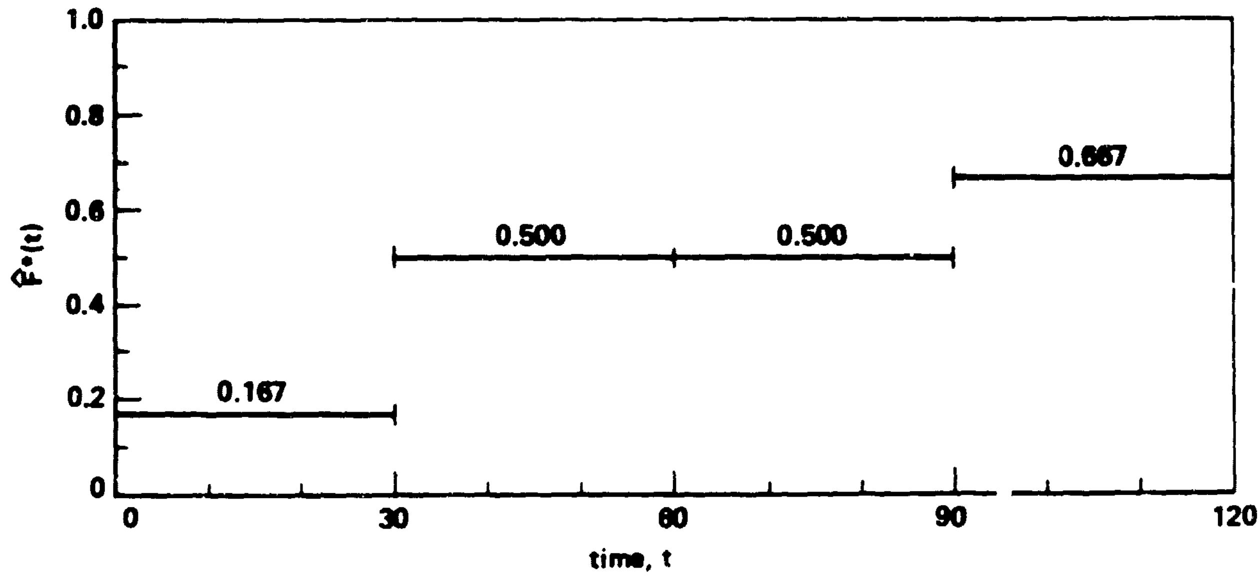


Figure 5. Probability of Failure vs. Time, Example II

It is clear that one obtains different plots of $\hat{F}^*(t)$ depending on the selection procedure for the time intervals. If the number of failure times is small, then one should use the failure times to determine the intervals $(\mu_{j-1,j})$. With this procedure, it is still possible to monitor the process through hypothesis testing, so nothing is lost in terms of one's ability to monitor.

Example III: A Ranking Approach

If one uses the individual times of failures, individual times of losses, pools the time information, and calculates the rank of the failure time in the combined sample, then (assuming there are no late entries) p_j may be estimated by

$$\hat{p}_j = \frac{N - R(t_j)}{N - R(t_j) + 1} \quad (7)$$

where $R(t_j)$ is the rank of the j^{th} failure time in the combined sample. We illustrate this technique by the following example.

In the previous example, with the entry time 80 omitted, we have in Figure 6,

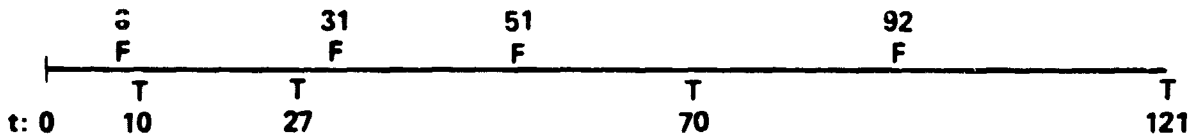


Figure 6. Time Line of Events, Example III

with (8, 31, 54, 92) the failure times. Consequently, the ranks of the failure times are

$$R(8) = 1, R(31) = 4, R(51) = 5, R(92) = 7.$$

The ranks and formula (7) give

$$\hat{p}_1 = 7/8, \hat{p}_2 = 4/5, \hat{p}_3 = 3/4, \hat{p}_4 = 1/2.$$

$\hat{p}_1$, $\hat{p}_2$, $\hat{p}_3$ correspond to the first three $\hat{p}_j$'s in the original example before a late entry time takes place. If late entries occur, we readjust formula (7) to allow for this. The readjusted formula is thus

$$\hat{p}_j = \frac{\{N + \Sigma y(t)\} - R(t_j)}{\{N + \Sigma y(t)\} - R(t_j) + 1} \tag{8}$$

where the ranking procedure is carried out in the same manner, but the sample size N is adjusted if and only if and where a late entry is made.

