CHAPTER 8

MECHANICAL RELIABILITY FUNDAMENTALS AND EXAMPLE ANALYSES

8.0 INTRODUCTION

The first step in mechanical design for a new product is to synthesize (configure) the product such that it performs its desired function. Design synthesis is markedly enhanced by first recognizing functional analogies among existing designs, then comparing alternative feasible designs, and in turn proposing the design that appears to have the greatest overall advantage. The second step in mechanical design for a new product is to try to assure that the proposed design will reliably perform its function. Tentative assurance of adequate reliability requires a new set of comparisons interwoven in a combination of design analyses and laboratory tests whose goal is to extrapolate service-based performance data for similar products to the proposed product. However, it is imperative to understand that adequate reliability is established only by actual service performance. It cannot be established by a combination of design analyses based on analytical bogies (design allowables, factors of safety) and laboratory tests based on experimental bogies (extreme load and environment histories). Nevertheless, a combination of design analysis and laboratory testing can be effectively employed either to maintain or to improve the reliability of a product.

When the mechanical design objective is to maintain the service-proven reliability of a product following its re-design, the re-design is required to meet the same set of analytical and laboratory test bogies that were met by the original design. But, when the mechanical design objective is to improve product reliability, the performance of the re-design must excel the performance of the original design. This improved performance must be evident in laboratory tests before it can be presumed that it will also be evident in service operation. Accordingly, reliability improvement laboratory tests should always be conducted using load and environment histories that are nominally identical to service load and environment histories. In particular, all failures in these reliability improvement laboratory tests should be identical in location, mode of failure, and appearance to corresponding service failures.

This chapter covers reliability concepts and reliability applications of statistical analyses that can be used to improve the reliability of a product. The more credible the performance comparison of the re-design to the original design, the more likely the predicted reliability improvement will actually be realized in service operation.

8.1 MECHANICAL RELIABILITY TERMINOLOGY

All solid materials resist failure. When the stimulus for failure is stated in terms of stress, the corresponding resistance is technically termed strength. However, the terms stress and strength are widely used in a generic sense to connote stimulus and resistance respectively.

Strengths and resistances are always established experimentally by strictly following a standardized
laboratory test method that has been severely constrained and simplified so that the resulting datum values will (allegedly) be repeatable within the given laboratory and reproducible between different laboratories. But, from a mechanical reliability perspective, all such standardized laboratory test methods are self-defeating — because there is no dependable way to use the resulting datum values to predict service behavior. Regardless of what may be alleged elsewhere, there is no direct relationship between the environmental and loading histories respectively pertaining to standardized laboratory test methods and actual service operation for any mechanical mode of failure.

Test duration in a reliability experiment test program is defined in terms of the environmental and loading histories imposed prior to failure, where failure connotes the discontinuation of satisfactory performance. When these environmental and loading histories exhibit recognizable increments such as cycles or repeated sequences, then life (endurance) is stated in terms of the number of these increments endured prior to failure. Similar recognizable increments seldom, if ever, occur in service. Thus there is no direct relationship between a well-defined laboratory test life (endurance) and the ill-defined service operation life (endurance).

Mechanical reliability is technically defined as the probability that a given device, randomly selected from the population of all such nominally identical devices, will perform satisfactorily under certain specified service operation environmental and loading histories for at least an arbitrarily specified life (duration). However, this technical definition for reliability is seldom if ever practical (for reasons that will be evident later). Rather, it is customary to compute a lower 100(scp)% (one-sided) statistical confidence (tolerance) limit that allegedly bounds the actual value for the metric pertaining to the p\textsuperscript{th} percentile of the presumed conceptual two-parameter life (duration of failure) or strength (resistance) statistical distribution — where p is usually selected to be either 0.01 or 0.10 and the associated scp is usually selected to be 0.95 (See Supplemental Topic 8.A).

### 8.2 THE CONCEPTUAL STATISTICAL MODEL FOR FATIGUE FAILURE AND THE ASSOCIATED LIFE (ENDURANCE) AND STRENGTH (RESISTANCE) EXPERIMENT TEST PROGRAMS

We now present a conceptual statistical model for fatigue failure that ostensibly pertains to conventional laboratory fatigue data, but applies in concept to all mechanical modes of failure when appropriate analogies are made. Figure 8.1(a) depicts the PDF of the conceptual statistical distribution that consists of all possible replicate fatigue life datum values that could be generated by continually replicating the conduct of a quantitative CRD experiment test program (Chapter Two) with a fixed value \( s_a \)\textsuperscript{*} for its alternating stress amplitude \( s_a \). Note that the metric for this PDF is \( \log_e(fnc) \), where \( fnc \) connotes failure number of cycles. Note in addition that the traditional \( s_a - \log_e(fnc) \) curve pertains to the median fatigue life. (Thus this traditional \( s_a - \log_e(fnc) \) curve is properly denoted a \( s_a - \log_e(fnc(50)) \) curve.) The CDF that corresponds to the PDF in Figure 8.1(a) is depicted in Figure 8.1(b). It also establishes the values for \( \log_e(fnc) \) that pertain to other probabilities of failure (pf) of specific interest, e.g., to \( \log_e(fnc(0.01)) \) for \( pf = 0.01 \) and \( \log_e(fnc(0.99)) \) for \( pf = 0.99 \). In turn, a conceptual \( s_a - \log_e(fnc(pf)) \) fatigue failure model, Figure 8.1(c), is developed when the respective \( \log_e(fnc(pf)) \) values of specific interest are augmented with corresponding \( \log_e(fnc(pf)) \) values pertaining to other values for \( s_a \). However, extensive fatigue test experience clearly indicates that the variance of the conceptual statistical distribution that consists of all possible replicate \( \log_e(fnc) \) datum values increases markedly as \( s_a \)\textsuperscript{*} decreases, especially at long fatigue lives. Thus the presumption of homoscedasticity for a conceptual \( s_a - \log_e(fnc(pf)) \) fatigue failure model is statistically invalid — as it is for all other mechanical modes of failure that are life (duration to failure) dependent. In contrast, experience indicates that the vast majority of empirical
scatter bands faired by eye to datum values that pertain to various life (duration to failure) dependent modes of failure (including fatigue) are approximately parallel when their width is stated in terms of the stimulus level (alternating stress amplitude) as in Figure 8.1(d). Accordingly, the notion that the associated conceptual fatigue strength statistical distribution, Figure 8.1(e), has a homoscedastic variance is statistically credible. Finally, the conceptual $s_a$-$\log_e[fnc(pf)]$ fatigue failure model can be visualized as a surface in $s_a$ (stimulus level), $\log_e(fnc)$ (duration to failure), and $pf$ (probability of failure) space. See Figure 8.1(f).

**Figure 8.1(a).** An example life (endurance) experiment test program that involves testing replicate fatigue specimens at a specific alternating stress amplitude $s_a^*$. Each replicate test is ideally conducted until failure occurs (which may not always be practical). Note the improper reversal of independent and dependent (random) variables for the traditional method of plotting the median s-N curve for which we use the unconventional notation, $s_a$-$\log_e[fnc(50)]$ model, where $fnc(50)$ connotes median failure number of cycles.

**Figure 8.1(b).** Schematic of the conceptual fatigue life CDF associated with Figure 8.1(a). Note that the conceptual fatigue life CDF pertains to the probability of failure, $pf$, whereas reliability is stated in terms of the probability of survival, $ps$. 
**Figure 8.1(c).** A conceptual $s_a$-$\log_{e}[fnc(pf)]$ model that is based on a series of life (endurance) experiments conducted at several different values for the alternating stress amplitude $s_a^*$. Figure 8.1(a). The respective life (duration to failure) datum values exhibit marked heteroscedasticity.

**Figure 8.1(d).** Experience demonstrates that $s_a$-$\log_{e}(fnc)$ fatigue experiment test programs almost always generate datum values that can be modeled as homoscedastic when stated in terms of the metric $s_a$.

**Figure 8.1(e).** Schematic of the CRD pertaining to the conceptual fatigue strength distribution associated with Figure 8.1(d). The actual values for percentiles of this conceptual fatigue strength distribution, given the value for $\log_{e}(fnc^*)$ of specific interest, can be estimated by conducting a strength (resistance) experiment test program (Section 8.2B).
8.2A LIFE (ENDURANCE) EXPERIMENT TEST PROGRAM DATA

Mechanical mode of failure models typically exhibit (i) a PDF that is markedly skewed to the right given a linear metric for life (endurance) and (ii) a heteroscedastic variance that increases as the concomitant stimulus level decreases. A logarithmic transformation for the life (endurance) datum values tends to mitigate the PDF skewness problem somewhat and thus is almost always employed. Nevertheless, a proper statistical analysis must still account for (or avoid) the problem of heteroscedasticity of the life (endurance) datum values. A further complication arises when the mode of failure has an initiation threshold. Then the life (endurance) PDF is not only truncated, but each individual test in the experiment test program must also be suspended at some pre-determined practical value for life (endurance). Finally, there is the issue of statistically planned and arbitrary suspensions of individual tests.

Types of Life (Endurance) Data:

1. Data with No Censoring: All items tested (will eventually) fail and the respective life (endurance) datum values are individually recorded. (These datum values can be depicted schematically by an X as illustrated below.) We specifically employ the terminology no censoring to pertain to replicate datum values generated in a life (endurance) test program in which it was decided in advance that each test would be continued until its test item failed, regardless of the test duration required. Accordingly, no censoring occurs only when no mandatory, planned, or arbitrary test suspension is permitted (see 2. and 3. below).
2. Data with Type I Censoring (Planned Suspensions): Each test item in a life (endurance) experiment test program with Type I censoring either fails before some pre-determined practical value for test life (endurance), or the test is deliberately suspended at this pre-determined life (endurance). Note that the number of failures (and the associated number of Type I censored tests) is a random variable under continual replication of a life (endurance) experiment test program with Type I censoring. Accordingly, it is possible that a given life (endurance) experiment test program with Type I censoring will have no deliberately suspended tests, whereas a replicate experiment test program will have several tests that are deliberately suspended. Thus the appropriate statistical analysis for a life (endurance) experiment test program with ostensibly no censored datum values depends on whether Type I censoring would have been employed if one or more test items had indeed survived to the pre-determined practical value for life (endurance), viz., we must consider potential as well as actual Type I censoring. Finally, note that although each test in a life (endurance) test program could require suspension if its life (endurance) becomes excessive, as illustrated below all Type I censored tests must be deliberately suspended at the same pre-determined value for life (endurance).

3. Data with an Arbitrarily Suspended Test: A life (endurance) experiment test program should be terminated when the current test is suspended due to equipment failure. (We limit our subsequent pragmatic statistical analyses to pertain only to actual or potential Type I censoring.)

Remark: Test equipment failure during very long experiment test programs is surely foreseeable, and often likely. Test equipment failure problems can be mitigated in comparative experiment test programs by employing time blocks (Chapter Two).

Perspective: There are only four life (endurance) data types (illustrated schematically below): (i) It is only known that the test item life (endurance) is less than some recorded value. (ii) It is only known that the test item life (endurance) is bounded by two recorded values. (iii) It is only known that the test item life (endurance) is greater than some recorded value. (iv) The actual value for each respective test item life (endurance) datum value is known and recorded.
Discussion: Textbooks often discuss Type II (planned) censoring. It occurs when the life (endurance) experiment test program (itself) is suspended as soon as the \( j^{th} \) item fails. Type II censoring is seldom if ever practical in mechanical testing because it requires either the simultaneous testing of \( n_i \) nominally identical test items, each in a different allegedly identical test machine, or the use of a single test machine whose test fixture simultaneously exposes each of its \( n_i \) test items to allegedly identical test conditions. In contrast, Type I censoring typically pertains to sequential testing of \( n_i \) items using a single test machine with a single test specimen in its test fixture (and presuming that test conditions are invariant during the entire experiment test program).

8.2B STRENGTH (RESISTANCE) EXPERIMENT TEST PROGRAM DATA

Strength (resistance) experiment test programs are generally more practical than life (endurance) experiment test programs for long life (endurance) reliability applications, but strength (resistance) experiment test programs typically require considerably more test time to conduct. Accordingly, more attention should be given to the statistical planning of strength (resistance) experiment test programs. In particular, specimen allocation strategies should be employed that generate more precise estimates of the actual values for the medians of the presumed conceptual strength (resistance) distributions.

Remark: The CDF’s for all conceptual statistical distributions used in the analysis of life (endurance) and strength (resistance) data differ markedly only at their extreme percentiles where, unfortunately, there is no practical way to distinguish one CDF statistically from another. Accordingly, strength (resistance) estimates have traditionally pertained only to medians. In contrast, life (endurance) estimates often naively pertain to an extreme percentile of some presumed conceptual statistical distribution.

Strength (Resistance) Experiment Test Program Strategies

1. The Conventional (Large Sample) Up-and-Down Test Method: The first test in the conventional (large sample) up-and-down test method is conducted at the guesstimated stimulus level (e.g., the alternating stress amplitude in fatigue) that hopefully corresponds to the actual value for the median of the presumed normally distributed conceptual strength (resistance) distribution. Then, if failure occurs, the second test is conducted at a decreased stimulus level. But if the first test item does not fail, then the second test is conducted using a increased stimulus level. This up-and-down strategy continues until all available test items have been tested. The increment used to decrease or increase the stimulus level is fixed, ideally selected to be equal to the standard deviation of the presumed conceptual strength (resistance) distribution (whose actual value, unfortunately, can only be guesstimated on the basis of preliminary testing or prior experience.)

The conventional (large sample) up-and-down test method is illustrated below for the example data given in ASTM D-3029-90, Standard Test Methods for Impact Resistance of Flat, Rigid Plastic Specimens by Means of a Tup (Falling Weight), in which the weight of the tup (dart) is decreased or increased depending on the prior test outcome. An analogous test method pertains to a tup (dart) of fixed weight dropped from a height that is either increased or decreased depending on the prior test outcome.
1(a). The Two-Point Strategy: Maximum likelihood analysis of the ASTM D-3029-90 example data indicates that the five tests conducted with a dart weight of 6 Kg have a negligible effect on the statistical estimate of the actual value for the median of the conceptual impact resistance distribution and its precision. Accordingly, the two-point strategy was proposed to alleviate this statistically ineffective test item allocation problem. In the two-point strategy, the conventional up-and-down test method strategy is terminated as soon as both O’s and X’s have been observed at two distinct stimulus levels. Then the test program continues with all future tests being conducted at only these two levels. Considering the ASTM D-3029-90 up-and-down test method example data, the two-point strategy gives

<table>
<thead>
<tr>
<th>test number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight, kg</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Test outcome code: X = failure; O = non-failure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

in which the three specimens not tested at 6 kg (indicated by dashes) are much more effectively tested at either 7 or 8 kg. (In fact, as discussed later, an even more statistically effective tup (dart) weight can be established using the minimum variance strategy.)

2. The Conventional Small Sample Up-and-Down Test Method: An experiment test program that is conducted following the conventional (large sample) up-and-down test method can involve as few as twenty specimens (but should involve at least sixty to eighty specimens when approximate estimation expressions are used), whereas an experiment test program that is conducted following the conventional small sample up-and-down test method usually involves only five to ten specimens. However, the actual value for the standard deviation of the presumed conceptual strength (resistance) distribution must be presumed known to accomplish this massive sample size reduction (and only the actual value for the median of the presumed conceptual strength (resistance) distribution is estimated in statistical analysis). The spacing between successive stimulus levels is also fixed for the conventional small sample up-and-down test method. It is ostensibly equal to the presumed known (guestimated) actual value for the standard deviation. The nominal sample size for a conventional small sample up-and-down test method program is established by counting only the last response of the beginning sequence of like responses, and all of the remaining responses. For example, given a conventional small sample up-and-down experiment test program outcome of O O O O X O X O, the nominal sample size is 5. Although, when employed to conduct fatigue tests, the conventional small sample up-and-down test method strictly pertains to estimating the actual value for the median of the conceptual endurance limit (threshold) distribution, it can also be used in an ad hoc manner to estimate the actual value for the median of the conceptual fatigue strength distribution at very long fatigue lives, e.g., at $10^7$ to $10^8$ alternating stress cycles.
\( \gamma(n) \) **Minimum Variance Strategy**: As mentioned above, the spacing (increment) between successive stimulus levels is fixed in the conventional small sample up-and-down test method. However, the minimum variance methodology indicates that the stimulus level that most effectively increases the statistical precision of the resulting estimate of the actual value for the median of the presumed conceptual strength (resistance) distribution changes from test to test. Thus as soon as a reasonably precise estimate of the actual value for the median of the presumed conceptual strength (resistance) distribution has been established, all subsequent stimulus levels should be selected to maximize the statistical precision of the resulting estimate of the actual value for the median of the presumed conceptual strength (resistance) distribution. This strategy is also effective when a conventional small sample up-and-down test method program is augmented by testing a few additional items to improve the statistical precision of the final estimate of the actual value for the median of the presumed conceptual strength (resistance) distribution.

**Preliminary Strength (Resistance) Experiment Test Program Strategies [Little (1990)]**

When the number of items available for testing is quite limited or costly, or when time available for testing is very limited, and when there is little if any information available to guestimate the initial stimulus level to begin up-and-down testing, the strategies used in preliminary testing should be particularly efficient and effective. These strategies should either markedly reduce the number of test items or the amount of test time required to home in on the actual value for the median of the presumed conceptual strength (resistance) distribution, thereby permitting the remaining test items to be allocated to stimulus levels that are statistically more effective in estimating the actual value for the median of the presumed conceptual strength (resistance) distribution. Two examples follow which pertain to an (unpublished) unique bolt pull-through fatigue experiment test program on a composite material that is markedly stronger in the longitudinal direction than the transverse direction. The respective stimulus levels are stated in terms of force (load) rather than stress because one of the main objectives of the experiment test program was to establish the mode of failure and the associated governing stress expression(s) for each different composite material tested. The third example strategy is intended to reduce the number of tests required to obtain a change in response from X to O or vice versa.

