QUALITY - SAFETY - RELIABILITY

U.S. ARMY

QUALITY ASSURANCE DIRECTORATE
PICATINNY ARSENAL
DOVER, NEW JERSEY

UNBIASED ESTIMATES OF RELIABILITY WHEN TESTING AT ONLY ONE EXTREME STRESS LEVEL

By

A. Bulfinch

December 1963

Reliability Engineering Branch
Nuclear Reliability Division
Quality Assurance Directorate
Picatinny Arsenal, Dover, New Jersey

TABLE OF CONTENTS

ABSTRACT

Based on the stress-strength concept of reliability for "one-shot" items, it is assumed that an item cannot fail until the stress equals or exceeds the strength. From this premise and the following additional assumptions, methods are given for claculating unbiased estimates of non-time dependent reliability:

1. The relation between the stress and strength standard deviations are known approximately.

2. A single stress level is applied during testing at approximately three standard deviations from the average stress level.

3. The stress and strength distributions are normal.

Calculations are included to show the effect of errors in the assumptions concerning the standard deviations, applied stress level, and rounding-off errors.

This approach further reduces the sample size required to demonstrate high non-time dependent reliability in laboratory testing. It has the added advantage of obtaining unbiased estimates of reliability with the simplest of testing methods.

Unbiased Estimates of Reliability When Testing at Only One Extreme Stress Level.

I. INTRODUCTION

The pressure of time and money in reliability testing requires a never ending quest for simpler methods and smaller sample sizes. Recent work at Picatinny Arsenal has suggested another contribution to this effort.

The usual interpretation of sample results for the determination of non-time dependent reliability, when only attribute type data can be obtained, is based on the binomial distribution. The usual laboratory method of testing is to apply a single level of stress to the sample. Under these conditions very large sample sizes are required to demonstrate reasonably high reliability values that may exist. In addition, this approach results in data very insensitive to changes in reliability. Both of these characteristics are costly shortcomings. However, the simple method of testing is an asset.

The purpose of this paper is to describe a procedure that retains the simple testing method but requires only small sample sizes for any reliability level and produces data that is sensitive to small changes in reliability. This is accomplished by changing the interpretation of the data and supplementing this, in a quantitative way, with knowledge gained from the experience of working with an item over a period of time.

However, the method presented here is limited to the laboratory determination of non-time dependent reliability when only success-failure type of data can be obtained. This type of reliability is based on stress-strength concept presented in an earlier paper (Reference 1).

The procedures proposed are an out-growth of recent work on the evaluation of laboratory methods by means of Monte Carlo sampling techniques. This work showed that when only attribute data can be obtained that:

1. The observed proportion of successes in a sample obtained at a single stress level is a biased estimate of the non-time dependent reliability defined by the stress-strength concept.

2. A sample obtained at a single stress level cannot measure the average or standard deviation of the strength distribution.

3. The observed failure rate, obtained at a single stress level measures the area of the tail of the strength curve to the left of (below) the applied stress ordinate.

From the above, it was realized that sample results obtained at a single stress level furnished information about the strength distribution. This suggested the possibility of making use of this fact for obtaining unbiased estimates of reliability, with the very simple method of testing at a single stress level, by changing the use made of sample results.

II. METHOD OF CALCULATION

The method of calculation described below is based on the normal deviate:

$$T = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{s_1^2 + s_2^2}}$$

Where: $\bar{X}_1$ = Average stress expected in use

$\bar{X}_2$ = Average strength

s_1^2 = Variance of the stress distribution

s_2^2 = Variance of the strength distribution

The previous work referred to above shows that precise and unbiased estimates of the true non-time dependent reliability can be obtained by entering a table of areas under the standard normal curve with this calculated T-value. This is true of course only when the stress and strength distributions are normally distributed. The sensitivity of this function to deviations from normality is yet to be demonstrated.

