Designation: E739 – 23

Standard Guide for Statistical Analysis of Linear or Linearized Stress-Life (*S-N*) and Strain-Life (ε-*N*) Fatigue Data[1]

This standard is issued under the fixed designation E739; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide covers only *S-N* and ε-*N* relationships that may be reasonably approximated by a straight line (on appropriate coordinates) for a specific interval of stress or strain. It presents elementary procedures that presently reflect good practice in modeling and analysis. However, because the actual *S-N* or ε-*N* relationship is approximated by a straight line only within a specific interval of stress or strain, and because the actual fatigue life distribution is unknown, it is *not recommended* that (*a*) the *S-N* or ε-*N* curve be extrapolated outside the interval of testing, or (*b*) the fatigue life at a specific stress or strain amplitude be estimated below approximately the fifth percentile ($P \simeq 0.05$). As alternative fatigue models and statistical analyses are continually being developed, later revisions of this guide may subsequently present analyses that permit more complete interpretation of *S-N* and ε-*N* data.

1.2 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E468 Practice for Presentation of Constant Amplitude Fatigue Test Results for Metallic Materials

E606/E606M Test Method for Strain-Controlled Fatigue Testing

E1823 Terminology Relating to Fatigue and Fracture Testing

3. Terminology

3.1 The terms used in this guide shall be used as defined in Terminology E1823. In addition, the following terminology is used:

3.1.1 *dependent variable*—the fatigue life *N* (or the logarithm of the fatigue life).

3.1.1.1 *Discussion*—Log (*N*) is denoted *Y* in this guide.

3.1.2 *independent variable*—the selected and controlled variable (namely, stress or strain). It is denoted *X* in this guide when plotted on appropriate coordinates.

3.1.3 *log-normal distribution*—the distribution of *N* when log (*N*) is normally distributed. (Accordingly, it is convenient to analyze log (*N*) using methods based on the normal distribution.)

3.1.4 *replicate (repeat) tests*—nominally identical tests on different randomly selected test specimens conducted at the same nominal value of the independent variable *X*. Such replicate or repeat tests should be conducted independently; for example, each replicate test should involve a separate set of the test machine and its settings.

3.1.5 *run out*—no failure at a specified number of load cycles (Practice E468).

3.1.5.1 *Discussion*—The analyses illustrated in this guide do not apply when the data include either run-outs (or suspended tests). Moreover, the straight-line approximation of the *S-N* or ε-*N* relationship may not be appropriate at long lives when run-outs are likely.

3.1.5.2 *Discussion*—For purposes of statistical analysis, a run-out may be viewed as a test specimen that has either been removed from the test or is still running at the time of the data analysis.

4. Significance and Use

4.1 Materials scientists and engineers are making increased use of statistical analyses in interpreting *S-N* and ε-*N* fatigue data. Statistical analysis applies when the given data can be reasonably assumed to be a random sample of (or representation of) some specific defined population or universe of material of interest (under specific test conditions), and it is desired either to characterize the material or to predict the performance of future random samples of the material (under similar test conditions), or both.

---

[1] This guide is under the jurisdiction of ASTM Committee E08 on Fatigue and Fracture and is the direct responsibility of Subcommittee E08.04 on Structural Applications.

Current edition approved Feb. 1, 2023. Published March 2023. Originally approved in 1980. Last previous edition approved in 2015 as E739 – 10(2015). DOI: 10.1520/E0739-23.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

5. Types of *S-N* and ε-*N* Curves Considered

5.1 It is well known that the shape of *S-N* and ε-*N* curves can depend markedly on the material and test conditions. This guide is restricted to linear or linearized *S-N* and ε-*N* relationships, for example,

$$\log N = A + B\,(S) \quad \text{or} \tag{1}$$

$$\log N = A + B\,(\varepsilon) \quad \text{or}$$

$$\log N = A + B\,(\log S) \quad \text{or} \tag{2}$$

$$\log N = A + B\,(\log \varepsilon)$$

in which *S* and ε may refer to (*a*) the maximum value of constant-amplitude cyclic stress or strain, given a specific value of the stress or strain ratio, or of the minimum cyclic stress or strain, (*b*) the amplitude or the range of the constant-amplitude cyclic stress or strain, given a specific value of the mean stress or strain, or (*c*) analogous information stated in terms of some appropriate independent (controlled) variable.