**Example One: Run-Up Preliminary Test Strategy, Table 8.1.** When the number of test items is severely limited or are extremely costly, the run-up preliminary test strategy is appropriate in which a single test item is used to begin to home in on the actual value for the median of the presumed conceptual strength (resistance) distribution. The first test is conducted with its stimulus level guestimated to be well below the actual value for the median of the presumed conceptual strength (resistance) distribution. If the test item does not fail before the pre-determined practical value for life (endurance), it is thoroughly examined and then re-tested at an increased stimulus level. This strategy continues until the test item eventually fails. (If sample size, test time, and cost constraints permit, this failure stimulus level could be replicated and the data for the run-up item relegated to ancillary information status because of potential damage or potential coaxing effects.) The fixed increment between the successive stimulus levels for these run-up tests should be relatively large, so that relatively few run-up test increments are required to cause the first failure. Accordingly, this run-up increment must be markedly reduced before beginning to employ the up-and-down test method strategy. The half-interval strategy is conveniently used to reduce this increment while continuing to home in on the actual value for the median of the presumed conceptual strength (resistance) distribution, viz., the increment between successive tests is reduced by one-half of the difference in adjacent stimulus levels with X and O outcomes. Once the stimulus level increment has been reduced to its pre-selected minimum level using the half-interval strategy, the up-and-down test method strategy is adopted and continued until a pre-selected nominal sample size is reached (or until
some other criterion is satisfied). Then the minimum variance strategy should use to allocate the remaining specimens to their statistically most effective stimulus levels.

Table 8.1 — Example One: Run-Up Preliminary Test Strategy

<table>
<thead>
<tr>
<th>Test Number</th>
<th>Alternating Force ($f_a$), lbs</th>
<th>Test Outcome</th>
<th>Test Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>50</td>
<td>O</td>
<td>Test specimen did not fail. Thus this unfailed specimen is re-tested with $f_a = 50 + 25$</td>
</tr>
<tr>
<td>1b</td>
<td>75</td>
<td>O</td>
<td>Test specimen did not fail. Thus this unfailed specimen is re-tested with $f_a = 75 + 25$</td>
</tr>
<tr>
<td>1c</td>
<td>100</td>
<td>O</td>
<td>Test specimen did not fail. Thus this unfailed specimen is re-tested with $f_a = 100 + 25$</td>
</tr>
<tr>
<td>1d</td>
<td>125</td>
<td>X</td>
<td>Test specimen failed. Thus the half-interval strategy begins. Accordingly the next specimen is tested with $f_a = 125 - (25/2)$</td>
</tr>
<tr>
<td>2</td>
<td>112.5</td>
<td>O</td>
<td>Test specimen did not fail. Thus the next specimen is tested with $f_a = 112.5 + (12.5/2)$. Note that this new increment was approximately equal to its pre-selected minimum practical value. We now begin the up-and-down test method strategy</td>
</tr>
<tr>
<td>3</td>
<td>118.75</td>
<td>X</td>
<td>Test specimen failed. Thus the next specimen is tested with $f_a = 118.75 - 6.25$</td>
</tr>
<tr>
<td>4</td>
<td>112.5</td>
<td>O</td>
<td>Test specimen did not fail. Thus the next specimen is tested with $f_a = 112.5 + 6.25$</td>
</tr>
<tr>
<td>5</td>
<td>118.75</td>
<td>X</td>
<td>Test specimen failed. The minimum variance strategy should be used for all remaining test specimens</td>
</tr>
</tbody>
</table>

**Summary:** Test specimen 1: run-up strategy, Test specimens 1 to 3: half-interval strategy, Test specimens 3 to 5: up-and-down test method strategy. Remaining test specimens: minimum variance strategy

**Example Two: Run-Down Preliminary Test Strategy, Table 8.2.** The run-down preliminary test strategy is recommended to begin to home in on the actual value for the median of the presumed conceptual strength (resistance) distribution when ample test items are available and it is desired to limit test time as much as practical. The life (endurance) datum values can also be used to decide when the change to the half-interval strategy is appropriate. For example, the fatigue test specimen tested at $f_a = 75$ lbs in Table 8.2 failed just before $10^7$ load cycles. Hence the third specimen could have been tested with an alternating force $f_a$ equal to 62.5 lbs rather than strictly following the run-down strategy.

This experiment test program followed the Example One experiment test program and employed the same composite material and test fixture, but the test specimen was oriented in the transverse direction rather than in the longitudinal direction. Thus the initial alternating force of 100 lbs was expected to be much larger than the actual value for the median of its conceptual failure load statistical distribution.
Table 8.2 — Example Two: Run-Down Preliminary Test Strategy

Preliminary test conduct choices:
- initial alternating force: 100 lbs
- run-down increment: 25 lbs
- minimum practical interval value: 5% of est($f_a(50)$)

Test outcome code:
- X = failure before $10^7$ cycles. O = run-out (did not fail)

<table>
<thead>
<tr>
<th>test number</th>
<th>alternating force ($f_a$)</th>
<th>Test Outcome</th>
<th>Test Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>X</td>
<td>Test specimen failed. Thus the next specimen is tested with $f_a = 100 - 25$</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>X</td>
<td>Test specimen failed. Thus the next specimen is tested with $f_a = 75 - 25$</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>O</td>
<td>Test specimen did not fail. Thus the half-interval strategy begins. Accordingly the next specimen is tested with $f_a = 50 + (25/2)$</td>
</tr>
<tr>
<td>4</td>
<td>62.5</td>
<td>X</td>
<td>Test specimen failed. Thus the next specimen is tested with $f_a = 62.5 - (12.5/2)$</td>
</tr>
<tr>
<td>5</td>
<td>56.25</td>
<td>O</td>
<td>Test specimen did not fail. Thus the next specimen is tested with $f_a = 56.25 + (6.25/2)$. Note that this new increment was approximately equal to its pre-selected minimum practical value. Accordingly, we now begin the up-and-down test method strategy</td>
</tr>
<tr>
<td>6</td>
<td>59.375</td>
<td>X</td>
<td>Test specimen failed. Thus the next specimen is tested with $f_a = 59.375 - 3.125$</td>
</tr>
<tr>
<td>7</td>
<td>56.25</td>
<td>O</td>
<td>Test specimen did not fail. Thus the next specimen is tested with $f_a = 56.25 + 3.125$</td>
</tr>
<tr>
<td>8</td>
<td>59.375</td>
<td>X</td>
<td>Test specimen failed. The minimum variance strategy will be used for all remaining test specimens</td>
</tr>
</tbody>
</table>

Summary: Test specimen 1 to 3: run-down strategy. Test specimens 3 to 6: half-interval strategy. Test specimens 6 to 8: up-and-down test method strategy. Remaining test specimens: minimum variance strategy

Example Three: Wide-Spacing Preliminary Test Strategy [Little and Thomas (1993)]. When sufficient prior testing has been conducted to have a reasonably precise estimate of the actual value for the standard deviation of the presumed conceptual strength (resistance) distribution, an effective preliminary test strategy is to use a stimulus level spacing equal to two times the estimated actual value for the standard deviation in the run-up or run-down tests used to begin the experiment test program, and then to change to the conventional small sample up-and-down test method strategy with a stimulus level spacing equal to the estimated actual value for the standard deviation as soon as the test outcome changes character from X to O, or vice versa. Then, after a few up-and-down tests have been conducted, the statistical precision of the final estimate of the actual value for the median of the presumed conceptual strength (resistance) distribution can be improved by subsequently employing the minimum variance strategy to allocate the remaining specimens to their statistically most effective stimulus levels.)
8.3 CONCEPTUAL STATISTICAL DISTRIBUTIONS FOR MODELING THE OUTCOMES OF LIFE (ENDURANCE) AND STRENGTH (RESISTANCE) EXPERIMENT TEST PROGRAMS

The outcomes of life (endurance) and strength (resistance) experiment test programs are modeled by conceptual statistical distributions that are actually mathematical abstractions. The conceptual statistical distribution that is properly employed in modeling, if any, is unknown for mechanical modes of failure. Moreover, even when some conceptual statistical distribution is widely alleged to model a given mechanical mode of failure, alternative conceptual statistical distributions should always be employed in analysis for comparative purposes. It is therefore important to understand the similarities and the differences among alternative conceptual life and strength statistical distributions.

8.3A CONCEPTUAL LIFE (ENDURANCE) DISTRIBUTIONS

The two-parameter conceptual log<sub>e</sub>-normal and Weibull distributions are most commonly used to model life (endurance) datum values. But log<sub>e</sub>-normal datum values are seldom analyzed directly. Rather, the natural logarithms of the respective life (endurance) datum values are first calculated and then these transformed datum values are presumed to have been randomly selected from a (two-parameter) conceptual normal distribution. In turn, following appropriate analysis, the exponentials (antilogs) of the particular estimated values of specific interest are evaluated. Weibull datum values can be analyzed similarly, viz., the natural logarithms of the respective life (endurance) datum values are first calculated and then these transformed datum values are presumed to have been randomly selected from a conceptual (two-parameter) smallest-extreme-value distribution. In turn, following appropriate analysis, the exponentials (antilogs) of the particular estimated values of specific interest are evaluated.

The Conceptual Two-Parameter Weibull Distribution: The CDF for the conceptual two-parameter Weibull can be written as

\[ F(fnc) = 1 - \exp\left(-\frac{fnc}{cdpl}\right)^{cdp2} \]

in which \( fnc \), the failure number of cycles, is the life (endurance) metric. The inverse CDF expression is

\[ y = \log_e\left(-\log_e\left[1 - F(fnc)\right]\right) = cdpl \cdot \left[\log_e(fnc) - \log_e(cdpl)\right] = a + b \cdot \log_e(fnc) \]

Accordingly, the conceptual two-parameter Weibull CDF plots as a straight line on Weibull probability paper with ordinate \( y \) and abscissa \( \log_e(fnc) \), where \( y \) is a function of \( pf = F(fnc) \) given by the inverse CDF expression.

The Corresponding Conceptual (Two-Parameter) Smallest-Extreme-Value Distribution: The CDF for the corresponding conceptual (two-parameter) smallest-extreme-value distribution is written as

\[ F[\log_e(fnc)] = 1 - \exp\left[-\exp\left(\frac{\log_e(fnc) - clp}{exp}\right)\right] \]

in which \( \log_e(fnc) \) is the life (endurance) metric rather than \( fnc \). The corresponding inverse CDF expression is
\[ y = \log_e \left( -\log_e \left( 1 - F[\log_e(fnc)] \right) \right) = \frac{\log_e(fnc) - clp}{csp} = a + b \cdot \log_e(fnc) \]

Accordingly, when \( csp = (1/cdp2) \) and \( clp = \log_e(cdp1) \) the conceptual (two-parameter) smallest-extreme-value distribution pertaining to the (transformed) random variable \( \log_e(FNC) \) is identical to the conceptual two-parameter Weibull distribution pertaining to the random variable \( FNC \). Observe, however, that the \( clp \) and the \( csp \) for the conceptual (two-parameter) smallest-extreme-value distribution have intuitive geometric interpretation because its life (endurance) metric \( \log_e(fnc) \) is plotted along a linear abscissa.

The Conceptual Three-Parameter Log_e-Normal and Weibull Distributions: The addition of a third conceptual parameter in the conceptual two-parameter log_e-normal and Weibull distributions provides their respective CDF’s more curve-fitting versatility relative to describing life (endurance) datum values and thus has led to the unwarranted popularity of three-parameter distributions. The third conceptual parameter, the so-called conceptual minimum life parameter, \( cmplp \), is absolutely fictitious relative to its physical interpretation. Accordingly, neither three-parameter distribution can be recommended for use in mechanical reliability applications.

Perspective: It is important to understand that no conceptual statistical distribution, regardless of the number or the nature of its parameters, ever exactly models life (endurance) or strength (resistance) datum values for any mechanical mode of failure. All conceptual statistical distributions are mathematical abstractions.

8.3B CONCEPTUAL LIFE (ENDURANCE) DISTRIBUTIONS WITH A CONCOMITANT INDEPENDENT VARIABLE

Recall that linear regression analysis can be viewed as augmenting the \( clp \) of a conceptual (two-parameter) normal distribution with one or more concomitant independent variables. This notion can be extended to construct any \( s_{\alpha} \cdot \log_e[fnc(pf)] \) model of specific interest. For example, a quadratic \( s_{\alpha} \cdot \log_e[fnc(pf)] \) model with a conceptual (two-parameter) smallest-extreme-value distribution for \( \log_e(fnc) \) datum values can be expressed as

\[
F[\log_e(fnc)] = 1 - \exp \left[ -\exp \left( \frac{\log_e(fnc) - clp0 - clp1 \cdot s_{\alpha} - clp2 \cdot s_{\alpha}^2}{csp} \right) \right]
\]

However, this \( s_{\alpha} \cdot \log_e[fnc(pf)] \) model is not statistically credible for long-life fatigue applications because its homoscedastic \( csp \) does not increase appropriately as \( \log_e[fnc(50)] \) increases. Although various \( s_{\alpha} \cdot \log_e[fnc(pf)] \) models with augmented conceptual (two-parameter) fatigue life distributions have been proposed to account for the pronounced heteroscedasticity of \( s_{\alpha} \cdot \log_e(fnc) \) datum values at long fatigue lives, none can be recommended. Rather, because almost all published \( s_{\alpha} \cdot \log_e(fnc) \) data have faiared scatter bands that are (approximately) parallel to their associated faiared \( s_{\alpha} \cdot \log_e[fnc(50)] \) curves, we opt to employ (adopt) either a straight line or a parabolic \( s_{\alpha} \cdot \log_e[fnc(pf)] \) model with a homoscedastic conceptual (two-parameter) fatigue strength distribution. For example, a parabolic \( s_{\alpha} \cdot \log_e[fnc(pf)] \) model with a homoscedastic conceptual (two-parameter) smallest-extreme-value (skewed) fatigue strength distribution can be expressed as
A directly analogous parabolic $s_a \cdot \log_{10}[fnc(pf)]$ model employs a homoscedastic conceptual (two-parameter) normal (symmetrical) fatigue strength distribution. The respective $s_a \cdot \log_{10}[fnc(50)]$ curves pertaining to these two straight line or parabolic $s_a \cdot \log_{10}[fnc(pf)]$ models can employ either a linear or a logarithmic metric for $s_a$. Thus there are eight alternative (candidate) $s_a \cdot \log_{10}[fnc(pf)]$ models for each new mechanical reliability application where experience does not dictate (suggest) the proper choice.

### 8.3C CONCEPTUAL STRENGTH (RESISTANCE) DISTRIBUTIONS

The conceptual (two-parameter) logistic distribution is sometimes used to analyze strength (resistance) datum values. Its symmetrical PDF is very similar to the conceptual (two-parameter) normal distribution PDF, but its longer tails have greater probability content. The conceptual (two-parameter) logistic distribution CDF can be written as

$$F(s) = 1 - \exp^{-\exp \left( \frac{s - c\ln(p)}{csp} \right)^{-1}}$$

Because of its longer tails, it serves as an excellent adjunct to the conceptual (two-parameter) normal distribution for purposes of comparing the respective lower 100($scp$)% (one-sided) statistical confidence limits based on alternative statistical models, just as does the conceptual (two-parameter) smallest-extreme-value distribution, which is skewed to the left and its mirror image, the conceptual (two-parameter) largest-extreme-value distribution, which is skewed to the right. This conceptual (two-parameter) largest-extreme-value distribution can be written as

$$F(s) = \exp^{-\exp \left( \frac{s - c\ln(p)}{csp} \right)}$$

These four CDF’s, with both linear and logarithmic abscissa scales, generate eight alternative analyses for estimating the actual value for the median (or other percentiles) of the presumed conceptual strength (resistance) distribution of specific interest.

### Exercise Set 1

(These exercises are intended to familiarize you with the PDF’s and the corresponding CDF’s for the alternative conceptual (two-parameter) life (endurance) and strength (resistance) distributions presented in this section.)

First, given the analytical expression for each of the CDF’s in Ex.’s 1 through 6 below, develop analytical expressions for the PDF, mean, median, mode and variance. (See Supplemental Topic 8.D to check your results.) Then tabulate $x, f(x)$, and $y(p)$ for $F(x) = p = 0.01, 0.02, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.98$, and $0.99$. Finally, using your tabulation of $p$ and $y(p)$, construct a sheet of probability paper for each of these conceptual (two-parameter) life (endurance) and strength (resistance) distributions with both $p$ and $y(p)$ plotted along its respective non-linear and linear ordinates and with either $x$ or $\log_{10}(x)$ as appropriate plotted along its linear abscissa.
1. Given the conceptual (two-parameter) normal distribution CDF,

\[ F(x) = \frac{1}{\sqrt{2\pi \cdot \text{csp}}} \int_{-\infty}^{x} \exp \left[ -\frac{1}{2} \left( \frac{u - \text{clp}}{\text{csp}} \right)^2 \right] du \]

in which \( u \) is the dummy variable of integration. Plot its PDF when the actual value for its mean is equal to 100 and the actual value for its standard deviation is (a) 10 and (b) 40. In turn, plot its corresponding CDF's on (a) normal probability paper and (b) rectilinear graph paper.

2. Given the conceptual (two-parameter) smallest-extreme-value distribution CDF,

\[ F(x) = 1 - \exp^{-\exp \left( \frac{x - \text{clp}}{\text{csp}} \right)} \]

Plot its PDF when the actual value for its mean is equal to 100 and the actual value for its standard deviation is (a) 10 and (b) 40. In turn, plot the corresponding CDF's on (a) smallest-extreme-value probability paper and (b) rectilinear graph paper.

3. Given the conceptual (two-parameter) largest-extreme-value distribution CDF,

\[ F(x) = \exp^{-\exp \left( \frac{x - \text{clp}}{\text{csp}} \right)} \]

Plot its PDF's when the actual value for its mean is equal to 100 and the actual value for its standard deviation is (a) 10 and (b) 40. In turn, plot the corresponding CDF's on largest-extreme-value probability paper and on rectilinear graph paper.

4. Given the conceptual (two-parameter) logistic distribution CDF,

\[ F(x) = \left[ 1 + \exp \left( \frac{x - \text{clp}}{\text{csp}} \right) \right]^{-1} \]

Plot its PDF when the actual value for its mean is equal to 100 and the actual value for its standard deviation is (a) 10 and (b) 40. In turn, plot the corresponding CDF's on (a) logistic probability paper and (b) rectilinear graph paper.