Since testing at a single stress level cannot measure the average and standard deviation of the strength distribution, the above formula was transposed to an equivalent function as follows:

Let X = Any applied stress level used in testing

Then:

$$T_1 = \frac{X - \bar{X}_1}{s_1}$$

$T_1 s_1 = X - \bar{X}_1$ (for the stress distribution)

$$T_2 = \frac{\bar{X}_2 - X}{s_2}$$

$T_2 s_2 = \bar{X}_2 - X$ (for the strength distribution)

$\bar{X}_2 - \bar{X}_1 = (X - \bar{X}_1) + (\bar{X}_2 - X)$

$= T_1 s_1 + T_2 s_2$ (by substitution)

When: $m s_1 = s_2$

$\bar{X}_2 - \bar{X}_1 = T_1 s_1 + m T_2 s_1$

$$T = \frac{T_1 s_1 + m T_2 s_1}{\sqrt{s_1^2 + (m s_1)^2}}$$

$$= \frac{s_1 (T_1 + m T_2)}{s_1 \sqrt{1 + m^2}}$$

$$= \frac{T_1 + m T_2}{\sqrt{1 + m^2}}$$

III. ASSUMPTIONS

The last formula can be used under the following assumptions:

1. The stress and strength distributions are normal
2. Where $m s_1 = s_2$, m is known
3. The testing is done at a stress of $(\bar{X}_1 + T_1 s_1)$, where T_1 is known approximately.

IV. DISCUSSION OF ASSUMPTIONS

1. If there is reason to question the assumption of normality, appropriate distribution free methods can be used. However, the form of the distribu- should be determined where possible.

2. Experience has shown that m is approximately two. The examples given below show that the value of m can vary widely before seriously affecting the accuracy of the resulting reliability value.

3. $(\bar{X} + T_1 s_1)$ can be defined as the maximum stress expected in use. This level of stress is usually known by the development engineer or is specified in the Military Characteristics. Such a maximum stress can be defined statistically as the stress occurring only once in a thousand or once in ten thousand times. As such, $T_1 = 3.09$ or $T_1 = 3.72$ respectively. The examples given below show that T_1 can also vary widely before seriously affecting the accuracy of the resultant reliability value.

V. USE OF MODIFIED T - FORMULA

In the above formula, T_2 is measured by the observed failure rate of the sample tested at a single stress level (X). Its numerical value can be obtained by entering a table of areas under the standard normal curve with the proportion of failures in the sample. With this value determined and the values of T_1 and m known or assigned, the above formula can be used without knowing $\bar{X}_2$ or s_2 the average and standard deviation of the strength distribution.

The average and standard deviation of the stress distribution must be separately determined. If this information is not available and cannot be determined, the determination of a numerical value for reliability is impossible.

VI. ACCURACY AND SENSITIVITY

The incentive for using the proposed method of calculating reliability is that it can furnish considerably more information about the existing reliability than the usual way of using sample success-failure results. The examples given in Table I show this quite well. Obtaining 50% sample failures in this method is not as bad as it might seem. If the 50% point of the strength curve is at the three (3.09) sigma point of the stress curve, the reliability equals .9162 (when m equals 2) - not 50%, the proportion of successes in the sample.

TABLE I

ACCURACY AND SENSITIVITY

Sample Size: $n = 22$

Number of Sample Failures: $= b$

Standard Deviations: $2S_1 = S_2$

Testing Level: $T_1 = 3.09$ ($P = .001$)

$$T = \frac{3.09 + 2T_2}{\sqrt{1 + (2)^2}}$$

SAMPLE			T-FORMULA		(1-b/n)	
b	b/n	T_2	T	Reliability	Reliability	Difference*
11	.5000	0.00	1.38	.9162	.5000	+.4162
10	.4545	0.11	1.48	.9306	.5455	+.3851
9	.4092	0.23	1.59	.9441	.5908	+.3533
8	.3636	0.35	1.69	.9545	.6364	+.3181
7	.3182	0.47	1.80	.9641	.6818	+.2823
6	.2728	0.60	1.92	.9726	.7272	+.2454
5	.2272	0.75	2.05	.9798	.7728	+.2070
4	.1818	0.91	2.19	.9857	.8182	+.1675
3	.1364	1.10	2.36	.9909	.8636	+.1273
2	.0909	1.34	2.58	.9951	.9091	+.0860
1	.0454	1.69	2.90	.9981	.9546	+.0435

* T-Formula reliability minus the observed proportion of successes in the sample.

The results in Table I show the sensitivity of the proposed method to changes in reliability values. A decrease of .03 in the reliability at the upper end of the scale increases the number of failures in the sample of 22, from zero to six. This is a significant difference at the 95% condifence level.

