# N O T I C E

THIS DOCUMENT HAS BEEN REPRODUCED FROM
MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT
CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED
IN THE INTEREST OF MAKING AVAILABLE AS MUCH
INFORMATION AS POSSIBLE

NASA CR—160090

# SPACECRAFT UTILITY AND THE DEVELOPMENT OF CONFIDENCE INTERVALS FOR CRITICALITY OF ANOMALIES

Dr. Raymond E. Williams
Ft. Lewis College,
Durango, Colorado

November 1980
Interim Report for Period January-October 1980

Prepared for

**GODDARD SPACE FLIGHT CENTER**
Greenbelt, Maryland 20771

# SPACECRAFT UTILITY AND THE DEVELOPMENT OF

# CONFIDENCE INTERVALS FOR CRITICALITY OF ANOMALIES

By

Dr. Raymond E. Williams

Ft. Lewis College, Durango, Colorado

November 1980

Interim Jan. – Oct. 1980

GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland

# CONTENTS

## Introduction

Bloomquist. [1], [2], developed the concept of availability,* which is a measure of the overall usefulness of a spacecraft or spacecraft payload. Kruger and Norris [3] extended this idea and used it in determining a thermal-vacuum test optimization procedure. Their optimization procedure makes use of a "lost value" function which is based on average availability* and other test costs. The optimization procedure is an interactive computer program in which the user supplies certain parameters as inputs. These parameters include payload information such as the number of components in the payload from which payload weight and payload cost can be determined, length of the mission in days, the type of item (protoflight, first flight item, or follow-on), whether an expendable launch vehicle or Shuttle (STS) is involved, if the latter, whether it is a Free-Flier or a Spacelab mission, the minimum and maximum test temperatures, and various other data.

One of the ideas central to the above optimization procedure is that of average utility which is calculated from instantaneous utility. Kruger and Norris [3] extended the concept of availability defined in Bloomquist [1] to make it dependent on the number of components in payload, mission duration, etc.

Utility is thought of as the successfulness of a mission as compared to a perfect mission. A perfect mission is one in which no failures or anomalies occur. (The words "failure" and "anomaly" are synonomous in the context of this report.) As a sequence of failures occur, the mission begins to degrade according to the seriousness or criticality of each failure. Utility is calculated from the observed occurrence of a random sequence of anomalies or types of failures by assigning a certain criticality to each of the failures in the sequence and then considering utility of the

---

*While Bloomquist [1], [2] and Kruger and Norris [3] have used the term "availability", that word in this report will be changed to "utility" because of the general usage of the term "availability" to mean other things in the field of reliability.

1

payload or spacecraft after any failure as the product of one minus the criticality term at the particular failure multiplied by the previous remaining utility. The instantaneous utility, u is defined as follows:

$$U = \prod_{i\in\Omega} (1 - D_i)^{\eta_i} \tag{1}$$

Where; $D_i$ denotes the criticality of a type i failure,

$\Omega$ is an index set for the various criticalities of failures that occur during space flight, and

$n_i$ is the total number of failures for any particular type of criticality during any space flight or any combined number of space flights.

In the above discussion, criticality is a predefined measurement of the percent of degradation caused by certain types of failures.

## Estimating Utility

In order to estimate U, it is first necessary to define the criticality various types of failures that can occur during any particular space flight. It is important to note that measures of criticality are subjective measurements and may vary from mission to mission and may vary due to various types of payloads such as those for scientific versus applications missions. Bloomquist [1] uses five classes of criticalities and has studied the effects of failures on the degradation of space missions for 304 spacecrafts. Timmins [4], studied 57 (GSFC) spacecraft and classified the criticalities according to four classes. The data in reference [4] consists of 449 malfunctions and failures and considered mission and component criticality with and without redundancy. Since Timmins' data involved GSFC spacecraft and GSFC thermal-vacuum test procedures based on these types of spacecraft, we use this data to define the various types of criticalities. This definition considers a malfunction and a failure as both being anomalies and we use the data based on redundancy. We summarize the data in Table 1.

In order to calculate U, we observe a sequence of failures or malfunctions, apply equation (1), and then calculate the utility at each stage of operation. This concept of utility is instantaneous in the sense that one has the overall utility at the instant the failure occurs.

Table 1

| Criticality Type | Criticality Classification | Range of the Percent Loss Due to Criticality | Percentage of Malfunction |
|---|---|---|---|
| Castastrophic | $D_1$ | 100 – 90 | $0.02 = Pr(D_1)$ |
| Major Loss | $D_2$ | 90 – 50 | $0.02 = Pr(D_2)$ |
| Substantial Loss | $D_3$ | 50 – 10 | $0.11 = Pr(D_3)$ |
| Minor Loss | $D_4$ | 10 – 0 | $0.85 = Pr(D_4)$ |

As an example, consider the following sequence of failures $(D_4, D_4, D_2, D_3)$ which is a subset of various criticality types $(D_1, D_2, D_3, D_4)$.

The utility after three consecutive failures of the failure sequence is $U = (1 - D_4)(1 - D_4)$ $(1 - D_2)$. In order to assign a numerical value to U, we must assign a percent loss due to each criticality. One possible choice is to assign the mid–points of the ranges of the percent losses due to the types of criticalities. Therefore, $D_4 = 0.05$, $D_3 = 0.30$, $D_2 = 0.70$, and $D_4 = 0.95$. With these D's, $U = (1 - 0.05)(1 - 0.05)(1 - 0.70) = (0.95)^2 (0.3) = 0.27075$ at the end of three failures.