5. Given the conceptual two-parameter Weibull distribution CDF,

\[ F(x) = 1 - \exp \left( -\frac{x}{\text{cdpl}} \right)^{\text{cdp2}} \]

Plot its PDF when the actual value for its mean is equal to 100 and the actual value for its standard deviation is (a) 10 and (b) 40. [Run microcomputer program WBLCDPS to verify that \( \text{cdpl} = 104.30376808 \) and \( \text{cdp2} = 12.15343419 \) for (a) and that \( \text{cdpl} = 112.45635006 \) and \( \text{cdp2} = 2.69562125 \) for (b)]. In turn, plot the corresponding CDF's on Weibull probability paper with a linear \( \log_e(x) \) scale along its abscissa, and on rectilinear graph paper.
6. Given the conceptual two-parameter loge-normal distribution CDF

\[ F(x) = \frac{1}{\sqrt{2\pi \cdot cdp2}} \int_{0}^{x} \frac{1}{u} \exp \left\{ -\frac{1}{2} \left[ \frac{\log_e(u) - \log_e(cdp1)}{cdp2} \right]^2 \right\} du \]

in which \( u \) is the dummy variable of integration. Plot its PDF when the actual value for its mean is equal to 100 and the actual value for its standard deviation is (a) 10 and (b) 40. [Run microcomputer program LNORCDPS to verify that \( cdp1 = 99.50371902 \) and \( cdp2 = 0.09975135 \) for (a) and that \( cdp1 = 92.84766909 \) and \( cdp2 = 0.38525317 \) for (b)]. In turn, plot the corresponding CDF’s on (b) loge-normal probability paper with a linear loge(x) scale along its abscissa, and on rectilinear graph paper.

**Exercise Set 2**

(These exercises supplement Exercise Set One and are intended to make specific comparisons between alternative conceptual life (endurance) and strength (resistance) CDF’s to enhance your perspective. For each of the required plots, let \( p \) cover the range from 0.001 to 0.999 using the values enumerated in Exercise Set One.)

1. Given a conceptual (two-parameter) smallest-extreme-value distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s on normal probability paper. Then given a conceptual (two-parameter) normal distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s on smallest-extreme-value probability paper.

2. Given a conceptual (two-parameter) largest-extreme-value distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s on normal probability paper. Then given a conceptual (two-parameter) normal distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s on largest-extreme-value probability paper.

3. Given a conceptual (two-parameter) logistic distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s on normal probability paper. Then given a conceptual (two-parameter) normal distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s on logistic probability paper.

4. Given a conceptual two-parameter Weibull distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s on loge-normal probability paper with a linear loge(x) scale along its abscissa. Then given a conceptual two-parameter loge-normal distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s on Weibull probability paper with a linear loge(x) scale along its abscissa.

5. Given a conceptual two-parameter Weibull distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s (a) on normal probability paper, (b) on smallest-extreme value probability paper, and (c) on largest-extreme-value probability paper. Then given a conceptual two-parameter loge-normal distribution with the actual value for its mean equal to 100 and the actual values for its standard deviation equal to (a) 10 and (b) 40, plot its CDF’s (a) on normal probability paper, (b) on smallest-extreme value probability paper, and (c) on largest-extreme-value probability paper.
8.4 QUANTITATIVE ANALYSIS FOR THE OUTCOMES OF LIFE (ENDURANCE) AND STRENGTH (RESISTANCE) EXPERIMENT TEST PROGRAMS

Only the maximum likelihood (ML) methodology is sufficiently versatile to deal with the outcomes of both statistically planned and ad hoc experiment test programs. Recall, however, that the fundamental statistical abstraction of a continually replicated experiment test program is obscure for ad hoc experiment test programs. ML methodology, coupled with its associated likelihood ratio test, provides the backbone of mechanical reliability analysis. We present several examples of quantitative maximum likelihood analyses in this section.

8.4A QUANTITATIVE MAXIMUM LIKELIHOOD ANALYSIS

Likelihood is analogous to probability. Consider a random variable $X$ whose PDF is more explicitly stated as $f(x \text{ given the actual cdp values})$. Then, for any collection of $n_{dv}$ independent replicate realization values of $X$, the joint PDF $g$ for these $x_i$'s is stated as

$$g(x_1, x_2, ..., x_{n_{dv}} \text{ given the actual cdp values}) = \prod_{j=1}^{n_{dv}} f(x_j \text{ given the actual cdp values})$$

Note that the $x_i$'s are analytical parameters and that the actual cdp values must be known (numbers) to evaluate this expression. In contrast, for the algebraically identical likelihood expression, the $x_i$'s are datum values, viz., known (numbers), whereas the actual cdp values are unknown and therefore must be treated as analytical parameters. Accordingly, in maximum likelihood (ML) analyses pertaining to replicate life (endurance) datum values, the actual csp values are estimated by (analytically or numerically) establishing the particular set of est(cdp's) that maximizes the likelihood (probability) of obtaining (observing) the associated experiment test program datum values.

Given datum values randomly selected from a continuous conceptual two-parameter statistical distribution whose metric range is not restricted by its parameter values, maximum likelihood estimators are asymptotically unexcelled, viz., ML estimators asymptotically (i) are unbiased, (ii) have minimum variance, and (iii) are normally distributed. But, for mechanical reliability experiment test programs with practical sizes, this extraordinary statistical behavior seldom prevails. Rather, the statistical behavior of each ML estimator of specific interest must either be deduced analytically or developed empirically by simulation. See Supplemental Topics 8.A through 8.E.

Because it is analytically expedient to work with a sum rather than a product, it is traditional in maximum likelihood analysis to maximize $\log_e(\text{likelihood})$ rather likelihood. Thus maximum likelihood estimation for statistical models with $n_{cp}$ conceptual parameters traditionally involves a search over $n_{cp}$-dimensional conceptual parameter space for the global maximum of the $\log_e(\text{likelihood})$. Given the maximum likelihood estimates of the actual values for the respective $c_p$'s, the second partial derivatives of the $\log_e(\text{likelihood})$ expression, evaluated at these ML est($c_p$) values, can be used to compute the associated $n_{cp}$ by $n_{cp}$ estimated asymptotic covariance matrix. In turn, this ML estimated asymptotic covariance matrix can be used in conjunction with the propagation of variability methodology to compute, for example, asymptotic statistical confidence limits (intervals) that allegedly bound (include) the actual value for the metric that pertains to the $(pf)^{th}$ percentile of the presumed conceptual statistical model of specific interest. It should be understood, however, that a statistical estimate of the actual value for the metric pertaining to a very small value for $pf$, say 0.01 or smaller, is markedly dependent on the presumed statistical model, whereas a statistical estimate of the actual value for the metric pertaining to $pf$ equal to 50, the median, is relatively insensitive to the presumed statistical model.

Convergence problems for numerical maximization procedures increase markedly as the number of conceptual parameters $n_{cp}$ increases. Moreover, the $\log_e(\text{likelihood})$ surface in $n_{cp}$-dimensional
conceptual parameter space may be so flat in the vicinity of its apparent maximum that classical 100(\(\text{secp}\))% (two-sided) asymptotic statistical confidence intervals are too wide to be credible. Thus, while maximum likelihood analysis can be employed to estimate the actual values for a large number of conceptual statistical model parameters, common sense dictates that the number of est(cp’s) be kept to a minimum.

Bartlett’s likelihood ratio (LR) test statistic should be used to examine the statistical adequacy of each proposed conceptual statistical model, just as Snedecor’s central F test statistic should be used to test the statistical adequacy of a conceptual linear regression model. Suppose, for example, the adequacy of a conceptual two-parameter statistical model is to be tested versus an alternative conceptual three-parameter statistical model. Then the null hypothesis is that the conceptual two-parameter statistical model is correct, whereas the alternative hypothesis is that the conceptual three-parameter statistical model is correct. When ML analyses are conducted for these two alternative conceptual statistical models, the respective values for the maximized \(\log_e(\text{likelihood})\)’s establish the experiment test program data-based realization value for Bartlett’s LR test statistic, viz.,

\[
\text{Bartlett's LR test statistic} = -2 \cdot \log_e \left( \frac{\text{est}(\text{ML})_{n_{cp}=2}}{\text{est}(\text{ML})_{n_{cp}=3}} \right) = 2 \cdot \left\{ \log_e \left[ \text{est}(\text{ML})_{n_{cp}=3} \right] - \log_e \left[ \text{est}(\text{ML})_{n_{cp}=2} \right] \right\}
\]

Then, under continual replication of the experiment test program the respective data-based realization values for Bartlett’s LR test statistic asymptotically generate Pearson’s central \(\chi^2_{n_{cp}=3-2=1}\) conceptual sampling distribution. However, in general, Bartlett’s LR test statistic is expressed as

\[
\text{Bartlett's LR test statistic} = -2 \cdot \log_e \left( \frac{\text{est}(\text{ML})_{n_{cp}}}{\text{est}(\text{ML})_{n_{cp}+1}} \right) = 2 \cdot \left\{ \log_e \left[ \text{est}(\text{ML})_{n_{cp}+1} \right] - \log_e \left[ \text{est}(\text{ML})_{n_{cp}} \right] \right\}
\]

Then, under continual replication of the experiment test program the respective data-based realization values for Bartlett’s LR test statistic asymptotically generate Pearson’s central \(\chi^2_{n_{cp}=n_{cp}+1-n_{cp}}\) conceptual sampling distribution. (A numerical example of testing the adequacy of a proposed conceptual statistical model using Bartlett’s LR test is found in Supplemental Topic 8.F.)

Both the engineering and statistics literature is sadly deficient in terms of proposing alternative conceptual reliability models and then comparing the respective analyses. Nevertheless, all reasonable alternative conceptual reliability models should be employed in statistical analysis and an engineering decision should be made only after comparing the respective analyses.

Bartlett’s LR test statistic can also be used to compute asymptotic statistical confidence intervals that are viable alternatives to (and usually excel) classical asymptotic statistical confidence intervals. Accordingly, Bartlett’s LR test is the most versatile and useful statistical tool available in mechanical reliability analysis.

8.4B QUANTITATIVE MAXIMUM LIKELIHOOD EXAMPLES

We now present four example quantitative ML analyses. The first example pertains to ML analysis of datum values generated by a series of independent strength (resistance) tests that were conducted with the same stress (stimulus). The second example pertains to ML analysis of replicate datum values generated in a fatigue life experiment test program with actual \(\text{Type I}\) censoring and presumes that these datum values were randomly selected from a conceptual two-parameter Weibull distribution. These first two examples are intended to illustrate the fundamentals of (asymptotic) ML analysis. In contrast, the last two examples are intended to illustrate practical applications. The third example pertains to outcome
of a $s_t$-log$_e$(fn) experiment test program with no censoring, whereas the fourth example pertains to the outcome of a strength (resistance) experiment test program conducted using the up-and-down test strategy.

Comparative ML analyses will be presented in Section 8.5. The analyses can be used either to establish statistically whether treatment effects (or batch-to-batch effects) exist, or to establish whether the presumed conceptual model is statistically adequate to explain the observed experiment test program datum values.

**Example One: ML Analysis for the Outcome of a Strength (Resistance) Experiment Test Program.** Consider the outcome of a strength (resistance) experiment test program that consists of $n_{st}$ independent strength (resistance) tests conducted at the same stress (stimulus) level. Since each test item has a priori the same (unknown) probability $ps$ of surviving a pre-determined life $l^*$ (endurance $e^*$), each of the respective strength (resistance) tests is statistically viewed as being a binomial trial. Next, let $y = 1$ connote that the test item survived some pre-determined life $l^*$ (endurance $e^*$) and let $y = 0$ connote that the test item failed prior to life $l^*$ (endurance $e^*$). Then, the conceptual joint PDF for the respective $n_{st}$ independent strength (resistance) tests can be written as

$$
\text{conceptual joint PDF} = \frac{n_{st}!}{(n_{st} - n_{ts})!} \prod_{i=1}^{n_{st}} ps^{y_i} \cdot (1 - ps)^{(1 - y_i)} = \frac{n_{st}!}{(n_{ts})!(n_{st} - n_{ts})!} \cdot ps^{n_{ts}} \cdot (1 - ps)^{(n_{st} - n_{ts})}
$$

where $n_{ts}$ is the number of test items that survived the pre-determined life $l^*$ (endurance $e^*$), viz., the number of $y_i$'s that are equal to 1, and $ps$ is the fixed probability of surviving for at least the pre-determined life $l^*$ (endurance $e^*$) in each respective reliability test. This conceptual joint PDF is subsequently re-interpreted, without analytical change, as the likelihood expression, viz.,

$$
\text{likelihood} = \frac{n_{st}!}{(n_{ts})!(n_{st} - n_{ts})!} \cdot ps^{n_{ts}} \cdot (1 - ps)^{(n_{st} - n_{ts})}
$$

Although this product expression for likelihood is directly amenable to numerical maximization, analytical maximization procedures have traditionally dominated. Accordingly, the product expression for likelihood is converted to a summation expression for log$_e$(likelihood), viz.,

$$
\log_e(\text{likelihood}) = n_{ts} \cdot \log_e(ps) + (n_{st} - n_{ts}) \cdot \left[ \log_e(1 - ps) \right]
$$

in which the non-essential factorial expression has been ignored. (Recall that the natural logarithm is a monotonic transformation. Thus the likelihood and the log$_e$(likelihood) reach their respective maximums simultaneously.)

The maximum value of the log$_e$(likelihood) is obtained by equating the derivative of the log$_e$(likelihood) expression with respect to $ps$ equal to zero, viz.,

$$
\frac{d(\log_e(\text{likelihood}))}{d(ps)} = 0 = \frac{n_{ts}}{ps} - \frac{n_{st} - n_{ts}}{1 - ps}
$$

which, when solved analytically for $ps$, yields the ML estimate

$$
\text{ML est}(ps) = \frac{n_{ts}}{n_{st}}
$$
This ML estimate clearly agrees with our intuition. Nevertheless, it is criticized for allowing estimates of the actual value for \( ps \) to be equal to either zero or one.

We now establish an intuitive 100\((scp)\)% (two-sided) statistical confidence interval that allegedly includes the actual value for the fixed probability of survival, \( ps \), in each respective reliability test. First, however, we revise our notation to conform to a generic 100\((scp)\)% (two-sided) statistical confidence interval pertaining to a series of binomial trials. Accordingly, \( n_{st} \) now becomes \( n_{hl} \), the number of binomial trials; \( n_{is} \) now becomes \( n_{fo} \), the number of favorable outcomes; and \( ps \) now becomes \( pfo \), the fixed probability of a favorable outcome in each independent binomial trial. The intuitive way to compute a 100\((scp)\)% (two-sided) statistical confidence interval that allegedly includes the actual value for the \( pfo \) is to compute two related probability values, one so low that the probability of observing \( n_{fo} \) or more favorable outcomes in \( n_{hl} \) binomial trials is equal to \( [(1 - scp)/2] \), and one so high that the probability of observing \( n_{fo} \) or fewer favorable outcomes in \( n_{hl} \) binomial trials is also equal to \( [(1 - scp)/2] \). Then the statistical confidence probability that the actual value of conceptual binomial probability \( pfo \) lies in the interval from \( pfo_{\text{low}} \) to \( pfo_{\text{high}} \) is equal to \( 1 - [(1 - scp)/2] - [(1 - scp)/2] \) = \( scp \).

Microcomputer program IBPSCL (Intuitive Binomial Probability Statistical Confidence Interval) calculates this intuitive 100\((scp)\)% (two-sided) statistical confidence interval as [0.4439, 0.9748] when the \( scp \) is selected to be equal to 0.95. For comparison, the corresponding LR-based 95\% (two-sided) asymptotic statistical confidence interval (Figure 8.2) is [0.5006, 0.9636].

\begin{verbatim}
C> IBPSCL
Input the number of binomial trials of specific interest
10
Input the number of favorable outcomes of specific interest
8
Input the statistical confidence probability of specific interest in per cent (integer value)
95
The intuitive 95\% (two-sided) statistical confidence interval that allegedly includes the actual value for the pfo is [0.4439, 0.9748]
\end{verbatim}

\begin{verbatim}
C> LRIBPSCL
Input the number of binomial trials of specific interest
10
Input the number of favorable outcomes of specific interest
8
Input the statistical confidence probability of specific interest in percent (integer value)
95
The likelihood-ratio-based 95\% (two-sided) asymptotic statistical confidence interval that allegedly includes the actual value for the pfo is [0.5006, 0.9636]
\end{verbatim}

Remark: See Natrela (1963) for an analogous 100\((scp)\)% (two-sided) statistical confidence interval that is shorter than our intuitive 100\((scp)\)% (two-sided) statistical confidence interval.
Figure 8.2 The $\log_e$(likelihood) for $n_{st} = 10$ independent strength (resistance) tests with $n_{sl} = 8$ test items surviving these $n_{st}$ strength (resistance) tests, plotted versus all alternative possible values for the probability of surviving, $ps$. Microcomputer program LRBBPSCL employs the numerical likelihood ratio (LR) method to establish the respective limits of the 95% (two-sided) asymptotic statistical confidence interval that allegedly includes the actual value for $ps$. Recall that under continual replication of the experiment test program the respective realizations values for the maximized $\log_e$(likelihood) statistic asymptotically generate Pearson's central $\chi^2_{n_{sl}=1}$ conceptual sampling distribution.