The above sensitivity is to be compared with the insensitivity of the method of using the observed proportion of successes in the sample as the point estimate of "reliability". In this method, where the binomial distribution pertains, the success probability ("reliability") must decrease approximately 0.23 (1.00 - .77) before the observed number of failures in the sample increases a significant (0 to 5) amount at the 95% level of confidence.

The above comparison of sensitivity shows that the proposed method is sensitive to changes in reliability. That is, the proposed method can detect relatively small changes in reliability with small sample sizes. This is an important property for a laboratory method. It means that small differences between design modifications and small changes occurring during storage can be readily detected.

VII. ERRORS DUE TO ASSUMPTIONS

The relative accuracies of the two methods for determining reliability are shown by the "differences" given in Table I. These differences are to be compared with the errors, resulting from incorrect assumptions shown in Table II. The assumption errors made here are the maximum expected in practice due to total ignorance about the system concerned. Any knowledge gained about a component or a system through experience will improve the accuracy of the assumptions and thereby reduce the resultant errors. This kind of knowledge, from experience, is always available and can be effectively used in the proposed method of calculation.

TABLE II

EFFECT OF ASSUMPTIONS AS FAILURE RATE INCREASES

Test Level: $U_1 + T_1 S_1$

Standard Deviation: $mS_1 = S_2$

FAILURE RATE (b/n)	m	T_1	P_1	T_2	T	POINT ESTIMATE
.02	1	2.33	.0100	2.05	3.10	.99903
.02	1	3.72	.0001	2.05	4.08	.99998
.02	2	3.09	.0010	2.05	3.21	.99934
.02	3	2.33	.0100	2.05	2.68	.99632
.02	3	3.72	.0001	2.05	3.12	.99910

Maximum Error:

Assuming the most favorable (highest reliability) condition when in fact the most unfavorable condition actually exists: .99998 - .99632 = +.00366.

Median Errors:

Assuming the median ($m = 2$; $T_1 = 3.09$) condition when the most favorable (1) and Unfavorable (2) conditions exist: (1) .99934 - .99998 = -.00064

(2) .99934 - .99632 = +.00302

Effect of Assumptions (continued):

Test Level: $U_1 + T_1 S_1$

Standard Deviation: $mS_1 = S_2$

FAILURE RATE (b/n)	m	T_1	P_1	T_2	T	POINT ESTIMATE
.05	1	2.33	.0100	1.65	2.81	.99752
.05	1	3.72	.0001	1.65	3.80	.99993
.05	2	3.09	.0010	1.65	2.85	.99781
.05	3	2.33	.0100	1.65	2.30	.98927
.05	3	3.72	.0001	1.65	2.74	.99693

Maximum Error:

Assuming the most favorable (highest reliability) condition when in fact the most unfavorable condition actually exists:

$$.99993 - .98927 = +.01066$$

Median Errors:

Assuming the median ($m = 2$; $T_1 = 3.09$) condition when the most favorable (1) and unfavorable (2) conditions exist:

(1) $.99781 - .99993 = .00212$

(2) $.99781 - .98927 = +.00854$

Effect of Assumptions (continued):

Test Level: $U_1 + T_1 S_1$

Standard Deviation: $mS_1 = S_2$

FAILURE RATE (b/n)	m	T_1	P_1	T_2	T	POINT ESTIMATE
.10	1	2.33	.0100	1.28	2.55	.99461
.10	1	3.72	.0001	1.28	3.53	.99979
.10	2	3.09	.0010	1.28	2.52	.99413
.10	3	2.33	.0100	1.28	1.95	.97441
.10	3	3.72	.0001	1.28	2.39	.99157

Maximum Error:

Assuming the most favorable (highest reliability) condition when in fact the most unfavorable condition actually exists:

$$.99979 - .97441 = +.02538$$

Median Errors:

Assuming the median ($m = 2$; $T_1 = 3.09$) condition when the most favorable (1) and unfavorable (2) conditions exists:

(1) $.99413 - .99979 = -.00566$

(2) $.99413 - .97441 = +.01972$

Effect of Assumptions (continued):