It can be seen that the value of U depends upon a random sequence of failures. Thus, there is no absolute model with which to compare U (such as assuming the linear model in regression analysis) and then determining the degree of fit by comparing the sum of squares due to regression to the total sum of squares.

## Establishing a Criterion for Goodness of Fit

We next establish a criterion for making a comparison for goodness of fit for estimates of U. Assume that S is the underlying set for all measurement techniques for U determined by equation (1). For $S_i \in S$, $S_j \in S$, we decide that measurement technique $S_i$ is better than $S_j$ if $S_i$ outperforms $S_j$. We use as a performance criteria (I) performance at each point or mode of failure and (II) measurement of the spread. We specify (I) by saying $S_i$ outperforms $S_j$ at the nth failure, if

3

the absolute value of the residual for the estimate $U_i$ of U, is less than the absolute value of the residual for the estimate of $U_j$ of U. In terms of events, we have a success, a failure, or a tie where a success means $S_i$ outperforms $S_j$, a failure means $S_j$ outperforms $S_i$, and a tie indicates equality of performance. If we exclude ties or use a randomization procedure to reclassify ties, we have a binomial structure which can be used to measure the performance at each point or mode of failure. To specify a measurement of spread, we use the mean squared error (M.S.E.).

To decide between selection procedure $S_i$ over $S_j$ based on performance criteria I and II we establish the following decision table.

Table 2

| Condition | Decision |
|---|---|
| (if) I:   $S_i = S_j$ | |
| (and) II:   (a) M.S.E. $(S_i)$ = M.S.E. $(S_j)$ | (then) estimation procedures are the same |
|            (b) M.S.E. $(S_i)$ < M.S.E. $(S_j)$ | $S_i$ is a better estimation procedure than $S_j$ |
|            (c) M.S.E. $(S_j)$ < M.S.E. $(S_i)$ | $S_j$ is a better estimation procedure than $S_i$ |
| (if) I:   $S_i > S_j$ | |
| (and) II:   (a) M.S.E. $(S_i)$ = M.S.E. $(S_j)$ | (then) $S_i$ is a better estimation procedure than $S_j$ |
|            (b) M.S.E. $(S_i)$ >> M.S.E. $(S_j)$ | no conclusion |
|            (c) M.S.E. $(S_i)$ < M.S.E. $(S_j)$ | $S_i$ is a better estimation procedure than $S_j$ |

In the above discussion we replace i by j in step II when $S_j > S_i$ for step I.

As in Kruger and Norris [3], one obtains a estimate $\hat{U}$ of U by considering a measurement of overall criticality $D^*$ such that

$$(1 - D^*)^n = \hat{U} \tag{2}$$

where n $\equiv$ total number of failures.

By equating $\hat{U}$ to U, we have

4

$$(1 - D^*)^n = \prod_{i \in \Omega} (1 - D_i)^{n_i}. \tag{3}$$

Solving for $D^*$, one obtains

$$D^* = 1 - \exp\left\{\sum_{i \in \Omega}\left[\frac{n_i}{n} \ln(1 - D_i)\right]\right\} \tag{4}$$

where $D_i$ denotes the percentage loss for the $i^{th}$ type of criticality, and $\Omega$ is the index set of defined criticalities.

If one calculates $D^*$ for the set of all spacecraft, then the probability of the event $D_i$ is $Pr(D_i) \approx \frac{n_i}{n}$ where $n =$ the total number of failures and $n_i =$ number of failures of the $i^{th}$ type of anomaly. $(\ln(1 - D_i))$ is a weight of $Pr(D_i)$.

Since one is interested in a measurement of overall criticality, it is feasible to look at an average measurement or a class of estimates that measure D overall. Thus, one may calculate an estimate of overall criticality by using the classical definition of expected value. We thus have:

$$E(D) = \sum_{i \in \Omega} (D_i) \cdot (Pr(D_i)) \tag{5}$$

Since the definition of the percentage loss due to a particular type of criticality is subjective, we may obtain various estimates of $E(D)$. As an example of various values that $E(D)$ assumes, we consider the maximum, upper quartile, middle, lower quartile, and minimum points for the intervals that describe the percentage loss that is due to the various types of criticalities based on Table 1. We summarize this information in Table 3.

As a sample calculation, we have:

$$E_{(1)}(D) = (1.0)(0.02) + (0.9)(0.02) + (0.50)(0.11) + (0.10)(0.85) = 0.178,$$

and the other $E_{(i)}(D)$'s are calculated in a similar manner using equation (5), Table 2, and the percentage of malfunction, $P(D_i)$ given in Table 1.

Table 3
Intervals Describing Percentage Loss Due to Criticality

| Max | Upper Quartile | Middle | Lower Quartile | Minimum |
|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) |
| 1.0 | 0.975 | 0.95 | 0.925 | 0.90 |
| 0.9 | 0.80 | 0.70 | 0.60 | 0.50 |
| 0.50 | 0.40 | 0.30 | 0.20 | 0.10 |
| 0.10 | 0.075 | 0.05 | 0.025 | 0 |
| $E_{(1)}(D)$ | $E_{(2)}(D)$ | $E_{(3)}(D)$ | $E_{(4)}(D)$ | $E_{(5)}(D)$ |
| 0.178 | 0.143 | 0.109 | 0.078 | 0.039 |

using $P_1 = 0.02$, $P_2 = 0.02$, $P_3 = 0.11$, $P_1 = 0.85$

where $P_i = Pr(D_i)$.