An alternative 100(scp)% (two-sided) asymptotic statistical confidence interval can be computed using the ML estimate of the actual value for $pfo$ and its associated ML estimated asymptotic variance. In general, the individual elements of the ML estimated asymptotic covariance matrix prior to its inversion are negatives of the second and mixed partial derivatives of the $\log_e$(likelihood) expression with respect to the $n_{cdp}$ conceptual distribution parameters. However, for the discrete (one-parameter) conceptual binomial distribution, this estimated asymptotic covariance matrix has a single element, viz., the variance of the asymptotic conceptual sampling distribution that consists of all possible replicate realization values for the statistic ML $\text{est}(pfo)$. Thus $\text{ML \, \text{est}}\{\text{var}(\text{est}(pfo))\}$ is merely the negative of the reciprocal of the second derivative of the $\log_e$(likelihood) with respect to $pfo$. Accordingly, we first take the second derivative of the generic $\log_e$(likelihood) expression with respect to $pfo$, viz.,

$$\frac{d^2[\log_e(\text{likelihood})]}{d(pfo)^2} = \frac{n_{sl}}{pfo^2} - \frac{n_{sl} - n_{sh}}{(1 - pfo)^2}$$

and then we evaluate this second derivative expression in theory at $pfo$. However, because $pfo$ is unknown, this second derivative expression must be evaluated by substituting $\text{est}(pfo) = n_{fo}/n_{bf}$ for $pfo$. This substitution gives

$$\frac{d^2[\log_e(\text{likelihood})]}{d(pfo)^2} = -\frac{n_{bf}}{\text{est}(pfo) \cdot [1 - \text{est}(pfo)]}$$

Finally, we take the inverse of the negative of this second derivative expression to generate the desired estimated variance expression, viz.,
Comparing this estimated asymptotic variance expression to the corresponding exact variance expression developed in Supplemental Topic 3.B, we see that it becomes exact as \( n_{bt} \) increases without bound, viz., when \( \text{est}(pfo) = pfo \). In turn, using the conceptual (two-parameter) normal distribution asymptotic approximation to the conceptual (one-parameter) binomial distribution (Figure 3B.2), we compute the classical 95\% (two-sided) asymptotic statistical confidence interval that allegedly includes the actual value for the \( pfo \) as

\[
0.8 - 1.9600 \cdot \left( \frac{(0.8) \cdot (0.2)}{10} \right)^{1/2}, \quad 0.8 + 1.9600 \cdot \left( \frac{(0.8) \cdot (0.2)}{10} \right)^{1/2} = [0.5521, 1.0479]
\]

Since the maximum value of the \( pfo \) is one by definition, the conceptual (two-parameter) normal distribution asymptotic approximation to the (one-parameter) binomial distribution is clearly unacceptable for the small number of replicate tests in our numerical example.

Remark: This elementary example illustrates that alternative 100(\( scp \))% (two-sided) statistical confidence intervals can differ markedly. Accordingly, good statistical and engineering practice requires that alternative intervals be computed and compared.

Example Two: ML Analysis for the Outcome of a Fatigue Life Experiment Test Program with Type I Censoring, Presuming Weibull Datum Values. Consider the outcome of a fatigue life experiment test program such that \( n_f \) test items failed and \( n_s \) test items were Type I censored (suspended) at \( fnc = snc^* \). Suppose that it is presumed the respective datum values were randomly selected from a conceptual two-parameter Weibull distribution whose CDF is expressed as

\[
F(fnc) = 1 - \exp \left( \frac{fnc}{cdp2} \right)
\]

Then the corresponding conceptual PDF is expressed as

\[
f(fnc) = \frac{cdp2}{cdp1 \cdot cdp2} \cdot fnc \cdot (cdp2 - 1) \cdot \exp \left( \frac{-fnc}{cdp1} \right)
\]

The likelihood expression is the product of two likelihood expressions when the respective outcomes (realizations) life (endurance) tests are all mutually independent: (i) the likelihood expression that pertains to the respective \( fnc_i \)’s, and (ii) the likelihood expression that pertains to the \( n_s \) Type I censored tests. Accordingly,

\[
\text{likelihood} = \left( \frac{n_f + n_s}{n_f} \right)! \cdot n_f! \prod_{i=1}^{n_f} f(fnc_i) \prod_{j=1}^{n_s} [1 - F(snc^*)]
\]

in which the non-essential factorial expression enumerates the equally-likely time orders of generating \( n_f \) failed test items failed and \( n_s \) Type I test items. This likelihood expression can be re-written as

\[
\text{likelihood} = \frac{cdp2^m f}{cdp1 \cdot n_f \cdot cdp2} \cdot \prod_{i=1}^{n_f} (fnc_i)^{cdp2 - 1} \cdot \prod_{i=1}^{n_f} \exp \left( \frac{-fnc_i}{cdp1} \right) \cdot \prod_{j=1}^{n_s} \exp \left( \frac{-snc^*}{cdp1} \right)
\]
The corresponding loge (likelihood) expression is

\[
\log_e(\text{likelihood}) = n_f \cdot \log_e(\text{cdp}2) - n_f \cdot \text{cdp}2 \cdot \log_e(\text{cdp}1) + (\text{cdp}2 - 1) \cdot \sum_{i=1}^{n_f} \log_e(fnc_i) - \sum_{i=1}^{n_f} \left( \frac{\text{fnc}_i}{\text{cdp}1} \right)^{\text{cdp}2} - \sum_{k=1}^{n_k} \left( \frac{\text{anc}_k}{\text{cdp}1} \right)^{\text{cdp}2}
\]

The last two terms in this loge likelihood expression can combined under a summation that pertains to all experiment test program datum values. Accordingly, we subsequently use the index \( k \) for all test items, where \( n_k = n_f + n_s \) and \( k = 1 \) to \( n_k \), the number of individual tests in the fatigue life experiment test program. Correspondingly, we include both \( fnc_i \) and \( \text{anc}_k \) in \( \text{anc}_k \).

The conceptual distribution parameters \( \text{cdp}1 \) and \( \text{cdp}2 \) can be estimated by simultaneously solving the partial derivative expressions

\[
\frac{\partial \log_e(\text{liklihood})}{\partial \text{cdp}1} = 0 = -n_f \cdot \frac{\text{cdp}2}{\text{cdp}1} + \text{cdp}2 \cdot \sum_{k=1}^{n_k} \left( \frac{\text{anc}_k}{\text{cdp}1} \right)^{\text{cdp}2}
\]

and

\[
\frac{\partial \log_e(\text{liklihood})}{\partial \text{cdp}2} = 0 = n_f \cdot \frac{\text{cdp}2}{\text{cdp}1} - n_f \cdot \log_e(\text{cdp}1) + \sum_{i=1}^{n_f} \log_e(fnc_i) - \sum_{k=1}^{n_k} \left[ \left( \frac{\text{anc}_k}{\text{cdp}1} \right)^{\text{cdp}2} \cdot \log_e \left( \frac{\text{anc}_k}{\text{cdp}1} \right) \right]
\]

Or, if so desired, the second and third terms in the latter partial derivative expression can be combined under a single summation, viz.,

\[
-n_f \cdot \log_e(\text{cdp}1) + \sum_{i=1}^{n_f} \log_e(fnc_i) = \sum_{i=1}^{n_f} \log_e \left( \frac{\text{fnc}_i}{\text{cdp}1} \right)
\]

The two resulting non-linear partial derivative equations, whatever their algebraic form, can be solved numerically by using the iterative Newton-Raphson (N-R) methodology in which these two equations are simultaneously expanded in a Taylor's series and higher order terms are ignored, viz.,

\[
0 = \frac{\partial \log_e(\text{liklihood})}{\partial \text{cdp}1} + \frac{\partial^2 \log_e(\text{liklihood})}{\partial \text{cdp}1^2} \cdot \Delta(\text{cdp}1) + \frac{\partial^2 \log_e(\text{liklihood})}{\partial \text{cdp}1 \partial \text{cdp}2} \cdot \Delta(\text{cdp}2)
\]

and

\[
0 = \frac{\partial \log_e(\text{liklihood})}{\partial \text{cdp}2} + \frac{\partial^2 \log_e(\text{liklihood})}{\partial \text{cdp}2 \partial \text{cdp}1} \cdot \Delta(\text{cdp}1) + \frac{\partial^2 \log_e(\text{liklihood})}{\partial \text{cdp}2^2} \cdot \Delta(\text{cdp}2)
\]

which are subsequently viewed simply as two equations in two unknowns, \( \Delta(\text{cdp}1) \) and \( \Delta(\text{cdp}2) \). Given initial numerical estimates for the \( \text{cdp}1 \) and the \( \text{cdp}2 \), these two analytical equations become two numerical equations that are easily solved for \( \Delta(\text{cdp}1) \) and \( \Delta(\text{cdp}2) \). These two \( \Delta \)'s are first-order corrections to the initial estimates for the \( \text{cdp}1 \) and the \( \text{cdp}2 \). In turn, these improved (corrected) estimates of the actual values for the \( \text{cdp}1 \) and the \( \text{cdp}2 \) can be further improved (corrected) using the iterative N-R solution methodology, viz.,

\[
\text{cdp}1 \text{ (first iteration)} = \text{cdp}1 \text{ (initial estimate)} + \text{N-R } \Delta(\text{cdp}1) \text{ numerical correction}
\]

and

\[
\text{cdp}2 \text{ (first iteration)} = \text{cdp}2 \text{ (initial estimate)} + \text{N-R } \Delta(\text{cdp}2) \text{ numerical correction}
\]
In turn,
\[ \text{cdp1 (second iteration)} = \text{cdp1 (first iteration)} + \text{new N-R } \Delta(\text{cdp1}) \text{ numerical correction} \]
and
\[ \text{cdp2 (second iteration)} = \text{cdp2 (first iteration)} + \text{new N-R } \Delta(\text{cdp2}) \text{ numerical correction} \]
et cetera

The iterative N-R estimation algorithm should continue until successive iterations produce absolute values for both \( \Delta(\text{cdp}^1) \)'s numerical corrections less than about \( 10^{-12} \), at which time, the current numerical values for \text{cdp1} and \text{cdp2} are regarded as the respective maximum likelihood estimates.

Remark: Simulation-based multiplicative statistical bias-correction factors for ML \text{est(cdp1)} and ML \text{est(cdp2)} are developed in Supplemental Topic 8.E.

The primary advantage of the iterative N-R methodology in ML analysis is that numerical values for the second partial derivatives are directly available for subsequent use in computing the estimated asymptotic covariance matrix, viz., the negatives of the numerical values of the respective second partial derivatives form elements of a 2 by 2 symmetrical array whose inverse is the ML estimated asymptotic covariance matrix. In turn, given the respective ML estimates and this ML estimated asymptotic covariance matrix, the propagation of variability methodology can be employed to compute Method One and Method Two lower 100(scp)% (one-sided) asymptotic statistical confidence limits (losasc1's) that allegedly bound the actual value for \( \text{fnct(}\text{pf}^*\text{)} \) of specific interest. Although the respective Method One and Method Two \text{losascl}'s are asymptotically identical, Method Two \text{losasc1}'s are always smaller (safer) and more accurate.

Remark: When a \text{losasc1} pertains to scp = 0.95 it can either be re-interpreted as an asymptotic A-basis statistical tolerance limit if \( \text{pf} = 0.01 \) or as an asymptotic B-basis statistical tolerance limit if \( \text{pf} = 0.10 \). (See Supplemental Topic 8.A.)

Method One \text{losasc1}'s Given Weibull Life (Endurance) Data:

First, we write the inverse CDF expression, viz.,
\[ y_{\text{weibull}}(\text{pf}) = \log_e \left( -\log_e \left( 1 - F\left( \text{fnct}\left( \text{pf}\right) \right) \right) \right) = \text{cdp2} \cdot \left\{ \log_e \left( \text{fnct}\left( \text{pf}\right) \right) - \log_e \left( \text{cdp1} \right) \right\} \]
and, for the \( \text{pf}^* \) value of specific interest, we compute \( y_{\text{weibull}}(\text{pf}^*) \) using the expression
\[ y_{\text{weibull}}(\text{pf}^*) = \log_e \left( -\log_e \left( 1 - \text{pf}^* \right) \right) \]
(For example, when \( \text{pf}^* = 0.01 \), \( y_{\text{weibull}}(0.01) = -4.6001492 \).) Then, given values for \text{est(cdp1)} and \text{est(cdp2)}, we compute \( \text{est}\{\log_e[\text{fnct(}\text{pf}^*\text{)}]\} \) using the expression
\[ \text{est}\{\log_e[\text{fnct(}\text{pf}^*\text{)}]\} = \left\{ y_{\text{weibull}}(\text{pf}^* ) / [\text{est(cdp2)}] \right\} + \log_e [\text{est(cdp1)}] \]

Next we state the propagation of variability expression for \text{est} [\text{var} (\text{est}\{\log_e[\text{fnct(}\text{pf}^*\text{)}]\})], viz.,
\[ \text{est}\left[ \text{var} \left( \text{est}\{\log_e[\text{fnct(}\text{pf}^*\text{)}]\} \right) \right] = \left[ \frac{1}{\text{est(cdp1)}} \right]^2 \cdot \text{est}\left[ \text{var}(\text{est(cdp1)}) \right] + \left[ \frac{y_{\text{weibull}}(\text{pf}^*)}{-\text{est(cdp2)}} \right]^2 \cdot \text{est}\left[ \text{var}(\text{est(cdp2)}) \right] \\
+ 2 \cdot \left[ \frac{1}{\text{est(cdp1)}} \right] \cdot \left[ \frac{y_{\text{weibull}}(\text{pf}^*)}{-\text{est(cdp2)}} \right] \cdot \text{est}\left[ \text{covar}(\text{est(cdp1)}, \text{est(cdp2)}) \right] \]
and evaluate this expression using the elements of the estimated asymptotic covariance matrix to compute \( \text{est}(\text{stddev}(\text{est}(\log_e[fnc(pf^*)]))) \). In turn, the Method One losascl that allegedly bounds the actual value for \( fnc(pf^*) \) is computed using the expression

\[
\text{Method One } \text{losascl} = \exp\left(\text{est}\left(\log_e[fnc(pf^*)]\right) - y_{\text{normal}}(scp) \cdot \text{est}\left(\text{stddev}(\text{est}(\log_e[fnc(pf^*)]))\right)\right)
\]

Method Two losascl's Given Weibull Life (Endurance) Data:
Method Two is akin to inverse regression in that \( \log_e[fnc(pf)] \) is treated as a parameter in analysis. Again, we first write the inverse CDF expression, viz.,

\[
y_{\text{weibull}}(pf) = \log_e\left(-\log_e\left(1 - F\left[fnc(pf)\right]\right)\right) = cdpl2 \cdot \left(\log_e\left[fnc(pf)\right] - \log_e\left[\text{est}(cdpl)\right]\right)
\]

and, for the \( pf^* \) value of specific interest, we compute \( y_{\text{weibull}}(pf^*) \) using the expression

\[
y_{\text{weibull}}(pf^*) = \log_e\left(-\log_e\left(1 - pf^*\right)\right)
\]

Then, given values for \( \text{est}(cdpl) \) and \( \text{est}(cdpl2) \), we express \( \text{est}\{y_{\text{weibull}}(pf) \text{ given } \log_e[fnc(pf)]\} \) as

\[
\text{est}\{y_{\text{weibull}}(pf) \text{ given } \log_e[fnc(pf)]\} = \text{est}(cdpl2) \cdot \left(\log_e\left[fnc(pf)\right] - \log_e\left[\text{est}(cdpl)\right]\right)
\]

Next we state the propagation of variability expression for \( \text{est}(\text{var}(\text{est}\{y_{\text{weibull}}(pf) \text{ given } \log_e[fnc(pf)]\})) \), viz.,

\[
\text{est}\left(\text{var}\left(y_{\text{weibull}}(pf) \text{ given } \log_e[fnc(pf)]\right)\right) = -\left(\frac{\text{est}(cdpl2)}{\text{est}(cdpl)}\right)^2 \cdot \text{est}\left(\text{var}(\text{est}(cdpl))\right) + \left(\log_e\left[fnc(pf)\right] - \log_e\left[\text{est}(cdpl)\right]\right)^2 \cdot \text{est}\left(\text{var}(\text{est}(cdpl2))\right)
\]

and, given the respective elements of the estimated asymptotic covariance matrix, we evaluate this expression for each parametric value of \( \log_e[fnc(pf)] \) of specific interest. In turn, the Method Two losascl that allegedly bounds the actual value for \( fnc(pf^*) \) is computed as that value for \( \exp\{\log_e[fnc(pf)]\} \) such that

\[
y_{\text{weibull}}(pf^*) = \text{est}\{y_{\text{weibull}}(pf) \text{ given } \log_e[fnc(pf)]\} + y_{\text{normal}}(scp) \cdot \text{est}\left(\text{stddev}(\text{est}(y_{\text{weibull}}(pf) \text{ given } \log_e[fnc(pf)]))\right)
\]

Perspective: When Method Two losascl's pertaining to \( scp = 0.95 \) and \( pf = 0.01 \) and 0.10 are computed by running microcomputer program WEIBULL (illustrated below), the resulting re-interpreted ML-based asymptotic A-basis and B-basis statistical tolerance limits can be compared to corresponding ML-based exact (unbiased) A-basis and B-basis statistical tolerance limits computed by running microcomputer programs ABWSTL and BBWSTL presented in Supplemental Topic 8.5A. Then it is clear that even Method Two losascl's are not sufficiently accurate to be used in mechanical reliability analyses. Thus we employ simulation-based pragmatic methodologies in Supplemental Topics 8.B, 8.C, 8.D, and 8.E that apply not only to both Weibull and \( \log_e\)-normal fatigue life (endurance) datum values with either no censoring or with potential or actual Type I censoring, but also to conditional ML analyses for \( s_{\alpha}-fnc \) datum values with either no censoring or with potential or actual Type I censoring. (These simulation-based pragmatic ML analyses are without peer.)
We now present exemplar outputs for four Weibull microcomputer programs that compute both Method One and Method Two losascl's for comparative purposes. Microcomputer program \textit{WEIBULLN} pertains to a ML analysis for replicate fatigue life datum values with no censoring that are presumed to have been randomly selected from a conceptual two-parameter Weibull life distribution. Recall that the terminology \textit{no censoring} explicitly connotes an experiment test program in which all tests are conducted until failure occurs, regardless of the individual test durations required. Note that numerous life (endurance) and $s_a$-log$_e(fnc)$ tests with ostensibly no censored datum values do not conform to this terminology.