Test Level: $U_1 + T_1 S_1$

Standard Deviation: $mS_1 = S_2$

FAILURE RATE (b/n)	m	T_1	P_1	T_2	T	POINT ESTIMATE
.20	1	2.33	.0100	0.84	2.24	.98745
.20	1	3.72	.0001	0.84	3.22	.99936
.20	2	3.09	.0010	0.84	2.13	.98341
.20	3	2.33	.0100	0.84	1.53	.93700
.20	3	3.72	.0001	0.84	1.97	.97558

Maximum Error:

Assuming the most favorable (highest reliability) condition when in fact the most unfavorable condition actually exists:

$$.99936 - .93700 = .06236$$

Median Errors:

Assuming the median ($m = 2$; $T_1 = 3.09$) condition when the most favorable (1) and unfavorable (2) conditions exist:

(1) .98341 - .99936 = -.01595

(2) .98341 - .93700 = +.04641

TABLE III

SUMMARY OF ERRORS

Failure Rate	Errors Due to Using Sample Proportion of Successes as Point Estimate	Errors Due to Assumptions Maximum	Errors Due to Assumptions Median
.02	+.019	+.004	+.003
.05	+.048	+.011	+.008
.10	+.094	+.025	+.020
.20	+.183	+.062	+.046

The sample errors in Table III were obtained by subtracting (1 - b/n) from the point estimates (in Table II) for M = 2 and T_1 = 3.09 - the median conditions. The assumption errors in Table III were obtained by rounding off the corresponding errors in Table II.

The data in Table III show that both types of errors increase as the observed proportion of failures (failure rate) increases. However, in each case the assumption errors are less than the sampling errors. The magnitude of the assumption errors up through a failure rate of 0.10 is not great enough to seriously affect the reliability value. Some knowledge of T_1 or m will greatly reduce these errors in the calculated reliability.

VIII. EFFECT OF ROUNDING OFF ERRORS

When sample sizes are small, rounding off errors may be important. Their effects at various failure rates are shown in Table IV and Table V.

TABLE IV

EFFECT OF ROUNDING OFF ERRORS

(SAMPLE CALCULATIONS)

Test Level: $U_1 + 3.09\ S_1$

Standard Deviation: $2S_1 = S_2$

b	n	b/n	T_1	P_1	T_2	T	Point Estimate
		.45	3.09	.001	+.13	1.50	.9332
5	10	.50	3.09	.001	.00	1.38	.9162
		.55	3.09	.001	-.13	1.27	.8979

b	n	b/n	T_1	P_1	T_2	T	Point Estimate
		.20	3.09	.001	.84	2.13	.9834
5	20	.25	3.09	.001	.68	1.99	.9767
		.30	3.09	.001	.53	1.85	.9678

b	n	b/n	T_1	P_1	T_2	T	Point Estimate
		.05	3.09	.001	1.65	2.85	.9978
5	50	.10	3.09	.001	1.29	2.54	.9945
		.15	3.09	.001	1.04	2.31	.9896

TABLE V

SUMMARY OF EFFECT OF ROUNDING OFF ERRORS

Test Level: $U_1 + 3.09\ S_1$

Standard Deviation: $2S_1 = S_2$

FAILURE RATE	MAXIMUM ERROR
.05	.0006
.10	.0082
.20	.0128
.30	.0185
.40	.0250
.50	.0353

The errors shown in Table V are the differences between the maximum and minimum reliability values for each failure rate (b/n). The method of calculating the maximum and minimum values is based on the assumption of rounding off errors of ± 0.05 in the failure rate as shown in Table IV.

Although the assumed rounding off error is the maximum expected, its magnitude is not excessive below a failure rate of 0.30. As shown in Table V, this type of error also increases with the failure rate.

IX. USE OF CHEBYSHEV'S INEQUALITY

There is little or no information available on the form of strength distributions of most missiles and missile components. Furthermore, it is costly to obtain. It would be helpful if a distribution free procedure such as Chebyshev's irrequality could be used. As shown in Table VI, the use of Chebyshev's inequality in the modified T-formula resulted in ridiculous values.

TABLE VI

CHEBYSHEV'S INEQUALITY

Test Level: $U_1 + T_1 S_1$

Standard Deviation: $2S_1 = S_2$

Failure Rate (b/n)	T_1	P_1	T_2	T
0.50	31.62	.001	2	15.4

The T-value of 15.4 shown in Table VI is to be compared with the T-value of 1.38 shown in Table I for a failure rate of 0.50. From this, it is concluded that Chevyshev's inequality cannot be used in this application.