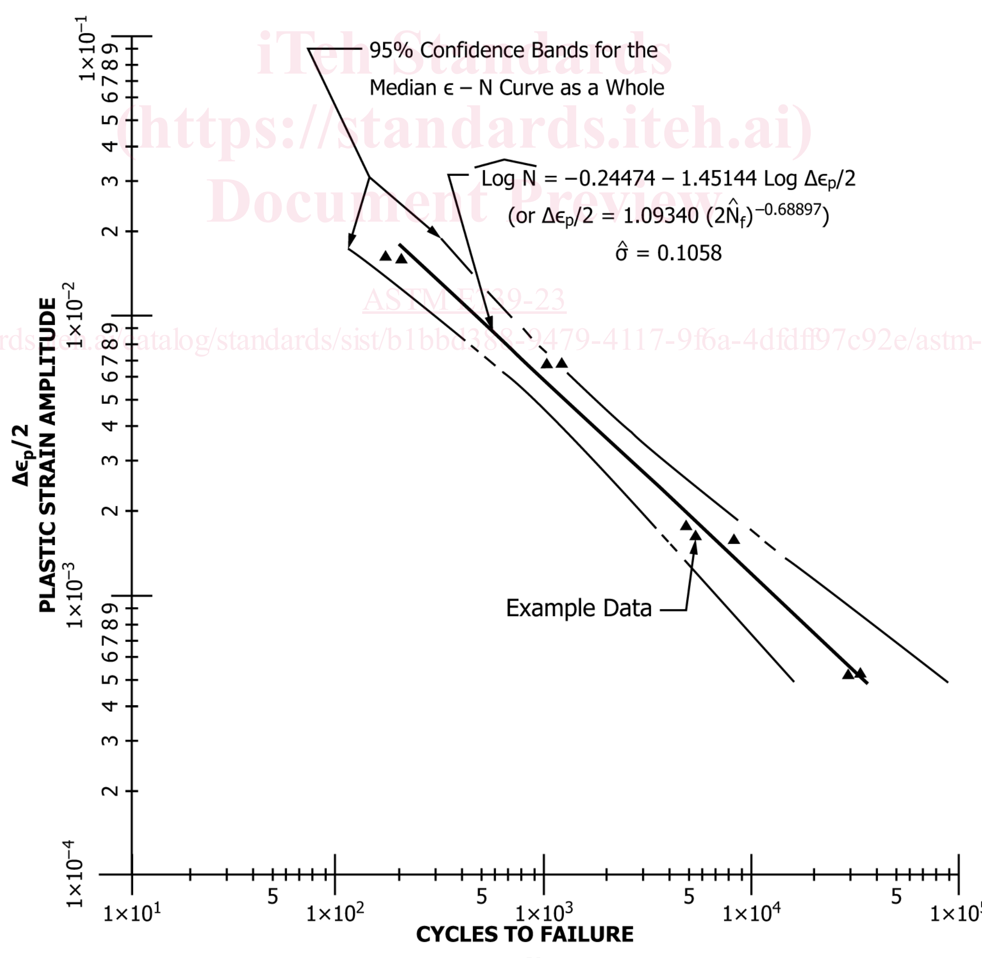
NOTE 1—In certain cases, the amplitude of the stress or strain is not constant during the entire test for a given specimen. In such cases some effective (equivalent) value of *S* or ε must be established for use in analysis.

5.1.1 The fatigue life *N* is the dependent (random) variable in *S-N* and ε-*N* tests, whereas *S* or ε is the independent (controlled) variable.