As seen from Table 3, the use of subjective probability enables one to consider various bounds and ranges of values for $D_i$ and hence for measurements of $E(D)$. Possible applications of the table extend to determining the overall criticality $E(D)$ as a function of the number of components in a spacecraft. If it is believed or it can be shown that when the number of components that comprise a spacecraft or payload increases, the overall criticality decreases, then one may make a subjective judgment on the choice of $E(D)$, choosing the lower quartile measurement for a greater number of components and the upper quartile for fewer number of components. If one uses this concept for future space flight missions such as Space Shuttle, then it would be feasible to study and redefine the types of criticalities and the range of the percent loss due to criticalities. In fact, Tables 1 and 2 should continually be studied and upgraded as one changes missions or enlarges the data banks from new missions.

A comparison was conducted on data from 31 spacecraft from PRC data base that were common to the GSFC data base to see how $D^*$ performed as compared to $E(D)$. Of 284 failures,

there were 24 ties and of the remaining data, E(D) outperformed D* 169 times. Excluding ties, and considering the experiment as a binomial, we have the following hypothesis scheme:

$H_0$: The performance of D* and E(D) are the same.

$H_a$: The performance of D* and E(D) differ.

The test data yielded a standard deviation of 8.06 and a mean of 130. Comparing 169 to the mean of 130, we have a Z (the standard normal random variable) of 4.83 which indicates we reject the null hypothesis at an alpha level of less than 0.005. Furthermore, calculating the m an square error (M.S.E.) of both estimates, we find M.S.E.(D*) = 0.01826 and M.S.E.(E(D)) = 0.01171. Since M.S.E.(E(D)) < M.S.E.(D*), we conclude that E(D) outperforms D*, based on the available spacecraft data. We also conclude that it is feasible to use E(D) in estimation for U for GSFC data.

The following theorem gives a mathematical justification of the previous observed results.

Theorem: If X is a random variable, then the best way to measure X by a single constant, a, is to choose a = E(X). Here, best is defined as minimum mean square error.

Proof: Let $f(a) = \sum_{i=1}^{n} (X_i - a)^2$.

Taking derivatives, we have

$$f'(a) = 2 \sum_{i=1}^{n} (X_i - a).$$

Setting $f'(X) = 0$ and applying the summation implies $\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} a = 0$

Since

$$\sum_{i=1}^{n} a = na,$$

we have

$$a = \frac{1}{n} \sum_{i=1}^{n} X_i = E(X).$$

Since

$$f''(a) > 0, a = E(X) \text{ minimizes } f(a) \qquad \text{Q.E.D.}$$

7

The Kruger-Norris report [3] established a functional relationship between $D^*$ and the number of components given by equation (I-14) of [3] as $D^* = 0.237 \exp(-0.0086N)$ where N is defined as the number of components and $D^*$ is the measure of criticality. A similar binomial comparison was conducted on data from the same 31 spacecraft from PRC as mentioned above using the hypothesis that $D^*$ is a function of the number of components.

Making a similar binomial comparison as was done previously, $c(D)$ outperformed $D^*$ as given by equation (I-14) of [3] 215 times out of a possible 284 times. Performing the appropriate hypothesis test, as before, we reject the null hypothesis that $E(D)$ and $D^*$ have the same performance at an $\alpha$-level at or less than 0.005. The M.S.E. of $D^*$ is much larger than the M.S.E. of $E(D)$;

$$M.S.E.(E(D)) = 0.01171 < M.S.E.(D^*) = 1.044.$$

As shown by the theorem, the best way to predict a random variable X by a single value is to choose the predictor "a" as the expected value of the random variable. This means that $a = E(X)$ or that "a" is an appropriate measure of central tendency of the random variable X. The theorem and numerical results implies that for any values of the random variable $X$; $(X_1, X_2, \ldots, X_n)$, $\sum_{i=1}^{n} X_i/n$, is the best single predictor for the overall collection of values considered, both as a total collection or as a subcollection such as the 31 spacecraft in PRC data. Hence, if we consider the totality or any subcollection of spacecraft failures and classify them according to their criticalities, we would choose $E(D)$ as the overall single appropriate measure of a criticality value. This means that the instantaneous utility, U, is estimated as $\hat{U} = \prod (1 - E(D))^{n_i}$. Two questions remain about the estimation procedure, (1) by a dependence relationship such as given in equation (I-14), can one gain predictability and (2) what is the value of $E(D)$ used in the estimation procedure?

For any subcollection of spacecraft we have just shown that $E(D)$ gives better predictability then $D^*$, on the basis of closeness of the prediction and M.S.E. of the estimates.

For a complete answer to question (1), one must examine if on the basis of some casual de-
pendence such as found in equation (1-14) one gains predictability. To accomplish this, we con-
sider the collection of 41 spacecraft as a single group and compute E(D). After computing E(D),
we compare the performance of E(D) to the performance of D*, equation (1-14) for each
spacecraft.

The above study yielded the following results. D* outperformed E(D) 162 times out of a
total of 284. Using a similar test as before, we reject the null hypothesis that E(D) and D* have
the same performance at the $\alpha = 0.05$ level. It should be remarked that the M.S.E.(E(D)) >
M.S.E.(D*), since M.S.E.(E(D)) = 0.055 > M.S.E.(D*) = 0.044.