In the previous example, a late entry is made at t = 80. Thus, we adjust N by adding y (80) = 1 (one late entry) to N for the calculation that involves the next rank and each rank thereafter. This yields

$\hat{p}_1 = 7/8$

$\hat{p}_2 = 4/5$

$\hat{p}_3 = 3/4$

$$\hat{p}_4 = \frac{(N + 1) - 7}{(N + 1) - 7 + 1} = \frac{8 + 1 - 7}{8 + 1 - 7 + 1} = \frac{2}{3}.$$

We note that these estimates are now the same as the original estimates in Example I.

SECTION 2. CONSTRUCTION OF CONFIDENCE INTERVALS

With the estimation procedures clearly defined and illustrated, we turn to the construction of confidence intervals and to the question of monitoring a process. To construct confidence intervals for $F^*(t)$, we use the fact that

$$\frac{\hat{P}(t) - P(t)}{\hat{P}(t)\sqrt{\hat{V}(t)}} \sim N(0, 1). \tag{9}$$

That is, the ratio is normally distributed with a mean of zero and a standard deviation of one; $\hat{P}(t)$ is given by equations (5) or by using equations (7) or (8). The variance, $\hat{V}(t)$, is given as

$$\hat{V}(t) = \sum_{j=1}^{t} \frac{1 - \hat{p}_{j-1,j}}{n_{j-1}(\hat{p}_{j-1,j})} \tag{10}$$

Thus, from (9) and (10) the confidence interval for $P(t)$ (the unknown but real product estimate) at any point t in time is

$$\hat{P}(t) - Z_{\alpha/2}\,\hat{P}(t)\sqrt{\hat{V}(t)} \leq P(t) \leq P(t) + Z_{\alpha/2}\,\hat{P}(t)\sqrt{\hat{V}(t)} \tag{11}$$

or

$$P(t) \text{ lies between } \hat{P}(t)\left[1 \pm Z_{\alpha/2}\sqrt{\hat{V}(t)}\right].$$

This allows one to calculate bounds around any step for $F^*(t)$. Note $\hat{F}^*(t) = 1 - \hat{P}(t)$ or $1 - \hat{F}^*(t) = \hat{P}(t)$.

Since $P(t) = 1 - F^*(t)$ and $\hat{P}(t) = 1 - \hat{F}^*(t)$, we have

$$F(t) = 1 + (F^*(t) - 1)\left[1 \pm Z_{\alpha/2}\,(F^*(t) - 1)\sqrt{\hat{V}(t)}\right] \tag{12}$$

As an alternate derivation, using P_L and P_U to designate lower and upper bounds respectively, it can be shown that

$$1 - P_U \leq F(t) \leq 1 - P_L \tag{13}$$

where

$$P_U = \hat{P}(t)\left[1 + Z_{\alpha/2}\sqrt{\hat{V}(t)}\right]$$

and

$$P_L = P(t)\left[1 - Z_{\alpha/2}\sqrt{\hat{V}(t)}\right].$$

SECTION 3. COMPARISON OF TWO PROCESSES

To compare the two groups, we let

$$\hat{\theta} = \hat{P}_2/\hat{P}_1 \tag{14}$$

where $\hat{P}_2$ and $\hat{P}_1$ are constructed as previously shown. If θ is equal to 1, then the two processes are the same. Stated in terms of hypothesis testing, we have

$$H_0: \theta = 1, \; H_a: \theta > 1$$

where it is understood that $\hat{P}_2$ will represent the group that has the greater $\hat{P}$. The distribution is

$$\frac{(\hat{P}_2 - \theta \hat{P}_1)}{\{\hat{P}_2^2 \hat{V}_2 + \theta^2 \hat{P}_1^2 \hat{V}_1\}^{1/2}} \sim N(0, 1).$$

This implies that

$$\Pr(O \leq |\hat{P}_2 - \theta \hat{P}_1| \leq Z_{\alpha/2} \{\hat{P}_2^2 \hat{V}_2 + \theta^2 \hat{P}_1^2 \hat{V}_1\}^{1/2}) = 1 - \alpha. \tag{15}$$

To find confidence bounds for θ, we consider the inequality inside the probability statement. We thus have

$$\hat{P}_2^2 - 2\hat{P}_2 \hat{P}_1 \theta + \theta^2 \hat{P}_1^2 \leq Z_{\alpha/2}^2 \{\hat{P}_2^2 \hat{V}_2 + \theta^2 \hat{P}_1^2 \hat{V}_1\} \tag{16}$$

and, representing the inequality as an equality, we have

$$(\hat{P}_1^2 - Z_{\alpha/2}^2 \hat{P}_1^2 \hat{V}_1)\theta^2 - (2\hat{P}_2\hat{P}_1)\theta + (\hat{P}_2^2 - Z_{\alpha/2}^2 \hat{P}_2^2 \hat{V}_2) = 0. \tag{17}$$