Microcomputer program \textit{WEIBULLP} pertains to a ML analysis for (replicate) fatigue life datum values with potential \textit{Type I} censoring that are presumed to have been randomly selected from a conceptual two-parameter Weibull distribution. The terminology \textit{potential Type I censoring} explicitly connotes that each test in the fatigue life experiment test program is conducted until either failure occurs or until the test is suspended because the pre-determined \textit{Type I} censoring value $fnc = snc^*$ is reached, but that no test was suspended during the given experiment test program. (Note that the value for $snc^*$ must be specified in microcomputer file \textit{WBLDTAPN}.) Although microcomputer programs \textit{WEIBULLN} and \textit{WEIBULLP} ostensibly generate the same ML estimates and lower 100($sep$)$\%$ (one-sided) asymptotic statistical confidence limits for a given set of datum values, their respective simulation-based pragmatic sampling distributions (Supplemental Topic 8.D) can differ markedly. Recall that the number of \textit{Type I} censored datum values is a random variable under continual replication of this fatigue life experiment test program.

Microcomputer program \textit{WEIBULLA} pertains to a ML analysis for (replicate) fatigue life datum values with actual \textit{Type I} censoring that are presumed to have been randomly selected from a conceptual two-parameter Weibull distribution. The terminology \textit{actual Type I censoring} explicitly connotes that each test the fatigue life experiment test program is conducted until failure occurs, or until the test is suspended because the pre-determined practical \textit{Type I} censoring value $fnc = snc^*$ is reached, and that one or more tests were indeed suspended at $snc^*$ during this experiment test program. The iterative N-R algorithm employed in microcomputer program \textit{WEIBULLA} almost always converges for datum values with only a small proportion of \textit{Type I} censoring. However, the two partial derivative expressions of the log$_e$-likelihood with respect to $cdp1$ and $cdp1$ can be reduced algebraically to a single derivative expression with respect to $cdp2$. Then the interactive N-R algorithm will always converge (unless all of the datum values are \textit{Type I} censored).

\textbf{Remark:} The precision of ML-based estimates is poor when a substantial proportion of the individual tests in a fatigue life (endurance) experiment test program require \textit{Type I} censoring. Accordingly, we recommend (i) switching to an up-and-down strength (resistance) test program as soon as it becomes clear that a substantial proportion of the individual life (endurance) tests will (may) require \textit{Type I} censoring, and (ii) then adopting the minimum variance strategy as soon as it can be employed.

Microcomputer program \textit{WBLARBST} pertains to data from a fatigue life experiment test program that (i) is terminated by an arbitrarily suspended test, but (ii) may include one or more prior \textit{Type I} suspended tests. It requires initial estimates of the actual values for $cdp1$ and $cdp2$ as input data. These values can be established graphically by plotting the failure and suspension life (endurance) datum values on Weibull probability paper using Mantel's log-rank algorithm (Section 8.5B). Note, however, because the concept of a continually replicated experiment test program is weakened when an arbitrarily suspended test occurs (for any reason), the statistical credibility of this ML analysis is diminished.
### Quantitative Analysis for the Outcomes of Life (Endurance) and Strength (Resistance)

#### Experiment Test Programs

**C> TYPE WBLDTANC**

- **6**  Number of Items Failed, Followed by the Respective \( fnc_i \)'s (Kilocycles)
- **277**
- **310**
- **374**
- **402**
- **456**
- **475** No Type I Censored Tests Permitted
- **95**  Statistical Confidence Probability of Specific Interest (Integer Value)
- **01**  Conceptual CDF Percentile of Specific Interest (Integer Value)

**C> TYPE WBLDTAPC**

- **6**  Number of Items Failed, Followed by the Respective \( fnc_i \)'s (Kilocycles)
- **277**
- **310**
- **374**
- **402**
- **456**
- **475** No Type I Censored Tests Occurred (See Below)
- **500** Pre-Determined \( snr^* \) Value (Kilocycles) for Potential Type I Censoring
- **95**  Statistical Confidence Probability of Specific Interest (Integer Value)
- **01**  Conceptual CDF Percentile of Specific Interest (Integer Value)

**C> TYPE WBLDTAAC**

- **5**  Number of Items Failed, Followed by the Respective \( fnc_i \)'s (Kilocycles)
- **277**
- **310**
- **374**
- **402**
- **456** One or More Actual Type I Censored Tests, Followed by the Identical \( snr^* \)'s
- **1**
- **500**
- **95**  Statistical Confidence Probability of Specific Interest (Integer Value)
- **01**  Conceptual CDF Percentile of Specific Interest (Integer Value)

**C> TYPE WBLDTAAS**

- **4**  Number of Items Failed, Followed by the Respective \( fnc_i \)'s (Kilocycles)
- **277**
- **310**
- **402**
- **456**
- **2** Number of Type I and Arbitrarily Suspended Tests, Followed by the Respective \( snr_j \)'s
- **500**
- **374**
- **450** Initial Estimate of the Actual Value for the \( cdpl \)
- **5.0** Initial Estimate of the Actual Value for the \( cdpl \)
- **95**  Statistical Confidence Probability of Specific Interest (Integer Value)
- **01**  Conceptual CDF Percentile of Specific Interest (Integer Value)
Given no censoring and presuming a conceptual two-parameter Weibull distribution fatigue life model in ML analysis with the parameterization

\[ F(fnc) = 1 - \exp \left( -\left( \frac{fnc}{cdp1} \right)^{cdp2} \right) \]

and with no statistical bias corrections

\[
\begin{align*}
\text{est}(cdp1) &= .4120458424D+03 & \text{est}(cdp2) &= .6324568781D+01 \\
\text{est}([\text{var}(cdp1)]) &= .7864400101D+03 & \text{est}([\text{var}(cdp2)]) &= .4429476691D+01 \\
\text{est}([\text{covar}(cdp1,cdp2)]) &= .1870862407D+02 & \text{est}(\text{conceptual correlation coefficient}) &= .3169807130D+00 \\
\end{align*}
\]

\[
\begin{array}{ccc}
\text{fnc} & \text{est}(y) & \text{est}(pf) \\
277.000 & -2.5115944 & .0779343 \\
310.000 & -1.7997339 & .1523970 \\
374.000 & -6.127167 & .4183444 \\
402.000 & -.1561063 & .5749159 \\
456.000 & .6410469 & .8502020 \\
475.000 & .8992284 & .9143688 \\
\end{array}
\]

Lower 95% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for \( \text{fnc}(01) \)

\[ 127.405 - \text{Computed Using Method One} \]
\[ 75.413 - \text{Computed Using Method Two} \]

Given actual Type I censoring and presuming a conceptual two-parameter Weibull distribution fatigue life model in ML analysis with the parameterization

\[ F(fnc) = 1 - \exp \left( -\left( \frac{fnc}{cdp1} \right)^{cdp2} \right) \]

and with no statistical bias corrections

\[
\begin{align*}
\text{est}(cdp1) &= .4289595976D+03 & \text{est}(cdp2) &= .4731943406D+01 \\
\text{est}([\text{var}(cdp1)]) &= .1666932095D+04 & \text{est}([\text{var}(cdp2)]) &= .3013638227D+01 \\
\text{est}([\text{covar}(cdp1,cdp2)]) &= .8393830857D+01 & \text{est}(\text{conceptual correlation coefficient}) &= .1184283562D+00 \\
\end{align*}
\]

\[
\begin{array}{ccc}
\text{fnc} & \text{est}(y) & \text{est}(pf) \\
277.000 & -2.0694929 & .1186053 \\
310.000 & -.5368900 & .1934980 \\
374.000 & -.6487823 & .4070717 \\
402.000 & -.3071535 & .5207523 \\
456.000 & .2892640 & .7369587 \\
\end{array}
\]

Lower 95% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for \( \text{fnc}(01) \)

\[ 86.868 - \text{Computed Using Method One} \]
\[ 34.584 - \text{Computed Using Method Two} \]
Discussion: The Weibull distribution CDF can be expressed using eight different parameterizations. We present the four most common of these parameterizations below:

1. \[ F(fnc) = 1 - \exp\left(-\frac{fnc}{cdp1}\right)^{cdp2} \]
2. \[ F(fnc) = 1 - \exp\left(-cdp1 \cdot \left(\frac{fnc}{cdp1}\right)^{cdp2}\right) \]
3. \[ F(fnc) = 1 - \exp\left(-\frac{fnc}{cdp1}\right)^{cdp2} \]
4. \[ F(fnc) = 1 - \exp\left(-\frac{1}{cdp1} \cdot fnc\right)^{cdp2} \]

The corresponding four parameterizations for the conceptual (two-parameter) loge smallest-extreme-value distribution variate are employed in ML analyses given by microcomputer programs LSEV1AAC, LSEV2AAC, LSEV3AAC, and LSEV4AAC. The respective outputs for these four microcomputer programs demonstrate that (i) ML estimates are the same regardless of the parameterization selected for the CDF, and (ii) the associated classical lower 100(scp)% (one-sided) asymptotic statistical confidence limits that allegedly bound the actual value for any fnc(pf) of specific interest are the same when these intervals are computed using propagation of variance methodology (Supplemental Topic 7.D).
C> COPY WBLDTAAC DATA
     1 files(s) copied
C> LSEV1AAC

Given actual Type I censoring and presuming a conceptual two-parameter Weibull distribution fatigue life model, but employing the conceptual (two-parameter) smallest-extreme-value distribution in ML analysis with the parameterization

\[ y = csp \cdot [\log_e(fnc) - clp] \]

and with no statistical bias corrections

\[
\begin{align*}
est(clp) &= .6061362736D+01 \\
est(csp) &= .4731943406D+01 \\
est(\text{var}(clp)) &= .9059101593D-02 \\
est(\text{var}(csp)) &= .3013638227D+01 \\
est(\text{covar}(est(clp),est(csp))) &= .1956788216D-01 \\
est(\text{conceptual correlation coefficient}) &= .1184283562D+00
\end{align*}
\]

\[
\begin{array}{ccc}
\text{fnc} & \text{est}(y) & \text{est}(pf) \\
277.000 & -2.0694929 & .1186053 \\
310.000 & -1.5368900 & .1934980 \\
374.000 & -.6487823 & .4070717 \\
402.000 & -.3071535 & .5207523 \\
456.000 & .2892640 & .7369587 \\
\text{snc} & \text{est}(y) & \text{est}(pf) \\
500.000 & .7251484 & .8731865
\end{array}
\]

Lower 95\% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for fnc(01)

86.868 - Computed Using Method One
34.584 - Computed Using Method Two

C> COPY WBLDTAAC DATA
     1 files(s) copied
C> LSEV2AAC

Given actual Type I censoring and presuming a conceptual two-parameter Weibull distribution fatigue life model, but employing the conceptual (two-parameter) smallest extreme-value distribution in ML analysis with the parameterization

\[ y = clp + csp \cdot \log_e(fnc) \]

and with no statistical bias corrections

\[
\begin{align*}
est(clp) &= -.2868202543D+02 \\
est(csp) &= .4731943406D+01 \\
est(\text{var}(clp)) &= .3013638227D+01 \\
est(\text{var}(csp)) &= .1120467627D+03 \\
est(\text{covar}(est(clp),est(csp))) &= -.1835934856D+02 \\
est(\text{conceptual correlation coefficient}) &= -.9991071169D+00
\end{align*}
\]

\[
\begin{array}{ccc}
\text{fnc} & \text{est}(y) & \text{est}(pf) \\
277.000 & 2.0694929 & .1186053 \\
310.000 & -1.5368900 & .1934980 \\
374.000 & -.6487823 & .4070717 \\
402.000 & -.3071535 & .5207523 \\
456.000 & .2892640 & .7369587 \\
\text{snc} & \text{est}(y) & \text{est}(pf) \\
500.000 & .7251484 & .8731865
\end{array}
\]

Lower 95\% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for fnc(01)

86.868 - Computed Using Method One
34.584 - Computed Using Method Two
Given actual Type I censoring and presuming a conceptual two-parameter Weibull distribution fatigue life model, but employing the conceptual (two-parameter) smallest-extreme-value distribution in ML analysis with the parameterization

\[ y = \left( \log(e_{nc}) - cl/p \right) \]

and with no statistical bias corrections

\[
\begin{align*}
\text{est}(cl/p) & = .6061362736D+01 & \text{est}(csp) & = .2113296619D+00 \\
\text{est}\{\text{var}(cl/p)\} & = .9059101593D-02 & \text{est}\{\text{var}(csp)\} & = .6010809287D-02 \\
\text{est}\{\text{covar}(cl/p, csp)\} & = -.8739060394D-03 & \text{est}(\text{conceptual correlation coefficient}) & = -.1184283562D+00
\end{align*}
\]


\[
\begin{array}{ccc}
\text{fnc} & \text{est}(y) & \text{est}(\text{pf}) \\
277.000 & -2.0694929 & .1186053 \\
310.000 & -1.5368900 & .1934980 \\
374.000 & -.6487823 & .4070717 \\
402.000 & -.3071535 & .5207523 \\
456.000 & .2892640 & .7369587 \\
\text{snc} & \text{est}(y) & \text{est}(\text{pf}) \\
500.000 & .7251484 & .8731865
\end{array}
\]

Lower 95% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for fnc(01)

86.868 - Computed Using Method One
34.584 - Computed Using Method Two
Supplemental Topic 8.B presents the analytical details of the analogous ML analysis that pertains to a conceptual two-parameter log_{e}-normal distribution. Given the same set of life (endurance) datum values, the respective \{\text{est}[\text{fncc}(p^{*})]\}'s for small values of \(p^{*}\) pertaining to the analogous conceptual two-parameter log_{e}-normal distribution are larger than for the conceptual two-parameter Weibull distribution. Thus quantitative life (endurance) estimates based on the conceptual two-parameter Weibull distribution are safer from a mechanical reliability perspective than corresponding quantitative life (endurance) estimates based on the conceptual two-parameter log_{e}-normal distribution — which is the only rational reason for preferring the conceptual two-parameter Weibull distribution over the conceptual two-parameter log_{e}-normal distribution. Dumonceaux and Antle (1973) demonstrated that the conceptual two-parameter Weibull and log_{e}-normal distributions are statistically interchangeable for experiment test programs with sizes that are practical in mechanical reliability applications. They used the ratio of the respective estimated maximized likelihoods as the test statistic in an attempt to discriminate between these two distributions statistically. Their simulation-based study showed that, even given an acceptable probability of committing a Type I error as large as 0.10, a sample size of almost fifty is required for their discrimination test to have a statistical power of 0.90, viz., to have a probability equal to 0.90 of correctly rejecting the presumed conceptual statistical distribution (the null hypothesis) in favor of the correct conceptual statistical distribution (the alternative hypothesis).

**Exercise Set 3**

These exercises are intended to use propagation of variability expressions (Supplemental Topic 7.A) to verify numerically the analytical relationships between (i) the respective \text{est}(cdp)'s for the conceptual two-parameter Weibull distribution and the corresponding conceptual (two-parameter) smallest-extreme-value distributions with a logarithmic metric, and (ii) the elements of their respective estimated asymptotic covariance matrices. Use at least eight digits in your calculations to avoid so-called round-off errors.

1. Given the numerical values for the respective \text{est}(cdp)'s in the example microcomputer program \textit{WEIBULLA} output, verify the corresponding numerical values for the respective \text{est}(cdp)'s given in the example output for microcomputer program (a) \textit{LSEV1AAC}, (b) \textit{LSEV2AAC}, (c) \textit{LSEV3AAC}, and (d) \textit{LSEV4AAC}. Then, given the numerical values for the respective elements of the estimated asymptotic covariance matrix in the example microcomputer program \textit{WEIBULLA} output, verify the corresponding numerical values for the respective elements of the estimated asymptotic covariance matrix given in the example microcomputer program (a) \textit{LSEV1AAC}, (b) \textit{LSEV2AAC}, (c) \textit{LSEV3AAC}, and (d) \textit{LSEV4AAC} output.

2. Given the numerical values for the respective \text{est}(cdp)'s in the example microcomputer program \textit{LSEV1AAC} output, verify the corresponding numerical values for the \text{est}(cdp)'s given in the example output for microcomputer program (a) \textit{WEIBULLA}, (b) \textit{LSEV2AAC}, (c) \textit{LSEV3AAC}, and (d) \textit{LSEV4AAC}. Then, given the numerical values for the respective elements of the estimated asymptotic covariance matrix in the example microcomputer program \textit{LSEV1AAC} output, verify the corresponding numerical values for the respective elements of the estimated asymptotic covariance matrix given in the example outputs for microcomputer program (a) \textit{WEIBULLA}, (b) \textit{LSEV2AAC}, (c) \textit{LSEV3AAC}, and (d) \textit{LSEV4AAC}.

3. Given the numerical values for the respective \text{est}(cdp)'s in the example microcomputer program \textit{LSEV2AAC} output, verify the corresponding numerical values for the \text{est}(cdp)'s given in the example output for microcomputer program (a) \textit{WEIBULLA}, (b) \textit{LSEV1AAC}, (c) \textit{LSEV3AAC}, and (d) \textit{LSEV4AAC}. Then, given the numerical values for the respective elements of the estimated asymptotic covariance matrix in the example microcomputer program \textit{LSEV2AAC} output, verify the
corresponding numerical values for the respective elements of the estimated asymptotic covariance matrix given in the example output for microcomputer program (a) WEIBULLA, (b) LSEV1AAC, (c) LSEV3AAC, and (d) LSEV4AAC.

4. Given the numerical values for the respective est(cdp)'s in the example microcomputer program LSEV3AAC output, verify the corresponding numerical values for the est(cdp)'s given in the example output for microcomputer program (a) WEIBULLA, (b) LSEV1AAC, (c) LSEV2AAC, and (d) LSEV4AAC. Then, given the numerical values for the respective elements of the estimated asymptotic covariance matrix in the example microcomputer program LSEV3AAC output, verify the corresponding numerical values for the respective elements of the estimated asymptotic covariance matrix given in the example output for microcomputer program (a) WEIBULLA, (b) LSEV1AAC, (c) LSEV2AAC, and (d) LSEV4AAC.