This result points to a complexity factor in determining overall criticality. Since critical-
ity is dependent on the number of components that make up a spacecraft, this dependence
can be used to gain predictability. This result also points to a need to ascertain a more refined
measure of the complexity of a spacecraft, since the measurement of complexity enters into the
formulation of utility which is a measurement essential to measuring the performance of
spacecraft and the amount of information a spacecraft gains on any particular mission.

In order to answer question (2), we turn to the development of confidence intervals for E(D).
When it is necessary to give a confidence statement about the amount of information one obtains
on any particular mission, it is necessary to bound E(D) or D* in the measurement for utility.
Since E(D) lends itself to this type of bound more readily than D*, but D* outperforms E(D), it
is necessary to adjust E(D) to make it perform as well as D*.

Another form of variability in the system defined by equation (5) comes from the distribu-
tion of the statistic $\sum_{i \in \Omega} D_i (P(D_i))$. We observe that $P(D_i) = \dfrac{Y_i}{n} = P_i$ where $Y_i$ is defined to be
the number of occurrences of the $i^{th}$ type of anomaly. The statistic $\sum_{i \in \Omega} D_i P(D_i)$ can be written
as

$$Y_n = \sum_{i=1}^{k} D_i \left( \frac{Y_i}{n} \right) \qquad (6)$$

where n designates the totality of failures for all spacecraft.

Thus $Y_n$ is a statistic that measures E(D) for a collection of spacecraft. We also observe that the expected value of $Y_n$ is, since $E(Y_i) = n\,P_i$,

$$E(Y_n) = \sum_{i=1}^{k} (D_i) \cdot \left( E \left( \frac{Y_i}{n} \right) \right) = \sum_{i=1}^{k} D_i\, P_i \qquad (7)$$

To distinguish between the role that n plays for the totality of failures for all spacecraft and the role $n_j$ plays for any single spacecraft j, we define $Y_{n_j}$ for a single spacecraft as

$$Y_{n_j} = \sum_{i=1}^{k} D_i \left( \frac{Y_{ij}}{n_j} \right) \qquad (8)$$

we further observe that

$$E(Y_{n_j}) = \sum_{i=1}^{k} D_i\, E \left( \frac{Y_{ij}}{n_j} \right) = \sum D_i\, P_{ij} \qquad (9)$$

where $P_{ij}$ is the information based only on the $j^{th}$ spacecraft.

Equations (7) and (9) imply that the expected value of the statistic $Y_{n_j}$ is different for different subclasses of spacecraft. It is noted that this is an a priori measurement based on the totality of the information that one has about all previous space flights; at present $P_1 = 0.02$, $P_2 = 0.02$, $P_3 = 0.11$, $P_4 = 0.85$ on 449 anomalies for 57 GSFC spacecraft and $D_i$'s are subjectively chosen. As information is gained and the nature of the mission better known, the $P_i$'s and the subjective definitions of the $D_i$'s should be re-evaluated.

Bounding the Statistic: Confidence Intervals

In order to place bounds on the statistic $\sum_i D_i\, P_i$, we outline two approaches. One approach is given in Briemann [5] and involves constructing simultaneous confidence intervals for

$P_i$'s from the individual confidence intervals for the $P_i$'s. This approach allows for dependence which we have in our system. The following definitions and theorems follow Briemann.

**Definition:** In a K-parameter problem, the intervals $J_1(\hat{\Theta}), \ldots, J_k(\hat{\Theta})$ are said to form simultaneous $100\,\gamma\,\%$ confidence intervals if

$$P_\theta [\theta_1 \epsilon J_1(\hat{\Theta}), \ldots, \theta_k \epsilon J_k(\hat{\Theta})] \geq \gamma \text{ for all } \theta \epsilon (\hat{\Theta}).$$

To obtain simultaneous $100\,\gamma\,\%$ intervals, we have the following proposition:

**Proposition:** If $J_1(\hat{\Theta}), \ldots, J_k(\hat{\Theta})$ individually are $100\,\gamma\,\%$ confidence intervals, then they form simultaneous $100\,[1 - K(1 - \gamma)]\,\%$ confidence intervals.

For the data given by Table 1, there are K = 4 criticalities and four $P_i$'s on which to construct confidence intervals. To obtain 90% simultaneous confidence intervals for $P_1, P_2, P_3, P_4$, we choose $Z_{\alpha/2} = 2.24$, $\gamma = 0.975$, $(\frac{\alpha}{2} = 0.0125)$. If we construct 95% confidence intervals for any single estimate $P_i$, $\gamma \approx 0.95$ and we obtain $(1 - K(1 - \gamma))\,\%$ or 80% simultaneous confidence intervals for $(P_1, P_2, P_3, P_4)$.

After simultaneous confidence intervals are obtained for $(P_1, P_2, P_3, P_4)$, we may obtain simultaneous confidence intervals for $E(D)$ and $E(D_j)$; where $E(D)$ denotes the expected criticality for cumulative spacecraft failure data and $E(D_j)$ denotes the expected criticality for any single space flight mission. The following are simultaneous a priori confidence intervals for $P_1, P_2, P_3, P_4$, where n denotes the totality of failures for all spacecraft $1 - P_i = Q_i$.

$$\hat{P}_i - Z_{\alpha/2} \sqrt{\frac{\hat{P}_i \hat{Q}_i}{n}} \leq P_i \leq \hat{P}_i + Z_{\alpha/2} \sqrt{\frac{\hat{P}_i \hat{Q}_i}{n}}; \tag{10}$$

where $\hat{P}_1 = 0.02$, $\hat{P}_2 = 0.02$, $\hat{P}_3 = 0.11$, $\hat{P}_4 = 0.85$, and $\alpha$ is determined as previously discussed.