NOTE 2—In certain cases, the independent variable used in analysis is not literally the variable controlled during testing. For example, it is common practice to analyze low-cycle fatigue data treating the range of plastic strain as the controlled variable, when in fact the range of total strain was actually controlled during testing. Although there may be some question regarding the exact nature of the controlled variable in certain *S-N* and ε-*N* tests, there is never any doubt that the fatigue life is the dependent variable.

NOTE 3—In plotting *S-N* and ε-*N* curves, the independent variables *S* and ε are plotted along the ordinate, with life (the dependent variable) plotted along the abscissa. Refer, for example, to Fig. 1.

5.1.2 The distribution of fatigue life (in any test) is unknown (and indeed may be quite complex in certain situations). For the purposes of simplifying the analysis (while maintaining sound statistical procedures), it is assumed in this guide that the logarithms of the fatigue lives are normally distributed, that is, the fatigue life is log-normally distributed, and that the variance of log life is constant over the entire range of the independent variable used in testing (that is, the scatter in log



NOTE 1—The 95 % confidence band for the ε-*N* curve as a whole is based on Eq 10. (Note that the dependent variable, fatigue life, is plotted here along the abscissa to conform to engineering convention.)

FIG. 1 Fitted Relationship Between the Fatigue Life *N* (*Y*) and the Plastic Strain Amplitude $\Delta\varepsilon_p/2$ (*X*) for the Example Data Given

N is assumed to be the same at low S and ε levels as at high levels of S or ε). Accordingly, log N is used as the dependent (random) variable in analysis. It is denoted Y. The independent variable is denoted X. It may be either S or ε, or log S or log ε, respectively, depending on which appears to produce a straight line plot for the interval of S or ε of interest. Thus Eq 1 and Eq 2 may be re-expressed as

$$Y = A + BX \tag{3}$$

Eq 3 is used in subsequent analysis. It may be stated more precisely as $\mu_{Y|X} = A + BX$, where $\mu_{Y|X}$ is the expected value of Y given X.

NOTE 4—For testing the adequacy of the linear model, see 8.2.

NOTE 5—The expected value is the mean of the conceptual population of all Y's given a specific level of X. (The median and mean are identical for the symmetrical normal distribution assumed in this guide for Y.)

6. Test Planning

6.1 Test planning for S-N and ε-N test programs is discussed in Chapter 3 of Ref **(1)**.[3] Planned grouping (blocking) and randomization are essential features of a well-planned test program. In particular, good test methodology involves use of planned grouping to (*a*) balance potentially spurious effects of nuisance variables (for example, laboratory humidity) and (*b*) allow for possible test equipment malfunction during the test program.

7. Sampling

7.1 It is vital that sampling procedures be adopted that assure a random sample of the material being tested. A random sample is required to state that the test specimens are representative of the conceptual universe about which both statistical and engineering inference will be made.

NOTE 6—A random sampling procedure provides each specimen that conceivably could be selected (tested) an equal (or known) opportunity of actually being selected at each stage of the sampling process. Thus, it is poor practice to use specimens from a single source (plate, heat, supplier) when seeking a random sample of the material being tested unless that particular source is of specific interest.

NOTE 7—Procedures for using random numbers to obtain random samples and to assign stress or strain amplitudes to specimens (and to establish the time order of testing) are given in Chapter 4 of Ref **(2)**.

7.1.1 *Sample Size*—The minimum number of specimens required in S-N (and ε-N) testing depends on the type of test program conducted. The following guidelines given in Chapter 3 of Ref **(1)** appear reasonable.

Type of Test	Minimum Number of Specimens[A]
Preliminary and exploratory (exploratory research and development tests)	6 to 12
Research and development testing of components and specimens	6 to 12
Design allowables data	12 to 24
Reliability data	12 to 24

[A] If the variability is large, a wide confidence band will be obtained unless a large number of specimens are tested (See 8.1.1).