Solving the quadratic equation in (17) yields

$$\theta = \frac{2\hat{P}_2\hat{P}_1 \pm \sqrt{4\hat{P}_2^2 \hat{P}_1^2 - 4(\hat{P}_1^2 - Z_{\alpha/2}^2 \hat{P}_1^2 \hat{V}_1)(\hat{P}_2^2 - Z_{\alpha/2}^2 \hat{P}_2^2 \hat{V}_2)}}{2(\hat{P}_1^2 - Z_{\alpha/2}^2 \hat{P}_1^2 \hat{V}_1)}. \tag{18}$$

Thus, the confidence intervals for θ are

$$\frac{2\hat{P}_2\hat{P}_1 - \sqrt{}}{2(\hat{P}_1^2 - Z_{\alpha/2}^2 \hat{P}_1^2 \hat{V}_1)} \leq \theta \leq \frac{2\hat{P}_2\hat{P}_1 + \sqrt{}}{2(\hat{P}_1^2 - Z_{\alpha/2}^2 \hat{P}_1^2 \hat{V}_1)} \tag{19}$$

where

$\sqrt{}$ is the term under the square root symbol in equation (18)

$\hat{P}_1$ is calculated from equation (5) for the first group

$\hat{P}_2$ is calculated from equation (5) for the second group

$\hat{V}_1$ is calculated from equation (10) for the first group

$\hat{V}_2$ is calculated from equation (10) for the second group

$Z_{\alpha/2}$ is the upper $\alpha/2$ value (e.g., $\alpha = 0.05$, $Z_{\alpha/2} = 1.96$).

An alternate way to write the equation is

$$\theta = \hat{\theta}\left[\frac{1 \pm \{1 - (1 - Z_{\alpha/2}{}^2\,\hat{V}_1)(1 - Z_{\alpha/2}{}^2\,\hat{V}_2\}^{1/2}}{1 - Z_{\alpha/2}{}^2\,\hat{V}_1}\right]. \tag{20}$$

If $L_L \leq \theta \leq L_U$ and $1 < L_L$ where L_L is the lower limit and L_U is the upper limit, then $P_2 > P_1$ at that particular point in time.

Stated another way, if the confidence interval for θ does not include one for a point in time, t, then $P_2(t) > P_1(t)$ and the two processes are different.

SECTION 4. CALCULATION OF RISK FOR INCORRECT SCREENING

It is often necessary to make a decision as to whether the individual or groups of components have progressed through the test screen with failure rate characteristics similar to components that have passed through this test screen and have performed successfully in later phases. In order to do this, we utilize the concepts of (1) controlling the difficulty of a screen and (2) the risk involved when we do not let a group of components pass the screen.

Having a difficult screen relates to the α measurement or the probability of a Type I error in statistical hypothesis testing. To make the screen more difficult, we decrease the confidence bands around the step function and consequently increase α. The larger α becomes, the tighter the screen. To assess the tightness of the screen, one usually asks a question such as, "If one makes the decision that the components passed the screen, and thus no corrective action is taken, what is the consequence of this action if trouble develops later in future testing or flight programs?" This assessment is usually made on weighted cost or subjective engineering judgment. In the case of space flight, it is thought best to have a tight screen, on the adage that "it is better to catch it and not have it happen, than it is to not catch it and have it happen." We adopt this philosophy in this screening process. On the other hand, if we do not let a group of components pass the screen, we encounter other costs such as those arising from redesign and retest efforts. If one can estimate the costs involved with the consequences of the decision to let a group of components pass the screen or not pass the screen, then it is possible to look at expected cost tradeoffs. To quantify the concepts, we use the notions of Type I and Type II errors.

We first take as a null hypothesis the belief that the two groups perform similarly. In the process, two incorrect decisions can occur. If we say that the group fails, i.e., it is unlike the control group when in fact it is similar, then we are making an incorrect decision. Call this an incorrect decision, D_{1_i}, of Type I.

On the other hand, if we say a group passes when in fact it is unlike the control group, we are again making an incorrect decision. Call this an incorrect decision, D_{2_i}, of Type II.