5. Given the numerical values for the respective est(cdp)'s in the example microcomputer program LSEV4AAC output, verify the corresponding numerical values for the est(cdp)'s given in the example output for microcomputer program (a) WEIBULLA, (b) LSEV1AAC, (c) LSEV2AAC, and (d) LSEV3AAC. Then, given the numerical values for the respective elements of the estimated asymptotic covariance matrix in the example microcomputer program LSEV4AAC output, verify the corresponding numerical values for the respective elements of the estimated asymptotic covariance matrix given in the example output for microcomputer program (a) WEIBULLA, (b) LSEV1AAC, (c) LSEV2AAC, and (d) LSEV3AAC.

Example Two (Extended): ML Analysis for the Outcome of a Life (Endurance) Experiment Test Program with Competing Modes of Failure. Mechanical devices (and sometimes even their components) exhibit competing modes of failure. Accordingly, if the failed device is removed from service due to a specific mode of failure, a suspended test datum value is created for each of its competing modes of failure. (Note that the life (endurance) metrics for all competing modes of failure must be explicitly inter-related.) In turn, if all competing modes of failure are presumed to be mutually independent, the device likelihood expression is simply the product of the respective individual mode of failure likelihood expressions. Nevertheless, ML analysis for competing models of failure has the statistical problems that (i) it may be difficult (or impossible) to test the null hypothesis of mutual independence, and (ii) the concept of a continually replicated experiment test program is dubious when the respective suspension lives (endurances) are not statistically planned. Despite these problems, a competing modes of failure analysis is generally more credible than an analysis that arbitrarily lumps all competing modes of failure into a single failure model.

Consider a shaft with an antifriction bearing at each end. The likelihood expression for bearing failure, presuming the respective bearing failures are independent, is the product of the respective individual likelihood expressions, viz.,

\[
\text{likelihood} = \left[\text{likelihood(right - hand bearing failures)}\right] \cdot \left[\text{likelihood(right - hand bearing suspensions)}\right] 
\]

in which, in addition to suspension durations pertaining to both bearings, the right-hand bearing suspension durations also include the left-hand bearing failure datum values and the left-hand bearing suspension durations also include the right-hand-bearing failure datum values. Note that this likelihood expression can be expanded to include the shaft by including its fatigue failure datum values (perhaps pertaining to more than one location), or by including competing bending fatigue, pitting, and scoring modes of failure datum values for one or more gears mounted on the shaft, etc. cetera. Note also that the greater the number of competing modes of failure in ML analysis, the more dubious the presumption
that all of these competing modes of failure are indeed mutually independent.

Nevertheless, given the presumption that all competing modes of failure are indeed mutually independent, in ML analysis we set the partial derivatives of the (aggregate) loge(likelihood) expression with respect to the respective parameters of the conceptual statistical distribution presumed for each mode of failure. Then ML analysis for mutually independent competing modes of failure is, for practical purposes, merely the amalgamation of several ML analyses similar to the ML analysis in microcomputer program **WBLARBST**, where the respective life (endurance) metrics for each difference mode of failure can be modeled using any conceptual statistical distribution of specific interest. Accordingly, ML analysis for mutually independent competing modes of failure merely requires having a library of alternative ML analyses available for each mode of failure of potential interest. The resulting estimated CDF would be given by the expression

$$\text{est}[F(fnc)] = 1 - \left(1 - \text{est}[F(fnc)_{\text{mode 1}}]\right) \cdot \left(1 - \text{est}[F(fnc)_{\text{mode 2}}]\right) \cdots \left(1 - \text{est}[F(fnc)_{\text{mode n}}]\right)$$

Unfortunately classical lower 100($scp$)% (one-sided) asymptotic statistical confidence limits analogous to those given in microcomputer program **WBLARBST** cannot be computed using propagation of variability. However, given any value for $pf$ of specific interest, a lower 100($scp$)% (one-sided) asymptotic statistical confidence limit that allegedly bounds the actual value for $fnc(pf)$ can be approximated numerically.

**Remark One:** Anti-friction bearing failures are typically a competing modes of failure problem. Pitting fatigue failure usually initiates at the surface of one of the several rolling elements (balls or rollers), but it can initiate at the corresponding contact surface of either the inner or the outer race. Moreover, each of these rolling elements also individually experiences competing modes of failure, viz., their pitting fatigue life typically depends on the specific type of non-metallic inclusion at which the fatigue crack initiates.

**Remark Two:** Scanning electron microscope (SEM) examination of failure surfaces for low-cycle (strain-controlled) fatigue tests conducted at elevated temperatures often indicates the existence of two distinct types of fatigue crack initiation processes (thereby indicating competing modes of fatigue crack initiation).

**Example Three: Conditional ML Analysis for the Outcome of a $s_a$-$fnc$ Experiment Test Program, Presuming a Homoscedastic Fatigue Strength Distribution.** Although a number of different $s_a$-$\log_e[fnc(pf)]$ models have been proposed to describe the outcomes of load-controlled fatigue tests, only two are both simple and physically and statistically credible: a parabolic $s_a$-$\log_e[fnc(pf)]$ model for $s_a$-$fnc$ data without an endurance limit, and a two-segment straight-line $s_a$-$\log_e[fnc(pf)]$ model for $s_a$-$fnc$ data with an endurance limit. The parabolic $s_a$-$\log_e[fnc(pf)]$ model that is quadratic in terms of $\log_e(fnc)$ is physically more credible for $s_a$-$fnc$ data without an endurance limit, both at long life and at shorter lives, than the parabolic $s_a$-$\log_e[fnc(pf)]$ model that is quadratic in $s_a$: (i) its slope approaches zero at very long lives and its curvature increases as the slope approaches zero, and (ii) its slope becomes very steep at short lives and its curvature decreases as the slope increases. The former parabolic $s_a$-$\log_e[fnc(pf)]$ model, however, requires an interchange (reversal) of the independent and dependent variables and the associated ML analysis is conditional on the observed experiment test program $fnc_i$ and $snc_i$ datum values. Given no censoring, this conditional ML analysis is identical to a linear regression (least-squares) analysis (Chapter Seven) when (i) the independent variable ($s_a$) and dependent variable [$\log_e(fnc)$] are interchanged, and (ii) when the presumed $s_a$-$\log_e[fnc(pf)]$ model employs a conceptual homoscedastic (two-parameter) normal fatigue strength distribution and Version
LS statistical bias corrections (Supplemental Topic 8.B) are included in analysis. Then Student's central t and Student's non-central t can be used to compute analogous lower 100(scp)% (one-sided) statistical confidence and tolerance limits. On the other hand, given either potential or actual Type I censoring, analogous Version LS statistical bias corrections are unknown. Accordingly, we can only approximate analogous lower 100(scp)% (one-sided) statistical confidence and tolerance limits by either using extrapolated Version LS statistical bias corrections (Supplemental Topic 8.C), or employing our pragmatic bias-correction methodology (Supplemental Topic 8.E).

Microcomputer programs SAFNCM1N, SAFNCM2N, SAFNCM3N, and SAFNCM4N respectively pertain to a straight-line $s_\alpha$-$\log_{e}[fnc(pf)]$ model with linear and logarithmic $s_\alpha$ metrics and to a parabolic $s_\alpha$-$\log_{e}[fnc(pf)]$ model with linear and logarithmic $s_\alpha$ metrics. These microcomputer programs first compute the conditional ML estimate of the actual value for $s_{fs}(50)$ given the $fnc^*$ value of specific interest and then employ (i) Student's central t to compute an analogous lower 100(scp)% (one-sided) statistical confidence limit that allegedly bounds the actual value for $s_{fs}(50)$ given this $fnc^*$ value and (ii) Student's non-central t to compute an analogous lower 100(scp)% (one-sided) statistical tolerance limit that allegedly bounds the actual value for $s_{fs}(pf)$ given this $fnc^*$ value. Although we recommend running microcomputer program SAFNCM3N given load-controlled $s_\alpha$fnc data with no censoring for a material that does not exhibit an endurance limit, the likelihood ratio test can be used to decide when it is statistically appropriate to switch from a linear to a logarithmic metric for $s_\alpha$ and when it is statistically appropriate to switch from a parabolic $s_\alpha$-$\log_{e}[fnc(pf)]$ model to a straight-line $s_\alpha$-$\log_{e}[fnc(pf)]$ model.

C> COPY SAFNCDTA DATA
1 file(s) copied

C> TYPE DATA

6
1
50.0 13000
2
40.0 55440
40.0 66800
1
35.0 110600
3
30.0 263000
30.0 478000
30.0 3137000
1
27.5 1205000
1
25.0 9497000
95 $scp$ of Specific Interest in Per Cent (Integer Value)
10000000 $fnc$ Value of Specific Interest
10 Associated Strength Percentile of Specific Interest in Per Cent (Integer Value)
Given no censoring and presuming a $s_\alpha \log_e[fnc(pf)]$ model in a conditional ML analysis with the homoscedastic conceptual (two-parameter) normal fatigue strength distribution augmented parameterization

$$y = \{s_\alpha - \text{clip} \cdot \text{clip} \cdot [\log_e(fnc)] - \text{clip} \cdot [\log_e(fnc)]^2\} / \text{clip}$$

and with Version LS statistical bias-correction factors

<table>
<thead>
<tr>
<th>fnc</th>
<th>$s_\alpha$</th>
<th>est$[s_\theta(50)]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>13000</td>
<td>50.0</td>
<td>49.9</td>
</tr>
<tr>
<td>55400</td>
<td>40.0</td>
<td>39.8</td>
</tr>
<tr>
<td>66800</td>
<td>40.0</td>
<td>38.7</td>
</tr>
<tr>
<td>110600</td>
<td>35.0</td>
<td>35.9</td>
</tr>
<tr>
<td>263000</td>
<td>30.0</td>
<td>32.1</td>
</tr>
<tr>
<td>478000</td>
<td>30.0</td>
<td>30.0</td>
</tr>
<tr>
<td>3137000</td>
<td>30.0</td>
<td>26.6</td>
</tr>
<tr>
<td>1205000</td>
<td>27.5</td>
<td>27.7</td>
</tr>
<tr>
<td>9479000</td>
<td>25.0</td>
<td>26.5</td>
</tr>
</tbody>
</table>

The minimum value for est$[s_\theta(50)]$ is equal to 26.48 (units). It is employed for all fnc greater than 4625328.8 cycles.

**Estimate of the Actual Value for $s_\theta(50)$ at 10000000. Cycles**

26.5

**Lower 95% (One-Sided) Statistical Confidence Limit**
that Allegedly Bound the Actual Value for $s_\theta(50)$ at 10000000. Cycles

**Computed Using Student’s Central t**
24.3

**Lower 95% (One-Sided) Statistical Tolerance Limit**
that Allegedly Bound the Actual Value for $s_\theta(10)$ at 10000000. Cycles

**Computed Using Student’s Non-Central t**
20.6

Discussion: The two-segment straight-line $s_\alpha \log_e[fnc(pf)]$ model is based on the assertion that the fatigue strength and endurance limit distributions are identical at its knee. However, there is a inconsistency between the physical presumptions underlying the ML analyses pertaining to the respective segments of this model, viz., it is presumed that a run-out will never fail no matter how long the test is continued (because its imposed alternating stress amplitude is smaller than its threshold endurance limit value), whereas if Type I censored tests had been continued indefinitely, failure would have always occurred eventually. Thus a run-out cannot legitimately be treated as a Type I censored test in ML analysis. Accordingly, our conditional ML analysis for the downward-sloping segment of the two-segment straight-line $s_\alpha \log_e[fnc(pf)]$ model must be restricted to only fatigue failure datum values (Supplemental Topic 8.C.). Nevertheless, it is advantageous to employ its estimate of the actual value for the standard deviation of the presumed homoscedastic conceptual (two-parameter) normal fatigue strength limit distribution in subsequently estimating the actual value for the median endurance limit and the associated actual location of the knee in the two-segment straight-line $s_\alpha \log_e[fnc(pf)]$ model.

Although the outcome of a typical $s_\alpha$fnc experiment test programs for a material with an endurance limit is stated in terms of the estimate of the actual value for the median endurance limit $s_\theta(50)$, typically less than 20% of its fatigue test specimens are allocated to alternating stress amplitudes in the vicinity of est$[s_\theta(50)]$. Accordingly, typically more than 80% of its fatigue test specimens have negligible statistical weights relative to estimating $s_\theta(50)$. (See Figure 8D.1, Supplemental Topic 8.D.) Thus an up-and-down experiment test program with the same number of test specimens provides a much more statistically precise (and physically credible) estimate of $s_\theta(50)$.
Example Four: ML Analysis for the Outcome of a Strength (Resistance) Experiment Test Program Conducted Using the Up-and-Down Strategy. Although the up-and-down test method strategy is strictly valid only for threshold phenomena, this strategy nevertheless can be used in a very effective manner to estimate the actual value for the median of the presumed conceptual (two-parameter) strength (resistance) distribution at very long durations for those modes of failure whose presumed stimulus-duration models asymptotically reach a limiting (threshold) value, e.g., fatigue, stress-rupture, and stress-corrosion cracking — provided that the standard deviation (or the $csp$) of the given conceptual (two-parameter) distribution is alleged to be known. This sequential strategy is particularly effective when only a few (nominally identical) specimens are available for testing because it efficiently allocates these specimens to alternating stress amplitudes (stimulus levels) in the vicinity of the actual value for the median of the presumed conceptual “one-parameter” strength (resistance) distribution. Little (1981) presents tabulated quantities for use in estimating the actual value for the median of the presumed conceptual strength (resistance) distribution given the outcome of a conventional small-sample up-and-down experiment test program with a fixed increment between successive alternating stress amplitudes. These tabulated quantities pertain to four alternative conceptual “one-parameter” fatigue strength distributions: normal (symmetrical with short tails), logistic (symmetrical with long tails), smallest-extreme-value (non-symmetrical, skewed toward low values), and the largest-extreme value (non-symmetrical, skewed toward high values), and include enumeration-based (exact) statistical bias corrections. But, in general, statistical bias corrections are not employed in ML analyses for strength (resistance) data.

Remark: We explain why up-and-down test programs with fixed increments do not generate minimum variance estimates in Supplemental Topic 8.D.

The example data set in microcomputer file $UADDATA$ pertains to a fatigue experiment test program that employed the conventional small sample up-and-down test method strategy for the first six test specimens, followed by allocation of the next four test specimens to their respective optimal alternating stress amplitudes successively computed by running microcomputer program $N/A$ (Normal Distribution, “One-Parameter” - Version A). Because microcomputer program $N/A$ has the actual values for the respective alternating stress amplitudes as input (rather than a beginning value and a fixed increment), it is conveniently used at any time during the experiment test program to allocate the next fatigue specimen to its statistically most effective alternating stress amplitude regardless of the strategy (strategies) employed for testing prior fatigue specimens.

<table>
<thead>
<tr>
<th>C&gt; TYPE</th>
<th>UADDATA</th>
<th>Fatigue Life, cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Number of Alternating Stress Amplitudes</td>
<td></td>
</tr>
<tr>
<td>3000</td>
<td>1,1</td>
<td>Alternating Stress Amplitude (space)</td>
</tr>
<tr>
<td>2750</td>
<td>1,1</td>
<td>Number of Items Tested (space)</td>
</tr>
<tr>
<td>2500</td>
<td>1,0</td>
<td>Number of Items Failed</td>
</tr>
<tr>
<td>2750</td>
<td>1,1</td>
<td></td>
</tr>
<tr>
<td>2500</td>
<td>1,0</td>
<td></td>
</tr>
<tr>
<td>2750</td>
<td>1,0</td>
<td></td>
</tr>
<tr>
<td>2710</td>
<td>1,1</td>
<td></td>
</tr>
<tr>
<td>2655</td>
<td>1,0</td>
<td></td>
</tr>
<tr>
<td>2700</td>
<td>1,0</td>
<td></td>
</tr>
<tr>
<td>2740</td>
<td>1,0</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>Presumed Standard Deviation of the Presumed Conceptual One-Parameter Strength (Resistance) Distribution</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>Statistical Confidence Probability of Specific Interest in Per Cent (Integer Value)</td>
<td></td>
</tr>
</tbody>
</table>

Note: DNF denotes Did Not Fail under Type I censoring.
COPY UADDATA DATA
  1 file(s) copied

C> N1A

Presuming a conceptual (two-parameter) normal distribution strength (resistance) model in ML analysis with its known (guestimated) standard deviation equal to 250 (and with no statistical bias corrections)

\[ \text{est}(s(50)) = 2778.0 \]

Lower 95\% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for \( s(50) \)

2604.7 - Computed Using Method Two
2602.9 - Computed Using the Likelihood Ratio Method

Microcomputer program N1A has the presumed known (guestimated) value for the standard deviation of the presumed conceptual one-parameter normal strength (resistance) distribution as input data. However, this presumed known (guestimated) value can be changed later (i) to obtain a revised estimate of the actual value for the median (mean) of the presumed conceptual (two-parameter) normal strength (resistance) distribution if a more accurate known (guestimated) value for its standard deviation becomes available, or (ii) to study the sensitivity of the ML estimate for the actual value for the median (mean) of the presumed conceptual (two-parameter) normal strength (resistance) distribution model to alternative presumed known (guestimated) values for its standard deviation.

Microcomputer programs L1A (Logistic Distribution, “One-Parameter” - Version A), SEV1A (Smallest-Extreme-Value Distribution, “One-Parameter” - Version A), and LEV1A (Largest-Extreme-Value Distribution, “One-Parameter” - Version A) demonstrate that the ML estimate of the actual value for the metric pertaining to the median of the presumed conceptual one-parameter strength (resistance) distribution is relatively insensitive to the presumed CDF.