Using the above confidence intervals for $P_i$ we have the following confidence intervals for $E(D)$ and $E(D_j)$.

$$\sum_i D_i \left\{ \hat{P}_i - Z_{\alpha/2} \sqrt{\frac{\hat{P}_i \cdot \hat{Q}_i}{n}} \right\} \leq E(D) \leq \sum_i D_i \left\{ \hat{P}_i + Z_{\alpha/2} \sqrt{\frac{\hat{P}_i \cdot \hat{Q}_i}{n}} \right\} \tag{11}$$

11

These relationships hold for any set or subset of spacecraft.

As an example, we calculate the 90% confidence intervals for $P_1$, $P_2$, $P_3$, $P_4$ using n = 400, 100, 10.

For n = 400:

$$0.02 - 2.24 \sqrt{\frac{(0.02)(0.98)}{400}} \leq P_1 \leq 0.02 + 2.24 \sqrt{\frac{(0.02)(0.98)}{400}}$$

$$0.02 - 0.016 \leq P_1 \leq 0.02 + 0.016$$

$$0.004 \leq P_1 \leq 0.036$$

$$0.004 \leq P_2 \leq 0.036$$

$$0.11 - 0.035 \leq P_3 \leq 0.11 + 0.035$$

$$0.075 \leq P_3 \leq 0.1435$$

$$0.85 - 0.04 \leq P_4 \leq 0.85 + 0.04$$

$$0.81 \leq P_4 \leq 0.89$$

For n = 100:

$$\max(0, 0.02 - 0.031) \leq P_1 \leq 0.02 + 0.031$$

$$0 \leq P_1 \leq 0.051$$

$$0 \leq P_2 \leq 0.051$$

$$0.11 - 0.070 \leq P_3 \leq 0.11 + 0.070$$

$$0.04 \leq P_3 \leq 0.18$$

$$0.85 - 0.08 \leq P_4 \leq 0.85 + 0.08$$

$$0.78 \leq P_4 \leq 0.93$$

For n = 10:

$$\max(0, 0.02 - 0.099) \leq P_1 \leq 0.02 + 0.099$$

$$0 \leq P_1 \leq 0.119$$

$$0 \leq P_2 \leq 0.119$$

$$\max(0, 0.11 - 0.22) \leq P_3 \leq 0.11 + 0.22$$

$$0 \leq P_j \leq 0.33$$

$$0.85 - 0.25 \leq P_4 \leq \min (1, 0.85 + 0.25)$$

$$0.60 \leq P_4 \leq 1$$

We next proceed to calculate confidence intervals for $E_{(j)}(D)$ for the given sample size n = 400 and for the various values of the range of D such as maximum, upper quartile, middle, lower quartile, and minimum. In this notation $E_{(1)}(D)$ denotes the maximum range of criticality of a failure, $E_{(2)}(D)$ denotes the upper quartile, $E_{(3)}(D)$ denotes the middle, $E_{(4)}(D)$ denotes the lower quartile, and $E_{(5)}(D)$ denotes the minimum. Thus, when we use the dependence between number of components and criticality of a failure, we establish a relationship to assign $E_{(j)}(D)$ to measure criticality as a function of the number of components. The relationship we use is the greater the number of components, the greater value of j we use. This relationship will be explored after another method to establish bounds for E(D) is discussed.

For n = 400;

$$0.178 - 0.052 \leq E_{(1)}(D) \leq 0.178 + 0.052$$

$$0.126 \leq E_{(1)}(D) \leq 0.230$$

$$0.143 - 0.0424 \leq E_{(2)}(D) \leq 0.143 + 0.0424$$

$$0.101 \leq E_{(2)}(D) \leq 0.187$$

$$0.109 - 0.039 \leq E_{(3)}(D) \leq 0.109 + 0.039$$

$$0.070 \leq E_{(3)}(D) \leq 0.148$$

$$0.078 - 0.032 \leq E_{(4)}(D) \leq 0.078 + 0.032$$

$$0.046 \leq E_{(4)}(D) \leq 0.110$$

$$0.039 - 0.026 \leq E_{(5)}(D) \leq 0.039 + 0.026$$

$$0.013 \leq E_{(5)}(D) \leq 0.065$$

One notes that the confidence bands for $E_{(3)}(D)$ actually contain all of the E(D)'s of the data from PRC's 31 spacecraft.*

---

*One may not bound $P_j$ by a number lower than zero, or bound $P_j$ above by a number greater than 1.

Another approach for confidence intervals involves using the multivariate normal distribution. This approach incorporates the variance of the statistic $\sum_i D_i P_i$ into the variance – covariance structure for the multivariate normal structure of the $P_i$'s. In order to develop confidence intervals for $\sum_i D_i P_i$, we give some preliminaries.

Definition:

Let

$$
V = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & & & \sigma_{1n} \\
\sigma_{12} & \sigma_{22} & \cdot & \cdot & \sigma_{2n} \\
\cdot & \cdot & & & \\
\cdot & \cdot & & & \\
\cdot & \cdot & & & \\
\sigma_{1n} & \sigma_{2n} & \cdot & \cdot & \sigma_{rn}
\end{bmatrix}
\tag{12}
$$

Where $\sigma_{ii} = \sigma_i^2$ which are the variance terms and $\sigma_{ij}$ are the covariance terms. The matrix $V$ is called the variance – covariance matrix.