X. EXAMPLES

Previous work (Reference 1) has shown that the true non-time dependent reliability of the set of conditions used in these examples can be obtained by means of the following formula:

$$Z \geq \frac{U_2 - U_1}{\sqrt{S_1^2 + S_2^2}}$$

Where:

U_1 = True mean of the stress distribution

S_1^2 = True variance of the stress distribution

U_2 = True mean of the strength distribution

S_2^2 = True variance of the strength distribution

The reliability value obtained by means of the above formula was used to determine the accuracy of the following two methods of using attribute data obtained from the application a single stress level:

1. Using the observed proportion of successes in the sample as the reliability point estimate

2. Using the observed proportion of failures in the sample as a measure of the area of the strength distribution, below the applied stress, to obtain T_2 in the T-formula.

The errors associated with the two methods of using sample data are to be compared to show the practical value of the method proposed here.

The conditions used in this example are:

Stress	Strength
$U_1 = 10$	$U_2 = 42$
$S_1 = 5$	$S_2 = 10$

The true non-time dependent reliability for this set of conditions can be calculated as follows:

$$Z \geq \frac{42 - 10}{\sqrt{(10)^2 + (5)^2}} = \frac{32}{125} = 2.86$$

The true reliability associated with this Z-value is 0.9979.

1. First Method

Using the observed proportion of successes as the point estimate:

If it is assumed that the testing is done at $U_1 + 3S_1$, then the applied stress will be equal to 25 units. For the set of conditions described above, the portion of the strength distribution below 25 units can be found as follows:

$$Z_2 = \frac{42 - 25}{10} = 1.70$$

Entering a table of areas under the standard normal curve with this Z_2 value, the following value is obtained:

$$P = .0446$$

The earlier work referred to above shows that this latter value is the expected failure rate of the single-stress-level method. The complement of this value (.9554) would be taken as the "true" mean reliability of this method. The difference between 0.9979 and 0.9554 (0.0425) is considered the expected error of the single-stress-level method when the proportion of successes in the sample is taken as the point estimate.

2. Second Method

Using the observed proportion of successes as a measure of the area in the tail of the strength curve:

The practical value of the method proposed here can best be demonstrated by calculating the magnitude of the errors due to the assumptions made concerning m and Z_1. Using the set of conditions described above, the variations in m and Z_1 used below are the maximum considered likely in practice. Therefore, the errors in the reliability values caused by these variations are the maximum expected.

TABLE VII

VARIATIONS DUE TO ASSUMPTION ERRORS

Failure Failure Rate	T_2	m	Z_1	P_1	T	Point Estimate
1/22	1.70	1	2.33	.0100	2.85	.9978
1/22	1.70	1	3.72	.0001	3.83	.9999
1/22	1.70	3	2.33	.0100	2.35	.9906
1/22	1.70	3	3.72	.0001	2.79	.9974

The following errors were obtained by calculating the differences between the true value and the point estimates shown in Table VII:

ERRORS

M	Z_1	Differences
1	2.33	.0001
1	3.72	.0020
3	2.33	.0073
3	3.72	.0005

These errors are to be compared with 0.0425, the error obtained when the sample result was used as the point estimate of non-time dependent reliability.

XI. CONCLUSIONS

1. The proposed use of attribute data to estimate non-time dependent reliability by the single-stress-level method is more accurate at all levels of reliability than the usual method of using the proportion of successes in the sample as the reliability point estimate.

2. The proposed method is more sensitive to changes in reliability than the usual method.

3. The proposed method permits the knowledge gained through the experience of working with an item to be used in a quantitative way and thereby reduce the sample size required to obtain an unbiased estimate of reliability.

4. When the true reliability of an item is in fact as high as 0.995 (the usual value of Military Characteristics requirements) and the stress is applied at the three-sigma level, the expected error in the proposed method is less than 1.0%.

XII. REFERENCES

1. Bulfinch A., "Evaluation of Various Laboratory Methods for Determining Reliability", Proceeding Eighth Conference on the Design of Experiments in Army Research, Development, and Testing, October 1962.

2. Lloyd, David K., and Lipow, Myron, "Reliability Management, Methods, and Mathematics", Prentice-Hall, Inc., Englewood Cliffs, N.Y., 1962.