To quantify the former idea, let

$$\alpha = \Pr(D_{1_i}),\ \beta = \Pr(D_{2_i}) \tag{21}$$

If we use the concept, "It is better to catch it and not have it happen, than it is to not catch it and have it happen," then we want α to be large enough to guarantee a rigid screen. A corresponding calculation of β is made when one chooses to not pass the test group. If the costs of decision D_1 and D_2 are respectively C_1 and C_2, then the expected costs for these decisions are

$$E(D_1) = \alpha C_1 \text{ and } E(D_2) = \beta C_2 \tag{22}$$

To monitor the process of component level testing, we operate under the null hypothesis H_0: $\theta = 1$ for the statistic

$$\hat{P}_2 - \theta \hat{P}_1 \sim N(O, \hat{P}_2^2 \hat{V}_2 + \theta^2 \hat{P}_1^2 \hat{V}_1)$$

i.e., the statistic is normally distributed with a mean of zero and a standard deviation of $(\hat{P}_2^2 \hat{V}_2 + \theta^2 \hat{P}_1^2 \hat{V}_1)^{1/2}$.

The screening process may be viewed in the following manner. If the control group and the test groups are the same, then $\hat{P}_1 = \hat{P}_2$. Since $\hat{\theta} = \hat{P}_2/\hat{P}_1$, this implies that $\theta = 1$ when the groups are alike.

If we construct confidence intervals on θ from the test data and these confidence intervals do not cover $\theta = 1$, then we reject the null hypothesis, H_0: $\theta = 1$, and the group does not pass the control test.

From a program to calculate product limit estimates for a control group and a test group, as shown in Appendix A, we obtained $\hat{P}_2 = 0.8759$ and $\hat{P}_1 = 0.9482$ with V_2 (the variance of P_2) = 0.00110 and V_1 (the variance of P_1) = 0.00028.

Confidence intervals for θ at $\alpha = 0.20$ were (0.8803, 0.9680). Since these confidence intervals do not cover $\theta = 1$, we reject H_0: $\theta = 1$. To calculate β, we do the following:

(a) Calculate a confidence interval about θ for H_0: $\theta = 1$, $\alpha = 0.20$. This confidence interval corresponds to the Z acceptance region, $-1.28 \leq Z \leq +1.28$ where

$$Z = \frac{\hat{P}_2 - \theta \hat{P}_1}{\{\hat{P}_2{}^2 \hat{V}_2 + \theta^2 \hat{P}_1{}^2 \hat{V}_1\}^{1/2}} \tag{23}$$

or

$$Z = (1 - \theta/\hat{\theta}) \{\hat{V}_2 + (\theta/\hat{\theta})^2 \hat{V}_1\}^{1/2}. \tag{24}$$

For $\alpha = 0.20$ ($Z = 1.28$), we have the following inequality

$$-1.28 \leq \frac{\hat{P}_2 - \theta \hat{P}_1}{\{\hat{P}_2{}^2 \hat{V}_2 + \theta^2 \hat{P}_1{}^2 \hat{V}_1\}^{1/2}} \leq 1.28. \tag{25}$$

Letting

$$C = \{\hat{P}_2{}^2 \hat{V}_2 + \theta^2 \hat{P}_1{}^2 \hat{V}_1\}^{1/2}$$

and, solving for $\hat{P}_2/\hat{P}_1$, we have

$$\theta - \frac{1.28C}{\hat{P}_1} \leq \hat{P}_2/\hat{P}_1 \leq \theta + \frac{1.28C}{\hat{P}_1}. \tag{26}$$

For the null hypothesis, H_0: $\theta = 1$, equation (26) reduces to

$$1 - \frac{1.28C}{\hat{P}_1} \leq \theta \leq 1 + \frac{1.28C}{\hat{P}_1}. \tag{27}$$

(b) Substitute for $\hat{\theta}$ a value that depends on Z, C, and the new belief about θ. Since $Z = (\hat{\theta} - \theta_N)/C$,

$$\hat{\theta} = ZC + \theta_N \tag{28}$$

where θ_N is a new estimate which equals $\hat{P}_2/\hat{P}_1$.

Substituting for $\hat{\theta}$ in equation (27) and making a probability statement, we have

$$\beta = \Pr\left[\left(1 - \frac{1.28C}{\hat{P}_1}\right) \leq ZC + \theta_N \leq \left(1 + \frac{1.28C}{\hat{P}_1}\right)\right]$$

or

$$\beta = \Pr\left[\left(1 - \theta_N - \frac{1.28C}{\hat{P}_1}\right)\frac{1}{C} \leq Z \leq \left(1 - \theta_N + \frac{1.28C}{\hat{P}_1}\right)\frac{1}{C}\right] \tag{29}$$

As an illustration of the calculation of β, we have from our calculations shown in Appendix A

$\hat{P}_1 = 0.9482$, $\hat{V}_1 = 0.00028$ **and**

$\hat{P}_2 = 0.8759$, $\hat{V}_2 = 0.00110$.