From Mood and Graybill [6], $\hat{P}_i = \dfrac{Y_i}{n}$, $i = 1, \ldots, 5$, are the maximum likelihood estimates of $P_i$ which are used to define $E(D)$ and $E(D_j)$. It is well known that the joint distribution $f(\hat{P}_1, \hat{P}_2, \ldots, \hat{P})$ is approximately normal with means $P_i$. Lehmann [7], Theorem 14, appendix, states that a multivariate normal distribution is completely specified by its means and its covariance structure. We thus have

$$
\sigma_{ii}(\hat{P}_i) = \frac{\hat{P}_i \cdot \hat{Q}_i}{n}; \hat{Q}_i = 1 - \hat{P}_i
$$

and

$$
\sigma_{ij}(\hat{P}_i, \hat{P}_j) = \text{cov}(\hat{P}_i, \hat{P}_j) = - \frac{\hat{P}_i \cdot \hat{P}_j}{n}.
\tag{13}
$$

These are the exact variances and covariances for any sample size n. This means that we can calculate bounds for the statistic $\sum_i D_i P_i$ for the number of failures for all spacecraft and study the behavior of any particular spacecraft as a member of its family.

The following result from Hogg and Graig [8], will be used to establish confidence intervals for $\sum_i D_i P_i$. Let $X_1, X_2, \ldots, X_n$ have a multivariate normal distribution with matrix $\mu$ of means and positive definite variance - covariance matrix V. Consider a linear function Y of $X_1$, $\ldots, X_n$ which is defined by $Y = C^T \cdot X = \sum_{i=1}^{n} C_i \cdot X_i$, where $C^T = [C_1, \ldots, C_n]$.

<u>Theorem:</u>   The random variable Y is $N(C^T\mu, C^TVC)$. That is, it is normally distributed with a mean of $C^T\mu$ and a variance $C^TVC$.

To adapt this to our situation, let

$$\mu = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} \; ; \; X = \begin{bmatrix} \hat{P}_1 \\ \hat{P}_2 \\ \hat{P}_3 \\ \hat{P}_4 \end{bmatrix}$$

$$C^T = [D_1, D_2, D_3, D_4]_{(i)*}$$

$$V = \begin{bmatrix} \dfrac{\hat{P}_1 \hat{q}_1}{n} & \dfrac{-\hat{P}_1 \hat{P}_2}{n} & \dfrac{-\hat{P}_1 \hat{P}_3}{n} & \dfrac{-\hat{P}_1 \hat{P}_4}{n} \\[3mm] \dfrac{-\hat{P}_1 \hat{P}_2}{n} & \dfrac{\hat{P}_2 \hat{q}_2}{n} & \dfrac{-\hat{P}_2 \hat{P}_3}{n} & \dfrac{-\hat{P}_2 \hat{P}_4}{n} \\[3mm] \dfrac{-\hat{P}_1 \hat{P}_3}{n} & \dfrac{-\hat{P}_2 \hat{P}_3}{n} & \dfrac{\hat{P}_3 \hat{q}_3}{n} & \dfrac{-\hat{P}_3 \hat{P}_4}{n} \\[3mm] \dfrac{-\hat{P}_1 \hat{P}_4}{n} & \dfrac{-\hat{P}_2 \hat{P}_4}{n} & \dfrac{-\hat{P}_3 \hat{P}_4}{n} & \dfrac{\hat{P}_4 \hat{q}_4}{n} \end{bmatrix}$$

(Note: * denotes the level such as minimum, maximum, middle, etc..)

Thus, we have

$$Y = C^T \cdot X = [D_1, D_2, D_3, D_4] \begin{bmatrix} \hat{P}_1 \\ \hat{P}_2 \\ \hat{P}_3 \\ \hat{P}_4 \end{bmatrix} = \sum_i D_i \hat{P}_i.$$

Using the normality of Y, we have the following confidence intervals for $\sum_i D_i \cdot P_r(D_i)$ or $E(D)$.

$$C^T\mu - Z_{\alpha/2} \sqrt{C^T \cdot V \cdot C} \leq E(D) \leq C^T\mu + Z_{\alpha/2} \sqrt{C^T \cdot V \cdot C}$$

As an example of the overall calculation of a 90% confidence interval, for E(D), we consider n = 10 and (i) = middle.

We thus have

$$X = \begin{bmatrix} 0.02 \\ 0.02 \\ 0.11 \\ 0.85 \end{bmatrix}, C^T = [0.95, 0.70, 0.30, 0.05],$$

$$V = \begin{bmatrix} 0.00196 & -0.00004 & -0.00022 & -0.00170 \\ -0.00004 & 0.00196 & -0.00022 & -0.00170 \\ -0.00022 & -0.00022 & 0.00979 & -0.00935 \\ -0.00170 & -0.00170 & -0.00935 & 0.01275 \end{bmatrix} ; Z_{\alpha/2} = 1.65.$$