COPY UADDATA DATA
  1 file(s) copied

C> L1A

Presuming a conceptual (two-parameter) logistic strength distribution (resistance) model in ML analysis with its known (guestimated) standard deviation equal to 250 (and with no statistical bias corrections)

\[ \text{est}(s(50)) = 2769.8 \]

Lower 95\% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for \( s(50) \)

2610.6 - Computed Using Method Two
2612.7 - Computed Using the Likelihood Ratio Method

COPY UADDATA DATA
  1 file(s) copied

C> SEV1A

Presuming a conceptual (two-parameter) smallest-extreme-value distribution strength (resistance) model in ML analysis with its known (guestimated) standard deviation equal to 250 (and with no statistical bias corrections)

\[ \text{est}(s(50)) = 2768.1 \]

Lower 95\% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for \( s(50) \)

2603.6 - Computed Using Method Two
2615.9 - Computed Using the Likelihood Ratio Method
<COPY UAADDATA DATA
1 file(s) copied

C> LEV1A

Presuming a conceptual (two-parameter) largest-extreme-value distribution strength (resistance) model in ML analysis with its known (guestimated) standard deviation equal to 250 (and with no statistical bias corrections)

\[ \text{est}(s(50)) = 2773.2 \]

Lower 95% (One-Sided) Asymptotic Statistical Confidence Limits that Allegedly Bound the Actual Value for \( s(50) \)

2630.8 - Computed Using Method Two
2613.3 - Computed Using the Likelihood Ratio Method

**Discussion:** The statistical literature does not exploit the versatility of ML analysis for strength (resistance) experiment test programs. The individual test outcomes do not have to be limited to realization values that are equal to zero or one. Rather, we can establish a damage index that lies in the interval from zero (no damage whatsoever) to one (complete failure). Then each test item (specimen) pertaining to each suspended test can be examined for signs of physical damage. For example, each of the automotive composite fatigue specimens pertaining to a suspended test in microcomputer file UAADDATA was visually examined relative to the presence of small fatigue cracks, and/or noticeable local delamination. The damage index for two of these specimens was subjectively assessed as 0.50 as indicated in microcomputer file MUADDATA. Accordingly, based on the damage index assessment for the re-interpreted experiment test program outcomes, microcomputer program NIA computes a revised estimate of the actual value for the median (mean) of the presumed conceptual one-parameter normal fatigue strength (fatigue limit) distribution at \( 10^7 \) stress cycles.

Little and Kosikowski (unpublished) conducted a modified up-and-down experiment test program to establish a non-traditional measure for the fatigue notch sensitivity of an automotive composite material. The size of a small central hole was decreased or increased depending on whether the fatigue crack that led to failure did or did not originate at the central hole. Predictably there were several failures where the crack origin could not be stated with certainty. Accordingly, although the test strategy merely required replicating the test with the same size hole, a subjective probability index was required in ML analysis to estimate the diameter of the hole for which fatigue failure is equally likely or not to originate at the hole.

C> TYPE MUADDATA

<table>
<thead>
<tr>
<th>10</th>
<th>Number of Alternating Stress Amplitudes</th>
<th>Fatigue Life, cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>3000 1 1</td>
<td>Alternating Stress Amplitude (space)</td>
<td>3,133,000</td>
</tr>
<tr>
<td>2750 1 1</td>
<td>Number of Item Tested (space)</td>
<td>9,932,000</td>
</tr>
<tr>
<td>2500 1 0</td>
<td>Number of Items Failed</td>
<td>( 10^7 ) DNF</td>
</tr>
<tr>
<td>2750 1 1</td>
<td></td>
<td>4,900,000</td>
</tr>
<tr>
<td>2500 1 0</td>
<td></td>
<td>( 10^7 ) DNF</td>
</tr>
<tr>
<td>2750 1 0.5</td>
<td><em>(Subjective Damage Index)</em></td>
<td>( 10^7 ) DNF</td>
</tr>
<tr>
<td>2710 1 1</td>
<td></td>
<td>4,548,000</td>
</tr>
<tr>
<td>2655 1 0</td>
<td></td>
<td>( 10^7 ) DNF</td>
</tr>
<tr>
<td>2700 1 0</td>
<td></td>
<td>( 10^7 ) DNF</td>
</tr>
<tr>
<td>2740 1 0.5</td>
<td><em>(Subjective Damage Index)</em></td>
<td>( 10^7 ) DNF</td>
</tr>
<tr>
<td>250</td>
<td>Presumed Standard Deviation of the Presumed Conceptual One-Parameter Strength (Resistance) Distribution</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>Statistical Confidence Probability of Specific Interest, Stated in Percent (Integer Value)</td>
<td></td>
</tr>
</tbody>
</table>

Note: DNF Denotes a Run-Out (Did Not Fail) Under Type I Censoring
Finally, Little (unpublished) conducted several up-and-down long-life fatigue experiment test programs with an automotive composite specimen attached to a steel grip at each of its ends with nominally identical bolted lap joints. Accordingly, fatigue failure was equally likely to occur at either the top or the bottom lap joint. The conceptual strength distribution for this example thus pertains the smallest of two independent realizations randomly selected from the conceptual strength distribution presumed for each independent realization, viz.,

\[ F_{\text{smallest}}(s) = 1 - \left[ 1 - F(s) \right]^2 \]

in which

\[ F(s) = \frac{1}{(2\pi)^{1/2} \cdot \text{stddev}(S)} \int_{-\infty}^{\infty} \exp\left[-\left( \frac{u - \text{median}(S)}{\text{stddev}(S)} \right)^2 \right] du \]

for the normal distribution;

\[ F(s) = \left( 1 + \exp\left(-\pi \cdot \left( s - \text{median}(S) / \left[3^{1/2} \cdot \text{stddev}(S)\right]\right)\right) \right)^{-1} \]

for the logistic distribution;

\[ F(s) = 1 - \exp\left( -\exp\left( + \pi \cdot [s - \text{median}(S) - 0.28577 \cdot \text{stddev}(S)] / \left[6^{1/2} \cdot \text{stddev}(S)\right]\right) \right) \]

for the smallest-extreme-value distribution; and

\[ F(s) = \exp\left( -\exp\left( -\pi \cdot [s - \text{median}(S) + 0.28577 \cdot \text{stddev}(S)] / \left[6^{1/2} \cdot \text{stddev}(S)\right]\right) \right) \]

for the largest-extreme-value distribution. Note that, as in microcomputer programs NIA, LIA, SEVIA, and LEVIA, the respective conceptual (two-parameter) strength (resistance) distributions are deliberately scaled to have the same (presumed known or guesstimated) standard deviations.

Little (1975) presents tabulated quantities pertaining to each of these four alternative conceptual strength (resistance) distribution models for use in estimating (i) the actual value for median(S) with two specimens “in series”, and (ii) the associated enumeration-based (exact) statistical bias corrections, given the outcome of a conventional small sample up-and-down experiment test program with a nominal sample size from two to six.

**Corollary:** The CDF for the weakest of \( n \) nominally identical items is given by

\[ F_{\text{weakest}}(s) = 1 - \left[ 1 - F(s) \right]^n \]
8.5 COMPARATIVE ANALYSIS FOR THE OUTCOMES OF LIFE (ENDURANCE) AND STRENGTH (RESISTANCE) EXPERIMENT TEST PROGRAMS

The ML analyses presented in Section 8.4 all pertain to quantitative CRD experiment test programs. Recall that, for a quantitative estimate to have credible statistical inference, it must be presumed that (i) all possible batch-to-batch effects are negligible, (ii) all nuisance test variables are negligible, and (iii) all calibrations are accurate. Fortunately, the likelihood ratio (LR) method provides a statistical means to make comparative (rather than quantitative) mechanical reliability analyses. This feature of the LR method is illustrated in Section 8.5A.

8.5A COMPARATIVE MAXIMUM LIKELIHOOD ANALYSIS (USING THE LIKELIHOOD RATIO TEST)

Suppose that two $s_{\alpha}^{\text{fnc}}$ experiment test programs pertaining to different batches of the same material have been conducted and the issue is whether the respective data sets can be legitimately combined. Microcomputer program C2SFNCM7 (Compare 2 $s_{\alpha}^{\log_e[fnc(p)]}$ Models - version 7) performs a LR-based analysis, presuming a conceptual $s_{\alpha}^{\log_e[fnc(p)]}$ model in a conditional ML analysis with the homoscedastic conceptual (two-parameter) smallest-extreme-value fatigue strength distribution augmented parameterization $y = \{s_{\alpha} - clp0 - clp1 \log_e(fnc) - clp2 [\log_e(fnc)]^2\}/csp$. The example input $s_{\alpha}^{\text{fnc}}$ data for this microcomputer program are constructed using the $s_{\alpha}^{\text{fnc}}$ values (data set one) that appear in microcomputer file SAFNCDTA, plus additional $s_{\alpha}^{\text{fnc}}$ values (data set two) whose $s_{\alpha}$’s are computed by subtracting 20 from each of the corresponding $s_{\alpha}$’s in data set one. (See microcomputer file C2SNDATA.) A linear metric for the alternating stress amplitude was deliberately used in constructing these example input $s_{\alpha}^{\text{fnc}}$ data to generate parallel $s_{\alpha}^{\log_e[fnc(p)]}$ models for the respective $s_{\alpha}^{\text{fnc}}$ data sets. Accordingly, the respective [est(clp1)]’s and [est(clp2)]’s pertaining to our two example $s_{\alpha}^{\text{fnc}}$ data sets will be identical and that the respective [est(clp0)]’s will differ by exactly 20 units. But, regardless of the alternating stress amplitude metric that is presumed for the conceptual $s_{\alpha}^{\log_e[fnc(p)]}$ model, the fundamental issue is that the respective data-based values for the asymptotic LR test statistic will establish null hypothesis rejection probability values for the two statistical tests of hypothesis that are of specific interest: (i) the null hypothesis that the correct statistical model is a single $s_{\alpha}^{\log_e[fnc(p)]}$ model versus the alternative hypothesis that the correct statistical model is two distinct $s_{\alpha}^{\log_e[fnc(p)]}$ models, and (ii) the null hypothesis that the correct statistical model is two parallel $s_{\alpha}^{\log_e[fnc(p)]}$ models with a common $csp$ versus the alternative hypothesis that the correct statistical model is two distinct $s_{\alpha}^{\log_e[fnc(p)]}$ models. Given our constructed exemplar $s_{\alpha}^{\log_e[fnc]}$ datum set two, we must rationally opt to reject null hypothesis (i) in favor of alternative hypothesis (i), but we cannot rationally reject null hypothesis (ii) in favor of alternative hypothesis (ii). Accordingly, we must statistically conclude that a batch-to-batch effect exists, but that this batch-to-batch effect is not affected by the actual value of the concomitant variable $s_{\alpha}$.

Analogously, suppose that two strength (endurance) experiment test programs pertaining to different batches of the same material have been conducted at two or more stress (resistance) levels and the issue is whether the respective data sets can be legitimately combined. Microcomputer program C2NSDDS (Compare 2 Normal Strength Distribution Data Sets) performs a LR-based analysis, presuming that the four respective strength (resistance) models of potential interest each employ an appropriate conceptual (two-parameter) normal distribution. It first computes the [est(clp)]’s and the [est(csp)]’s for each of these strength (resistance) models. Then it computes the data-based values for the two asymptotic LR test statistics of specific interest and establishes the associated null rejection probabilities for (i) the null hypothesis that the correct statistical model is a single conceptual (two-parameter) normal strength
distribution versus the alternative hypothesis that the correct statistical model is two distinct conceptual (two-parameter) normal strength distributions, and (ii) the null hypothesis that the correct statistical model is two distinct conceptual (two-parameter) normal strength distributions with a common csp versus the alternative hypothesis that the correct statistical model is two distinct conceptual (two-parameter) normal strength distributions. Given the example data sets that appear in microcomputer file C2SDDATA, LR-based analysis indicates that we cannot rationally opt to reject null hypothesis (i) in favor of alternative hypothesis (i). Accordingly, we opt to aggregate the two example data sets and employ a statistical model with a single conceptual (two-parameter) normal strength distribution in subsequent analysis. Moreover, we statistically conclude that the batch-to-batch effect, if it exists, is evidently relatively small compared to the actual value for the standard deviation of the presumed conceptual (two-parameter) normal strength distribution.

Remark: LR tests typically require statistical bias corrections when quantitatively employed (to compute statistical confidence intervals and limits), but not when comparatively employed (to test the statistical adequacy of a proposed model).

```
C> TYPE C2SNDATA
8
1 320  56430
1 300  99000
1 280  183140
1 260  479490
1 240  909810
1 220  3632590
1 200  4917990
1 180  19186790
0
8
1 300  56430
1 280  99000
1 260  183140
1 240  479490
1 220  909810
1 200  3632590
1 180  4917990
1 160  19186790
0
```

Data Set One (Same as in Microcomputer File SAFNCDTA)

Data Set Two (Constructed by subtracting 20 from each alternating stress amplitude in Data Set One)
Given no censoring and assuming a conceptual $s_\alpha \log_e[\text{fn}(pf)]$ model in a conditional ML analysis with the homoscedastic smallest-extreme-value fatigue strength distribution augmented parameterization

$$y = s_\alpha - c_{lp}0 - c_{lp}1 \log_e(\text{fn}) - c_{lp}2 \left[\log_e(\text{fn})\right]^2/csp$$

(and with no statistical bias corrections)

Employing a distinct $s_\alpha \log_e[\text{fn}(pf)]$ model for data set one
- ML est($c_{lp}0$) = 0.748120D+03
- ML est($c_{lp}1$) = -0.500048D+02
- ML est($c_{lp}2$) = 0.973366D+00
- ML est($csp$) = 0.407381D+01

Estimated Maximum log$_e$(Likelihood) = -2.3424546D+02

Employing a distinct $s_\alpha \log_e[\text{fn}(pf)]$ model for data set two
- ML est($c_{lp}0$) = 0.728120D+03
- ML est($c_{lp}1$) = -0.500048D+02
- ML est($c_{lp}2$) = 0.973366D+00
- ML est($csp$) = 0.407381D+01

Estimated Maximum log$_e$(Likelihood) = -2.3424546D+02

Employing two parallel $s_\alpha \log_e[\text{fn}(pf)]$ models with a common csp
- ML est($c_{lp}01$) = 0.748120D+03
- ML est($c_{lp}02$) = 0.728120D+03
- ML est($c_{lp}1$) = -0.500048D+02
- ML est($c_{lp}2$) = 0.973366D+00
- ML est($csp$) = 0.407381D+01

Estimated Maximum log$_e$(Likelihood) = -4.6849092D+02

Employing a single $s_\alpha \log_e[\text{fn}(pf)]$ model (for the aggregated data)
- ML est($c_{lp}$) = 0.762663D+03
- ML est($c_{lp}1$) = -0.528957D+02
- ML est($c_{lp}2$) = 1.06880D+01
- ML est($csp$) = 0.962249D+01

Estimated Maximum log$_e$(Likelihood) = -6.1152993D+02

Likelihood Ratio Tests

(i) $H_0$: Two distinct $s_\alpha \log_e[\text{fn}(pf)]$ models versus $H_1$: A single $s_\alpha \log_e[\text{fn}(pf)]$ model (8-4 $n_{ald}$)
$H_1$ Rejection Probability = 0.9391D-05

(ii) $H_0$: Two distinct $s_\alpha \log_e[\text{fn}(pf)]$ models versus $H_1$: Two parallel $s_\alpha \log_e[\text{fn}(pf)]$ models with a common csp (8-5 $n_{ald}$)
$H_1$ Rejection Probability = 1.0000D+01

Also (ii) $H_0$: Two parallel $s_\alpha \log_e[\text{fn}(pf)]$ models with a common csp versus $H_1$: A single $s_\alpha \log_e[\text{fn}(pf)]$ model (5-4 $n_{ald}$)
$H_1$ Rejection Probability = 0.8862D-07

Remark: Note that if the experiment test program objective is to compare $s_\alpha \log_e(\text{fn})$ models pertaining to treatments $B$ and $A$, then data set one would pertain to treatment $B$ and data set two would pertain to treatment $A$. 
C> TYPE C2SDDATA

<table>
<thead>
<tr>
<th>4</th>
<th>Number of Data Set One $s_1$'s</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>90</td>
<td>6</td>
</tr>
<tr>
<td>80</td>
<td>4</td>
</tr>
<tr>
<td>70</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Number of Data Set Two $s_2$'s</td>
</tr>
<tr>
<td>95</td>
<td>8</td>
</tr>
<tr>
<td>85</td>
<td>7</td>
</tr>
<tr>
<td>65</td>
<td>6</td>
</tr>
</tbody>
</table>

C> COPY C2SDDATA DATA

1 file(s) copied

C> C2NSDDS

Presuming a conceptual (two-parameter) normal distribution strength (resistance) model in ML analysis with the conventional parameterization

$$y = c_{lp} + c_{sp} s$$

(and with no statistical bias corrections)

Employing a distinct strength (resistance) model for data set one

ML est[$s(50)$] = 86.1  ML est[$c_{sp}$] = 6.8
Estimated Maximum log$_e$(Likelihood) = -6.2841

Employ a distinct strength (resistance c) model for data set two

ML est[$s(50)$] = 80.2  ML est[$c_{sp}$] = 14.5
Estimated Maximum log$_e$(Likelihood) = -10.5833

Employing two distinct strength (resistance) models that have a common $c_{sp}$

ML est[$s(50)$] for data set one = 86.0
ML est[$s(50)$] for data set two = 81.0
ML est(common $c_{sp}$) = 11.1
Estimated Maximum log$_e$(Likelihood) = -17.7893

Employing a single strength (resistance) model (for the aggregated data)

ML est[$s(50)$] = 83.3
ML est[$c_{sp}$] = 11.2
Estimated Maximum log$_e$(Likelihood) = -18.2358

**Likelihood Ratio Tests**

(i) $H_0$: Two distinct normal distributions versus $H_1$: A single normal distribution (4–2 $n_{sub}$)

$H_0$ Rejection Probability = .2545

(ii) $H_0$: Two distinct normal distributions versus $H_1$: Two distinct normal distributions that have a common $c_{sp}$ (4-3 $n_{sub}$)

$H_0$ Rejection Probability = .1745

Also (iii) $H_0$: Two distinct normal distributions that have a common $c_{sp}$ versus $H_1$: A single normal distribution (3–2 $n_{sub}$)

$H_0$ Rejection Probability = .3447
8.5B DISTRIBUTION-FREE (NON-PARAMETRIC) COMPARATIVE ANALYSIS FOR THE OUTCOME OF A LIFE (ENDURANCE) EXPERIMENT TEST PROGRAM WITH ARBITRARILY SUSPENDED TESTS

We now re-consider an experiment test program that is conducted to test the null hypothesis that \( (B - mpd) = A \) statistically versus the alternative hypothesis that \( (B - mpd) > A \) statistically, where \( mpd \) denotes the minimum practical difference of specific interest in adopting \( B \) in preference to \( A \). Suppose that certain of the tests in this experiment test program are arbitrarily suspended. (Recall however that the fundamental statistical concept of a continually replicated experiment test program is obscure when arbitrarily suspended tests occur.) Suppose also that the log-rank algorithm [Mantel (1981)] is used to assign ranks to the aggregated \( A \) and \( B \) experiment test program life datum values. Then Savage’s distribution-free (non-parametric) rank test statistic can be generalized to include arbitrarily suspended tests. It is employed in microcomputer program C2DSWAST (Compare 2 Data Sets With Arbitrarily Suspended Tests) to test the null hypothesis that \( (B - mpd) = A \) statistically versus the alternative hypothesis that \( (B - mpd) > A \) statistically. In turn, a lower 100(\( sep \))\% (one-sided) confidence limit that allegedly bounds the actual value for the \( mpd \) can be computed as outlined in Supplemental Topic 3.C.