Then, from the definition of θ_N as in equation (28),

$$\theta_N = \frac{0.8759}{0.9482} = 0.9238.$$

As before, we let

$$C = \left\{(0.8759)^2\,(0.00110) + (0.9482)^2\,(0.00028)\right\}^{1/2}$$

$$= 0.0331$$

and from equation (29),

$$\beta = \Pr\left\{\left[(1 - 0.9238) - \frac{1.28\,(0.0331)}{0.9482}\right]\frac{1}{0.0331} \leq Z \leq \left[(1 - 0.9238 + \frac{1.28\,(0.0331)}{0.9482}\right]\frac{1}{0.0331}\right\}$$

$$= \Pr\left\{0.9581 \leq Z \leq 3.6580\right\}.$$

From the definition of the right-hand term,

$$\beta = \frac{1}{\sqrt{2\pi}}\int_{0.9581}^{3.6580} e^{-1/2\,x^2}\,dx = 0.169.$$

SECTION 5. CONCLUSIONS

Using the PL process, one can monitor failure characteristics during component and system level testing and also during the orbital mission. This process enables one to make confidence statements about the failure characteristic of a test group and to compare the characteristic with that of a control group. This can be done graphically with statistical hypothesis testing.

One can calculate risks and have quantitative guidelines for making decisions as to whether to continue testing or not.

Using the step function to represent $F^*(t)$, one can obtain regression estimates of a Duane type mode to project failure characteristics from component level testing to system level and on to the orbital mission. The process gives the same form of an estimation procedure at each stage of testing.

The process converges in expected mean square to the theoretical distribution. This means that estimates of $F^*(t)$ based on this process will be a good representation of the failure characteristic.

REFERENCES

1. Kaplan, E. L. and Meier, P., "Nonparametric Estimation from Incomplete Observations," Journal of the American Statistical Association, 53 (June 1958), 457–81.

2. Thomas, D. R. and Grunkemeier, G. L., "Confidence Interval Estimation of Survival Probabilities for Censored Data," Journal of the American Statistical Association, 70 (December 1975), 865–871.

3. Turnbull, B. W., "Nonparametric Estimation of a Survivorship Function with Doubly Censored Data," Journal of the American Statistical Association, 69 (March 1974), 169–173.

4. Kruger, R. and Norris, H. P., "A Thermal Vacuum Test Optimization Procedure," NASA Technical Memorandum 80297, June 1979.

5. Duane, J. T., "Learning Curve Approach to Reliability Monitoring," IEEE Transactions on Aerospace, Volume 2, No. 2, April 1964, 563-566.

APPENDIX A

EXAMPLE OF ANALYSIS USING ACTUAL DATA

The data used in this example come from two entirely different groups of spacecraft components. The first group, noted as record numbers 1 to 931, is taken at random from a group of tests conducted in the period 1966 to 1976. The second, records 941 to 1631, is data from a group of components for one spacecraft tested in 1976 to 1978. Ten record numbers are used to maintain the data for one test so 94 tests are included in the first group and 60 tests in the second. Any test may contain one or more components.

Figure A-1 is a simplified flow diagram for the computer program used to reduce the data. The program itself is written in a modified BASIC language for a Hewlett-Packard 9831A desktop calculator. Block 1 of Figure A-1 basically performs program housekeeping functions. Blocks 2, 3, and 4 select the data to be analyzed and perform functions necessary to put the data into a form compatible with the statistical analysis requirements. The value for Z entered in block 6 determines the stringency of the comparison between the two groups of data. The comparison is done in block 7; if the upper and lower limits for θ do not include 1.0 as in equation (19), then the hypothesis that the two groups are alike is rejected and the acceptance region for the alternate hypothesis is computed.

Figure A-2 is a copy of the print-out of the data as adjusted and entered. The note at the top signifies that a 24 hour interval will be used in the analysis. The first case, as noted previously, includes records 1 through 931. The number of components actually listed in the data is 187. However, tests exist in the data wherein more than one failure per component exists. This condition is impossible in the Kaplan-Meier (1) approach and so the data was adjusted to increase the number of components so that at least one component was on test during each test period. The problem, in part, exists because of the uncertainty in designating a component; a complex item is normally designated as being made up of more than one component. However,