$$C^T X = 0.109$$

$$C^T V = (0.00168, 0.00118, 0.00211, -0.00497)$$

$$C^T V C = 0.00281$$

$$\sqrt{C^T V C} = 0.0530$$

Therefore

$$0.109 - 1.65 (0.0530) \leq E(D) \leq 0.109 + 1.65 (0.0530)$$

$$0.021 \leq E(D) \leq 0.197.$$

If n = 5, the entries in the matrix V are multiplied by 2 and hence

$$CV_{(5)} = [2] \ C^T V_{(10)} = C^T V_{(5)} \ C = [2] \ C^T V_{(10)} \ C = 0.00563$$

$$= \sqrt{C^T V_{(5)} \ C} = \sqrt{0.00563} = 0.0750$$

$$0.109 - 1.65 (0.0750) \leq E(D) \leq 0.109 + 1.65 (0.0750)$$

$$\text{min. } (0, 0.109 - 0.124)) \leq E(D) \leq 0.109 + 0.124$$

$$0 \leq E(D) \leq 0.233.$$

## Using E(D) to Estimate D Based on the Dependence Between Criticality and Number of Components

It has been shown that there is a dependence between the number of components (a measure of the complexity of the spacecraft) and the criticality of a failure. Using this dependence, one develops a functional relationship to better estimate D in the equation

$$U = \prod (1 - D)^n$$

where D is now a function of component count.

The same type of relationship may be established using $E_{(j)}(D)$ as used in the section of this report dealing with establishing a criterion for goodness of fit. The way this is done is to take the 31 PRC spacecraft and separate them into 5 categories, I, II, III, IV, V, depending on their component count. For consistency, we match the lowest component counts with category I and correspond this with $E_1(D)$, since the smaller the number of components the greater the criticality of any failure and a similar process for greater component count. Using this technique for the 31 PRC spacecraft, we have the following correspondences in terms of spacecraft component count, categories, and measures of criticality.

| Categories | I | II | III | IV | V |
|---|---|---|---|---|---|
| Component Count | 28, 39, 46, 46 | 52, 52, 52, 52, 52, 59 | 64, 70, 70, 70, 70, 76, 78 | 86, 86, 86, 86, 86, 86 | 120, 129, 129, 130, 130, 131, 131, 137 |
| Criticality of Failure | 0.178 | 0.143 | 0.109 | 0.078 | 0.039 |

After making these correspondences, we now compare the performances of E(D) and D = 0.204 $e^{-0.005247N}$. We obtain the following results: E(D) outperformed D = 0.204 $e^{-0.005247N}$ 143 times to 141 times out of 284 failures. The M.S.E.'s of the estimates were approximately the same, thus there is no significant difference in the performance of E(D) and D as defined above. We may use E(D) when we want to calculate confidence bounds for utility, U, or use D for estimation and cost purposes without loss of performance. Due to the changing character of spacecraft in terms of complexity and failure modes, it is advisable to continually study and refine these measurements.

## Procedure to Calculate Confidence Intervals for $U(t) = (1 - D^*)^{F(t)}$

To bound U as given by $U(t) = (1 - D^*)^{F(t)}$, where $D^*$ denotes the criticality of a failure and $F(t)$ denotes the number of failures after time t as in equation I-13 of reference 3, we bound $D^*$ by forming bounds $E(D)$ and estimating $D^*$ by $E(D)$ as discussed previously.

Let

$$E_L(D) \leq E(D) \leq E_U(D) \tag{16}$$

where $E_L(D)$ and $E_U(D)$ denote the lower and upper $(1 - c)$ confidence bounds on $E(D)$. To bound $U(t)$ with a lower bound corresponds to choosing the upper bound for $E(D)$, since $(1 - E(D))$ will be minimal for $E_L(D)$. A similar argument applies for the upper bound for $U(t)$. Thus, in terms of $D^*$ only, we have

$$U_L \leq U_{(t)} \leq U_U \qquad \cdot \tag{17}$$

where

$$U_L = (1 - E_U(D))^{F(t)}$$

and

$$U_U = (1 - E_L(D))^{F(t)}$$

for appropriate confidence levels.

The final consideration is the bounding of $F(t)$, the number of spacecraft failures after an orbital time t. Before we bound t, we must determine if the criticality of failures is independent of time. If this is the case, then we may bound $D^*$ with the bounds for $E(D)$, and bound $F(t)$ with appropriate bounds. In order to answer this question, we refer to table I-A-1 on page I-A-2 of reference (3) and perform the appropriate $X^2$ test for independence. At an $\alpha$-level of 0.01, we do not reject the null hypothesis of independence.

In order to bound $F(t)$, we use the procedure developed by Williams-Kruger, [9], for component, system, and orbital failure modes, and obtain

$$F_L(t) \leq F(t) \leq F_U(t) \tag{18}$$

where $F_U$, $F_L$ denote the upper and lower confidence bounds for E(t) where F(t) is the failure mode in orbit. This process involves using regression through the mid-points for the upper and lower confidence bounds for the Product Limit estimate for F(t). Thus, the upper and lower bounds for u(t) are given by:

$$(1 - E_U(D))^{F_L(t)} \leq U(t) \leq (1 - E_L(D))^{F_U(t)} \tag{19}$$

## Conclusions and Remarks

When one calculates bounds for expected criticality, E(D), the multivariate confidence interval approach is preferred because it deals with the dependence of the parameters $P_1$, $P_2$, ..., $P_n$, collectively through the variance—covariance structure.

In order to use the multivariate confidence approach for confidence intervals for any set of data, one calculates $\hat{P}_1$, ..., $\hat{P}_n$ from that set of data, then finds the matrices V, and $C^T$, finds D, and calculates confidence intervals as indicated in the previous discussion.