7.1.2 *Replication*—The replication guidelines given in Chapter 3 of Ref **(1)** are based on the following definition:

% replication = 100 [1 – (total number of different stress or strain levels used in testing/total number of specimens tested)]

Type of Test	Percent Replication[A]
Preliminary and exploratory (research and development tests)	17 to 33 min
Research and development testing of components and specimens	33 to 50 min
Design allowables data	50 to 75 min
Reliability data	75 to 88 min

[A] Note that percent replication indicates the portion of the total number of specimens tested that may be used for obtaining an estimate of the variability of replicate tests.

7.1.2.1 *Replication Examples*—Good replication: Suppose that ten specimens are used in research and development for the testing of a component. If two specimens are tested at each of five stress or strain amplitudes, the test program involves 50 % replications. This percent replication is considered adequate for most research and development applications. Poor replication: Suppose eight different stress or strain amplitudes are used in testing, with two replicates at each of two stress or strain amplitudes (and no replication at the other six stress or strain amplitudes). This test program involves only 20 % replication, which is not generally considered adequate.

8. Statistical Analysis (Linear Model $Y = A + BX$, Log-Normal Fatigue Life Distribution with Constant Variance Along the Entire Interval of X Used in Testing, No Runouts or Suspended Tests or Both, Completely Randomized Design Test Program)

8.1 For the case where (*a*) the fatigue life data pertain to a random sample (all Y_i are independent), (*b*) there are neither run-outs nor suspended tests and where, for the entire interval of X used in testing, (*c*) the S-N or ε-N relationship is described by the linear model $Y = A + BX$ (more precisely by $\mu_{Y|X} = A + BX$), (*d*) the (two parameter) log-normal distribution describes the fatigue life N, and (*e*) the variance of the log-normal distribution is constant, the maximum likelihood estimators of A and B are as follows:

$$\hat{A} = \bar{Y} - \hat{B}\bar{X} \tag{4}$$

$$\hat{B} = \frac{\sum_{i=1}^{k} \left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^{k} \left(X_i - \bar{X}\right)^2} \tag{5}$$

where the symbol "caret" (^) denotes estimate (estimator),

[3] The boldface numbers in parentheses refer to the list of references appended to this standard.

the symbol "overbar" ($^{-}$) denotes average (for example, $\bar{Y} = \sum_{i=1}^{k} Y_i/k$ and $\bar{X} = \sum_{i=1}^{k} X_i/k$), $Y_i = \log N_i$, $X_i = S_i$ or ε_i, or $\log S_i$ or $\log \varepsilon_i$ (refer to Eq 1 and Eq 2), and k is the total number of test specimens (the total sample size). The recommended expression for estimating the variance of the normal distribution for log N is

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{k} \left(Y_i - \hat{Y}_i\right)^2}{k-2} \quad (6)$$

in which $\hat{Y}_i = \hat{A} + \hat{B}X_i$ and the $(k - 2)$ term in the denominator is used instead of k to make $\hat{\sigma}^2$ an unbiased estimator of the normal population variance $\sigma\hat{}^2$.

NOTE 8—An assumption of constant variance is usually reasonable for notched and joint specimens up to about 10^6 cycles to failure. The variance of unnotched specimens generally increases with decreasing stress (strain) level (see Section 9). If the assumption of constant variance appears to be dubious, the reader is referred to Ref (**3**) for the appropriate statistical test.

8.1.1 *Confidence Intervals for Parameters A and B*—The estimators $\hat{A}$ and $\hat{B}$ are normally distributed with expected values A and B, respectively, (regardless of total sample size k) when conditions (*a*) through (*e*) in 8.1 are met. Accordingly, confidence intervals for parameters A and B can be established using the t distribution, Table 1. The confidence interval for A is given by $\hat{A} \pm t_p\hat{\sigma}_{\hat{A}}$, or

$$\hat{A} \pm t_p\,\hat{\sigma}\left[\frac{1}{k} + \frac{\bar{X}^2}{\sum_{i=1}^{k}\left(X_i - \bar{X}\right)^2}\right]^{½}, \quad (7)$$

and for B is given by $\hat{B}\hat{} \pm t_p\hat{\sigma}_{\hat{B}}$, or

$$\hat{B} \pm t_p\hat{\sigma}\left[\sum_{i=1}^{k}\left(X_i - \bar{X}\right)^2\right]^{-½} \quad (8)$$

in which the value of t_p is read from Table 1 for the desired value of P, the confidence level associated with the confidence interval. This table has one entry parameter (the statistical degrees of freedom, n, for t). For Eq 7 and Eq 8, $n = k - 2$.