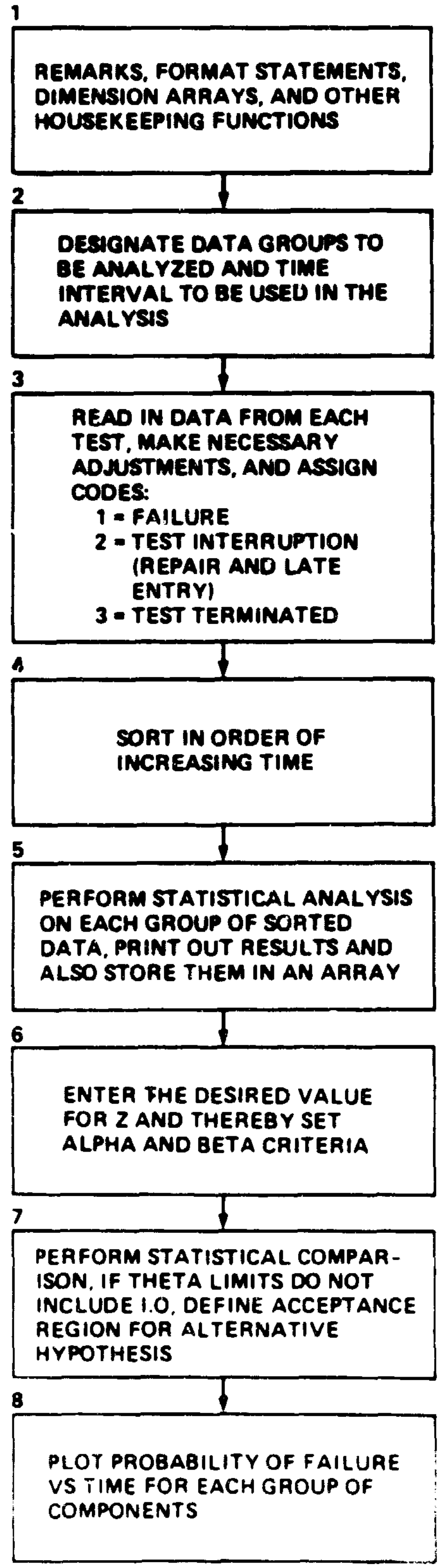


Figure A-1. Simplified Flow Diagram for Program to Reduce Data

TIME INCREMENT SET AT 24

1st CASE: BEGIN RECORD 1 , END RECORD 931

INITIAL COMPONENT COUNT: 187 , ADJUSTED COUNT: 198

ROW	HOUR	CODE	#COMP	ROW	HOUR	CODE	#COMP
1	109	3	4	86	232	3	2
2	98	3	1	87	26	3	15
3	48	3	2	88	26	1	1
4	182	2	0	89	53	3	1
5	296	3	4	90	57	3	2
6	192	1	1	91	60	3	1
7	212	2	-1	92	83	2	0
8	212	1	1	93	126	3	1
9	230	3	4	94	26	3	1
10	149	3	1	95	26	3	1
11	84	1	1	96	53	3	1
12	162	1	1	97	54	3	1
13	231	3	2	98	135	2	0
14	45	2	0	99	207	3	6
15	135	3	2	100	33	1	1
16	216	3	4	101	41	2	-1
17	37	[illegible]	1	102	92	[illegible]	[illegible]

Figure A-2. Data as Entered by Test

no exact definition exists. A better unit might be a part (e.g., a resistor or a capacitor), a device whose function is irreparably destroyed by disassembly. The large number of parts per component essentially obviates the possibility of having more failures than parts and eliminates the need for the adjustment in count. The adjusted component count is shown as 198.

The data itself is stored, test by test, in an array of 450 rows by 3 columns. Only rows having data in them are presented, and the output splits the data in half listing both halves simultaneously; therefore we have row 1 alongside row 86, there being 169 rows of data.

The data can be interpreted as follows:

Row 1: A test lasting 109 hours and then terminating (code 3) with the 4 components being taken off test.

Row 2: A test lasting 98 hours and then terminating with 1 component being taken off test.

Row 3: A test lasting 48 hours and then terminating with 2 components being taken off test.

Row 4: A test interrupted (code 2) at 182 hours. No failures occurred before the interruption and so no repairs or late entries are involved; #COMP = 0.

Row 5: The test in which the interruption in row 4 occurred is terminated at 296 hours and the 4 components that were on test are removed.

Row 6: A test in which a failure (code 1) occurs at 192 hours. One failure is noted under #COMP.

Row 7: The test is interrupted (code 2) at 212 hours and previous failures are repaired. The -1 designation is a late entry of the repaired component. The time spent during the interruption is subtracted from the in-test time.

Row 8: A failure occurs as soon as the test is begun again.

Row 9: The test continues until time 230 when it is terminated and the remaining operating item, 4, is taken off test.

It can be seen that the code 3 designators separate individual tests.