The use of confidence intervals is necessary when one uses the concept of utility to determine how much time is necessary to gain the pertinant amount of information needed from each mission. As a mission is conducted, then one has time intervals in which to gain information. As an example we consider the diagram in Figure 1.

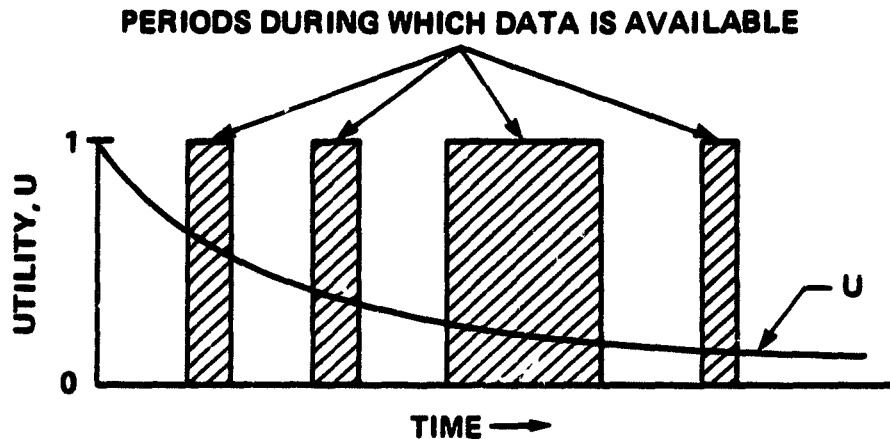**PERIODS DURING WHICH DATA IS AVAILABLE**



Figure 1. Change in Utility as a Mission Progresses.

The vertical bars in Figure 1 indicate the periods when data may be taken, such as when a spacecraft is over a ground point. The utility of the spacecraft may be expected to decrease with time as malfunctions occur. Therefore, if one needed a specific number of hours worth of information, the observing times would have to be multiplied by the value of U existing during that time.

In order to use the confidence intervals for E(D) with the concept of utility in gaining information, one would be interested in using the confidence bounds for E(D) through the mission to calculate U. If one wanted 12 hours of information, then one would numerically integrate the area in each bar of Figure 1 that is under the curve U using confidence bounds calculated as in equation (19). This would give a final estimate for the total number of hours of information gained with a degree of confidence based on the bounds for U(t) in equation (19).

One should continue to study the question of complexity of spacecraft and criticality of failures during the mission since these concepts and measurements have a direct bearing on mission cost optimization.

Bibliography

1. Bloomquist, C. E., "Use of the Space Shuttle to Avoid Spacecraft Anomalies," PRC Systems Sciences Co., Los Angeles, CA, Report PRC R-1467, 3 May 1972.

2. Bloomquist, C. E., DeMars, D., Graham, W., and Henmi, P., "On-Orbit Spacecraft Reliability," Planning Research Corp., Los Angeles, CA, Report PRC R-1863, 30 Sept. 1978.

3. Kruger, R., Norris, H. P., "A Thermal Vacuum Test Optimization Procedure," NASA, GSFC, Greenbelt, MD, TM 80297, June 1979.

4. Timmins, A. R., "A Study of Total Space Life Performance of GSFC Spacecraft," NASA TN D-8017, July 1975.

5.  Breiman, L., "Statistics: With a View Toward Applications," Houghton Mifflin Co., 1973.

6.  Mood, A. M., and Graybill, F. A., "Introduction to the Theory of Statistics," McGraw-Hill, Inc., 1963.

7.  Lehmann, E. L., "Nonparametrics: Statistical Methods Based on Ranks," Holden-Day, 1976.

8.  Hogg, R. V., and Graig, A. T., "Introduction to Mathematical Statistics," Macmillian Co., 1970.

9.  Williams, R., and Kruger, R., "Estimation Procedures to Measure and Monitor Failure Rates of Components during Thermal-Vacuum Testing," NASA, GSFC, Greenbelt, MD, TM 82052, Nov. 1980.

| 1. Report No. CR 160090 | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| 4. Title and Subtitle Spacecraft Utility and the Development of Confidence Intervals for Criticality of Anomalies | | 5. Report Date |
| | | 6. Performing Organization Code |
| 7. Author(s) Dr. Raymond E. Williams | | 8. Performing Organization Report No. |
| 9. Performing Organization Name and Address Ft. Lewis College Durango, Colorado 81301 | | 10. Work Unit No. |
| | | 11. Contract or Grant No. NAG 5-1 |
| 12. Sponsoring Agency Name and Address NASA/Goddard Space Flight Center Greenbelt, Maryland 20771 | | 13. Type of Report and Period Covered Contractor Report |
| | | 14. Sponsoring Agency Code |

15. Supplementary Notes

16. Abstract

The concept of spacecraft utility, a measure of its performance in orbit, is discussed and its formulation is described. Performance is defined in terms of the malfunctions that occur and the criticality to the mission of these malfunctions. Different approaches to establishing average or expected values of criticality are discussed and confidence intervals are developed for parameters used in the computation of utility.

| 17. Key Words (Selected by Author(s)) Utility Criticality of Failure Confidence Intervals | 18. Distribution Statement |
|---|---|

| 19. Security Classif. (of this report) Unclassified | 20. Security Classif. (of this page) Unclassified | 21. No. of Pages | 22. Price* |
|---|---|---|---|