NOTE 9—The confidence intervals for A and B are exact if conditions (*a*) through (*e*) in 8.1 are met exactly. However, these intervals are still reasonably accurate when the actual life distribution differs slightly from the (two-parameter) log-normal distribution, that is, when only condition (*d*) is not met exactly, due to the robustness of the t statistic.

NOTE 10—Because the actual median *S-N* or ε-*N* relationship is only approximated by a straight line within a specific interval of stress or strain, confidence intervals for A and B that pertain to confidence levels greater than approximately 0.95 are not recommended.

8.1.1.1 The meaning of the confidence interval associated with, say, Eq 8 is as follows (Note 11). If the values of t_p given in Table 1 for, say, $P = 95$ % are used in a series of analyses involving the estimation of B from independent data sets, then in the long run we may expect 95 % of the computed intervals to include the value B. If in each instance we were to assert that B lies within the interval computed, we should expect to be correct 95 times in 100 and in error 5 times in 100: that is, the statement "B lies within the computed interval" has a 95 % probability of being correct. But there would be no operational meaning in the following statement made in any one instance: "The probability is 95 % that B falls within the computed interval in this case" since B either does or does not fall within the interval. It should also be emphasized that even in independent samples from the same universe, the intervals given by Eq 8 will vary both in width and position from sample to sample. (This variation will be particularly noticeable for small samples.) It is this series of (random) intervals "fluctuating" in size and position that will include, ideally, the value B 95 times out of 100 for $P = 95$ %. Similar interpretations hold for confidence intervals associated with other confidence levels. For a given total sample size k, it is evident that the width of the confidence interval for B will be a minimum whenever

$$\sum_{i=1}^{k}\left(X_i - \bar{X}\right)^2 \quad (9)$$

is a maximum. Since the X_i levels are selected by the investigator, the width of confidence interval for B may be reduced by appropriate test planning. For example, the width of the interval will be minimized when, for a fixed number of available test specimens, k, half are tested at each of the extreme levels X_{min} and X_{max}. However, this allocation should be used only when there is strong *a priori* knowledge that the *S-N* or ε-*N* curve is indeed linear—because this allocation precludes a statistical test for linearity (8.2). See Chapter 3 of Ref (**1**) for a further discussion of efficient selection of stress (or strain) levels and the related specimen allocations to these stress (or strain) levels.

NOTE 11—This explanation is similar to that of STP 313 (**4**).

8.1.2 *Confidence Band for the Entire Median S-N or ε-N Curve (that is, for the Median S-N or ε-N Curve as a Whole)*—If conditions (*a*) through (*e*) in 8.1 are met, an exact confidence band for the entire median *S-N* or ε-*N* curve (that is, all points on the linear or linearized median *S-N* or ε-*N* curve considered simultaneously) may be computed using the following equation:

TABLE 1 Values of t_p (Abstracted from STP 313 (4))

n[A]	P, %[B]	
	90	95
4	2.1318	2.7764
5	2.0150	2.5706
6	1.9432	2.4469
7	1.8946	2.3646
8	1.8595	2.3060
9	1.8331	2.2622
10	1.8125	2.2281
11	1.7959	2.2010
12	1.7823	2.1788
13	1.7709	2.1604
14	1.7613	2.1448
15	1.7530	2.1315
16	1.7459	2.1199
17	1.7396	2.1098
18	1.7341	2.1009
19	1.7291	2.0930
20	1.7247	2.0860
21	1.7207	2.0796
22	1.7171	2.0739

[A] n is not sample size, but the degrees of freedom of t, that is, $n = k - 2$.
[B] P is the probability in percent that the random variable t lies in the interval from $-t_p$ to $+t_p$.