After all the data from a group has been entered into the array, the array is sorted in order of increasing time (column 1) and the data can be output as shown in Figure A-3. The information in rows 2 and 11 in Figure A-2 now appear in rows 102 and 86, respectively, in Figure A-3. Some, such as row 89 in Figure A-2 cannot be exactly identified since both rows 94 and 95 in Figure A-3 contain the same data. This is of no consequence and simply indicates that some other test (not shown) contained data having the same values.

Figure A-4 shows the statistical analysis of the data of the first case. Column designations are compatible with the terminology of this report. Note that the analysis is terminated when fewer than 20 components are on test. This was done because of the rapid increase in variance.

ROW	HOUR	CODE	#COMP	ROW	HOUR	CODE	#COMP
1	0	1	1	86	84	1	1
2	2	1	1	87	86	1	1
3	2	2	-1	88	87	3	1
4	2	1	1	89	89	3	1
5	3	1	1	90	90	3	2
6	3	1	1	91	91	2	-1
7	4	1	1	92	92	3	1
8	5	1	1	93	92	3	2
9	5	1	1	94	93	3	1
10	5	2	-1	95	93	3	1
11	7	1	1	96	94	3	1
12	8	1	1	97	94	3	1
13	12	3	1	98	94	3	1
14	13	3	1	99	95	3	1
15	14	3	1	100	95	3	1
16	17	3	2	101	97	3	1
17	19	2	0	102	98	3	1
18	20	3	1	103	98	3	2
19	20	3	1	104	100	3	1
20	20	3	1	105	107	3	5
21	23	2	0	106	109	3	4
22	23	2	-1	107	[illegible]	1	1
23	[illegible]	[illegible]	0	[illegible]	[illegible]	[illegible]	[illegible]

Figure A-3. Sorted Data

MAXIMUM TIME IN DATA: 459

Interval From	To	n(j-1)	d(j-1,j)	r(j-1,j)	phat(j-1,j)	Phat*	Variance
0	24	198	10	5	0.9482	0.9482	0.00028
24	48	183	8	26	0.9490	0.8999	0.00061
48	72	149	4	20	0.9690	0.8720	0.00099
72	96	125	4	14	0.9640	0.8405	0.00152
96	120	107	1	18	0.9888	0.8311	0.00190
120	144	88	2	13	0.9733	0.8089	0.00268
144	168	73	1	6	0.9851	0.7969	0.00349
168	192	66	2	6	0.9667	0.7703	0.00452
192	216	58	1	19	0.9744	0.7535	0.00573
216	240	38	0	12	1.0000	0.7505	0.00875
240	264	26	0	3	1.0000	0.7505	0.01278
264	288	23	0	8	1.0000	0.7505	0.01445

NUMBER OF COMPONENTS FALLS BELOW 30 AFTER TIME 288

Figure A-4. Analysis of First Case Data

The data for the second case, records 941 to 1631, were operated on the same way as that of the first case and the statistical analysis is shown in Figure A-5.

Finally, the two groups of data are compared. A value of Z = 1.28 was selected and the comparison can be seen in Figure A-6.

From Figure A-6, we can see that the null hypothesis (the two groups are alike) can be rejected in the period 0 to 24 hours. With the acceptance region of 0.9581 to 3.6580 then, there exists a 16.9 percent probability that the observed value will fall in the non-rejection region when the alternative hypothesis is true.

Figure A-7 graphically compares the two groups of components. While differences are apparent, the analysis as in Figure A-6 is needed to attribute confidence interval statements as to the degree of similarity.

MAXIMUM TIME IN DATA: 647

Interval From	To	n(j-1)	d(j-1,j)	y(j-1,j)	phat(j-1,j)	Phat*	Variance
0	24	129	17	-8	0.8759	0.8759	0.00110
24	48	120	3	7	0.9735	0.8527	0.00144
48	72	110	4	6	0.9615	0.8199	0.00200
72	96	100	1	2	0.9898	0.8115	0.00232
96	120	97	2	13	0.9762	0.7922	0.00270
120	144	82	3	-2	0.9643	0.7639	0.00377
144	168	81	1	-1	0.9878	0.7546	0.00402
168	192	81	1	-7	0.9886	0.7460	0.00420
192	216	87	0	0	1.0000	0.7460	0.00391
216	240	87	0	0	1.0000	0.7460	0.00391
240	264	87	0	8	1.0000	0.7460	0.00391
264	288	79	1	0	0.9873	0.7365	0.00453
288	312	78	0	10	1.0000	0.7365	0.00459
312	336	68	1	8	0.9833	0.7243	0.00560
336	360	59	2	11	0.9583	0.6941	0.00747
360	384	46	0	19	1.0000	0.6941	0.00958
384	408	27	1	9	0.9444	0.6555	0.01946

NUMBER OF COMPONENTS FALLS BELOW 20 AFTER TIME 408

Figure A-5. Analysis of Second Case Data

Z alpha/2 = 1.28

From	To	<=	Theta	<=	Acceptance Region From	To	%
0	24	0.8803		0.9680	0.9581	3.6580	16.9
24	48	0.9673		1.0862	-	-	-
48	72	0.9245		1.0634	-	-	-
72	96	0.9478		1.1109	-	-	-
96	120	0.9044		1.0763	-	-	-
120	144	0.8929		1.0973	-	-	-
144	168	0.8969		1.1202	-	-	-
168	192	0.9074		1.1533	-	-	-
192	216	0.9061		1.1660	-	-	-
216	240	0.8687		1.1603	-	-	-
240	264	0.8529		1.1899	-	-	-
264	288	0.8335		1.1891	-	-	-

Figure A-6. Comparison of the Two Groups of Data

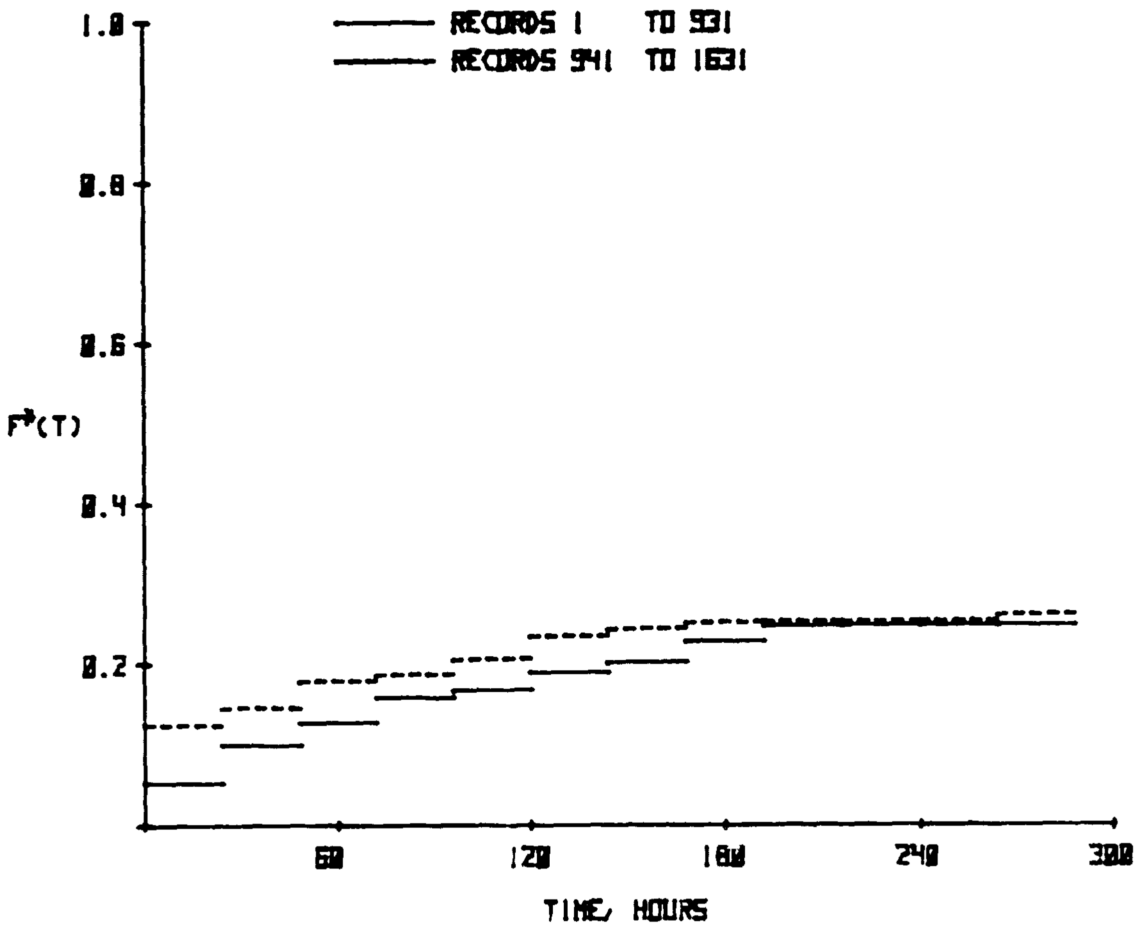


Figure A-7. Graphical Comparison of Failure Characteristics of the Two Groups of Components