Remark: Savage’s rank test statistic is employed in microcomputer program C2DSWAST because, given a continually replicated CRD experiment test program with complete data and two treatments, it is asymptotically unexcelled.

Microcomputer program C2DSWAST can be used to compare service life (endurance) data for alternative designs by asserting that arbitrarily suspended tests are “statistically analogous” to products that are still operating satisfactorily in service. It can also be used to compare life (endurance) data for alternative designs generated in a statistically planned experiment test program with actual Type I censoring.

```
C> TYPE C2ASTDATA
200 Minimum Practical Difference of Specific Interest
10 Number of A Failures, Followed by Each A Failure Life
112
143
151
177
231
345
378
401
498
512
2 Number of A Suspended Tests, Followed by Each A Life at Test Suspension
1000.01
1000.01
2 Number of B Failures, Followed by Each B Failure Life
592
712
5 Number of B Suspended Tests, Followed by Each B Life at test Suspension
1000.01
1000.01
1000.01
1000.01

Note. An increment equal to 0.01 is added to each datum value pertaining to a suspension to avoid the possibility of a tie between a failure life and a life at test suspension
```
C> COPY C2ASTDTRA DATA
1 file(s) copied
C> C2DSWAST

For a minimum practical difference (mpd) = 200.0

<table>
<thead>
<tr>
<th>Aggregated a and (b-mpd) Datum Values</th>
<th>code</th>
<th>log-rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>.11200000000D+03</td>
<td>1</td>
<td>.10000000000D+01</td>
</tr>
<tr>
<td>.14300000000D+03</td>
<td>1</td>
<td>.20000000000D+01</td>
</tr>
<tr>
<td>.15100000000D+03</td>
<td>1</td>
<td>.30000000000D+01</td>
</tr>
<tr>
<td>.17700000000D+03</td>
<td>1</td>
<td>.40000000000D+01</td>
</tr>
<tr>
<td>.23100000000D+03</td>
<td>1</td>
<td>.50000000000D+01</td>
</tr>
<tr>
<td>.34500000000D+03</td>
<td>1</td>
<td>.60000000000D+01</td>
</tr>
<tr>
<td>.37800000000D+03</td>
<td>1</td>
<td>.70000000000D+01</td>
</tr>
<tr>
<td>.39200000000D+03</td>
<td>3</td>
<td>.80000000000D+01</td>
</tr>
<tr>
<td>.40100000000D+03</td>
<td>1</td>
<td>.90000000000D+01</td>
</tr>
<tr>
<td>.49800000000D+03</td>
<td>1</td>
<td>.10000000000D+02</td>
</tr>
<tr>
<td>.51200000000D+03</td>
<td>1</td>
<td>.11500000000D+02</td>
</tr>
<tr>
<td>.51200000000D+03</td>
<td>3</td>
<td>.11500000000D+02</td>
</tr>
<tr>
<td>.10000100000D+04</td>
<td>0</td>
<td>.16000000000D+02</td>
</tr>
<tr>
<td>.10001000000D+04</td>
<td>0</td>
<td>.16000000000D+02</td>
</tr>
<tr>
<td>.80001000000D+03</td>
<td>2</td>
<td>.16000000000D+02</td>
</tr>
<tr>
<td>.80001000000D+03</td>
<td>2</td>
<td>.16000000000D+02</td>
</tr>
<tr>
<td>.80001000000D+03</td>
<td>2</td>
<td>.16000000000D+02</td>
</tr>
<tr>
<td>.80001000000D+03</td>
<td>2</td>
<td>.16000000000D+02</td>
</tr>
</tbody>
</table>

The data-based value of the generalized Savage test statistic for the CRD experiment test program that was actually conducted is equal to 11.195

Given the null hypothesis that \((B-\text{mpd}) = A\) statistically, this microcomputer program constructed exactly 50388 equally-likely outcomes for this experiment test program by using Ehrlich’s method to re-assign its \(a\) and \((b-\text{mpd})\) datum values to treatments \(A\) and \(B\). The number of these outcomes that had its generalized Savage test statistic value equal to or greater than 11.195 is equal to 232. Thus, given the simple null hypothesis that \((B-\text{mpd}) = A\) statistically, the enumeration-based probability that a randomly selected outcome of this experiment test program when continually replicated will have its generalized Savage test statistic value equal to or greater than 11.195 is equal to .0046. When this probability is sufficiently small, reject the null hypothesis in favor of the simple (one-sided) alternative hypothesis that \((B-\text{mpd}) > A\) statistically.

Given the null hypothesis that \((B-\text{mpd}) = A\) statistically, the asymptotic probability that a randomly selected outcome of this experiment test program when continually replicated will have its generalized Savage test statistic value equal to or greater than 11.195 is equal to .0054

Note:  Code 0 = suspended \(A\) test life

Code 1 = \(A\) datum value

Code 2 = suspended \(B\) test life – \text{mpd}

Code 3 = \(B\) datum value – \text{mpd}

When a paired-comparison test program is clearly appropriate, arbitrarily suspended tests within blocks can pose problems in establishing the relevant \((b-a)\) difference, e.g., when the suspended test pertains to a shorter duration than the corresponding failure life (endurance). However, given Type I censoring, the durations of all suspended tests are greater than their corresponding observed failure lives. Then all paired-comparison \((b-a)\) differences are either known or can be treated as arbitrarily suspended tests. When the paired-comparison \((b-a)\) difference is positive, it is treated as \(B\) data in a CRD test program. On the other hand, when the paired-comparison \((b-a)\) difference is negative, it is treated as \(A\) data in this
CRD test program. (If both specimens fail at the same time or if both tests are suspended at the same time, the resulting paired-comparison \((b-a)\) difference is equal to zero and this paired-comparison outcome can be treated as a tie.) Accordingly, microcomputer program \(C2DSWAST\) can be used to test the null hypothesis \((B-mpd) = A\) statistically versus the alternative hypothesis that \((B-mpd) > A\) statistically for any \(mpd\) value of specific interest, even when a paired-comparison experiment test program is conducted.

8.6 ESTIMATING SUB-SYSTEM RELIABILITY

A system is comprised of sub-systems. In turn, sub-systems are comprised of sub-sub-systems, and sub-sub-systems are comprised of sub-sub-sub systems, et cetera. Because systems, sub-systems, sub-sub-systems, and sub-sub-sub-systems can be parsed in various ways, their distinction is often a matter of semantics. We choose to use the terminology sub-system in our presentation.

The proper method to estimate the reliability of a mechanical (electro-mechanical) sub-system is to conduct a life (endurance) experiment test program using the sub-system itself as the test specimen. However, if this experiment-based methodology is not practical, then the reliability of a mechanical (electro-mechanical) sub-system can be analytically guestimated using estimates of the reliabilities of its components. In theory, the reliability of a mechanical (electro-mechanical) sub-system can analytically expressed as a function of duration \(d\) when \((i)\) all of its components are presumed to operate independently and \((ii)\) the reliabilities of these components are expressed as a function of, or in terms of, the same stimulus and duration metrics. However, in practice, the respective component reliabilities must be estimated by the appropriate life (endurance) experiment test programs whose stimulus and duration metrics are unlikely to be identical and may not even be coherent (compatible). Accordingly it is extremely unlikely that the necessary component reliability estimates exist — unless these estimates were deliberately generated for use in estimating the sub-system reliability of specific interest. Even then, the only way to test the credibility of the presumption that the respective components operate independently within this sub-system is to test this sub-system itself. Despite these practical limitations, sub-system reliability estimates based on questionable presumptions and analyses can nevertheless have useful application in comparing alternative designs and re-designs in an iterative reliability improvement process.

We now define a mechanical (electro-mechanical) sub-system as being comprised of components in a specific configuration, where a component by definition has no redundancy. Accordingly, at any given duration, each of the components in a sub-system is either operating satisfactorily or it is not. In contrast, depending on the configuration of its components, a sub-system can (should) have redundancy, viz., certain components can fail and the sub-system will continue to operate satisfactorily.

The physical understanding of a sub-system with redundancy is enhanced by limiting the sub-system to having a maximum of six to eight components. This can be accomplished by parsing larger sub-systems into smaller sub-systems (that could be called sub-sub-systems or sub-sub-sub-systems if so desired.) Consider a sub-system that is comprised of \(n_c\) components. Its reliability can be established by examining the operational state for each of these \(n_c\) components. This examination requires that each of the \(2^n\) distinct sets of component operational states be enumerated in terms of whether each individual component is operating satisfactorily or not.

Consider a component denoted \(A\). Let (capital) \(A\) connote that it is operating satisfactorily and let (lower case) \(a\) denote that it is not operating satisfactorily. Next consider a sub-system composed of four components: \(A, B, C\) and \(D\). The reliability of this sub-system can be estimated by enumerating all \(2^4 (= 16)\) of its distinct sets of component operational states. The required enumeration is conveniently accomplished using Yate's enumeration algorithm in which, e.g., \(a\) corresponds to \(-1\) and \(A\)
corresponds to +1 in column one, \( b \) corresponds to -1 and \( B \) corresponds to +1 in column two, et cetera. The resulting 16 distinct sets of component operational states for this sub-system example are:

\[
\begin{align*}
& a b c d \\
& A b c d \\
& a B c d \\
& A B c d \\
& a b C d \\
& A b C d \\
& a B C d \\
& A B C d \\
& a b c D \\
& A b c D \\
& a B c D \\
& A B c D \\
& a b C D \\
& A b C D \\
& a B C D \\
& A B C D \\
\end{align*}
\]

To simplify subsequent notation, let \( A \) also denote the estimated probability that component \( A \) is operating satisfactorily (\( A = \) estimated reliability) and let \( a \) also denote the complementary probability that component \( A \) is not operating satisfactorily (\( a = \) estimated unreliability).

**The Reliability of a New Sub-System:** Now consider a new sub-system with all of its components operating satisfactorily at duration \( d = 0 \). For simplicity, let this sub-system have only two components, say \( A \) and \( B \). The 4 distinct sets of component operational states for this sub-system are:

<table>
<thead>
<tr>
<th>component operational states</th>
<th>associated probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ab )</td>
<td>( a\cdot b )</td>
</tr>
<tr>
<td>( Ab )</td>
<td>( A\cdot b )</td>
</tr>
<tr>
<td>( aB )</td>
<td>( a\cdot B )</td>
</tr>
<tr>
<td>( AB )</td>
<td>( A\cdot B )</td>
</tr>
</tbody>
</table>

The estimated reliability of this sub-system can be established by summing the respective estimated probabilities pertaining to each distinct set of component operational states such that the sub-system continues to operate satisfactorily. When the estimated reliabilities for components \( A \) and \( B \) have been established by either appropriate life (endurance) or strength (resistance) experiment test programs, the reliability of the sub-system can be tabulated up to the duration used in Type I censoring for the life (endurance) experiment test programs, or at the specific duration \( d^* \) employed in the strength experiment test programs to suspend each individual test.

**Elementary Example One:** Suppose that a sub-system is configured to have no redundancy, viz., all of its components must operate satisfactorily for the sub-system to operate satisfactorily. Then only the distinct component operational state set \( AB \) pertains. Accordingly, the estimated reliability of this sub-system is the product \( A\cdot B \). Note that if this Example One sub-system had been comprised of numerous components, then its estimated reliability would have been the product of each of the individual estimated component reliabilities. Clearly this estimated reliability becomes smaller and smaller as the number of components increases. Clearly this type of sub-system configuration should be avoided whenever practical.
**Elementary Example Two:** Suppose that a sub-system is configured to have maximum redundancy, viz., only one of its components must operate satisfactorily for the sub-system to operate satisfactorily. Then only distinct component operational state set \( ab \) fails the condition for satisfactory operation. Accordingly the estimated unreliability of this sub-system with maximum redundancy is the product \( a \cdot b \). Note that if this Example Two sub-system had been comprised of numerous components, then its estimated unreliability would have been the product of the respective estimated component unreliabilities. Clearly this estimated unreliability becomes smaller and smaller as the number of its redundant components increases.

**Message:** Redundant components increase sub-system reliability and therefore should be used whenever practical.

Elementary Examples One and Two are intended to demonstrate that the reliability of a sub-system decreases with an increase in the number of its components, unless the sub-system is appropriately configured with redundant components. Unfortunately mechanical (electro-mechanical) sub-systems are never quite as simple as either of these two elementary examples. Nevertheless, we now present three additional elementary examples that are intended to demonstrate the computational simplicity of estimating sub-system reliability when \( (i) \) all distinct sets of component operational states have been enumerated, and \( (ii) \) the component operational states that pertain to satisfactory sub-system operation can be unambiguously defined (or the distinct sets of component operational states that pertain to unsatisfactory sub-system operation can be unambiguously defined), and \( (iii) \) all components are presumed to operate independently.

**Elementary Example Three:** Suppose satisfactory sub-system operation is defined such that at least two of its four components must operate satisfactorily. Then, by inspection, the estimated reliability of this sub-system is the sum of the products of the respective estimated probabilities in the second main-column of the following example array:

<table>
<thead>
<tr>
<th>component operational states</th>
<th>Elementary Example Three (sum these estimated probability products)</th>
<th>Elementary Example Four (sum these estimated probability products)</th>
<th>Elementary Example Five (sum these estimated probability products)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a b c d )</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( A b c d )</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( a B c d )</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( A B c d )</td>
<td>( A B c d )</td>
<td>-</td>
<td>( A B c d )</td>
</tr>
<tr>
<td>( a b C d )</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( A b C d )</td>
<td>( A b C d )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( a B C d )</td>
<td>( a B C d )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( A B C d )</td>
<td>( A B C d )</td>
<td>( A B C d )</td>
<td>( A B C d )</td>
</tr>
<tr>
<td>( a b c D )</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( a b D )</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( a B c D )</td>
<td>( a B c D )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( a B c D )</td>
<td>( A B c D )</td>
<td>( A B c D )</td>
<td>( A B c D )</td>
</tr>
<tr>
<td>( A B c D )</td>
<td>( a b C D )</td>
<td>-</td>
<td>( a b C D )</td>
</tr>
<tr>
<td>( a b C D )</td>
<td>( a b C D )</td>
<td>( a b C D )</td>
<td>( a b C D )</td>
</tr>
<tr>
<td>( a B C D )</td>
<td>( A B C D )</td>
<td>( A B C D )</td>
<td>( A B C D )</td>
</tr>
</tbody>
</table>

**Elementary Example Four:** Suppose that at least three components must operate satisfactorily for this sub-system to operate satisfactorily. Then, by inspection, the estimated reliability for this sub-system is the sum of the products of the respective estimated probabilities given by the entries in the third main-column of our example array.
Elementary Example Five: Suppose that this sub-system operates satisfactorily only when both A and B or when both C and D operate satisfactorily. Accordingly, by inspection, the estimated reliability of this sub-system is the sum of the products of the respective estimated probabilities given by the entries in the fourth main-column of our example array.

Discussion: These five elementary examples are intended to demonstrate that the estimated reliability of a sub-system can be established by enumerating all distinct sets of component operational states. This estimation methodology has three basic steps. (i) Enumerate all component operational states using Yate’s enumeration algorithm. (ii) Identify those distinct sets of component operational states that pertain either to satisfactory or unsatisfactory sub-system operation. (iii) Sum the products of the associated estimated probabilities. Step (ii) requires a special microcomputer program for each different sub-system — because the specifications defining satisfactory or unsatisfactory operation for each different sub-system involve a distinct set of computer logic statements. This special microcomputer program is easily obtained by inserting the logic statements appropriate to the sub-system into the code for a generic enumeration microcomputer program.

When the generic enumeration microcomputer program pertains to more than six to eight components, the number of component operational states may be so large that physical understanding is compromised. Accordingly, each sub-system should be established (defined) such that its completely enumerated operational states are all easily comprehensible.

The Reliability of a Repaired Sub-System: We now presume that when a component fails, the failed component is eventually replaced (say at a routinely scheduled maintenance) even if the sub-system continues to operate satisfactorily. Then the respective durations endured by different components in the repaired sub-system can (will) differ. Satisfactory replacement of each failed component depends on (i) its availability, and (ii) the duration required for its satisfactory replacement, where (i) must be stated in terms of an (estimated probability)-(duration to obtain) expression and (ii) must be expressed using a (estimated probability)-(duration to its satisfactory replacement) expression. Accordingly, reliability estimation for a repaired sub-system is best handled by simulation (with a minimum of 1,000 trials).

To begin our discussion of a simulation-based estimate of sub-system reliability, presume that all components operate satisfactorily at duration \( d = 0 \). Then, given estimated reliabilities for each of the respective components generated by appropriate life (endurance) experiment test programs, let duration \( d \) be successively incremented by an appropriately small duration interval \( di \) during each simulation trial. In turn, for each such appropriately small duration interval, let a uniform pseudorandom number generator, zero to one, be used to establish whether each respective component fails during that interval. However, a component cannot fail in the interval of specific interest unless it has survived to the beginning of that interval. Accordingly, for each component in the system, the (enumeration-based) probability of failure in the interval of specific interest is given by the expression

\[
\text{(enumeration-based) probability of failure in the interval of specific interest} = \frac{\int_{d}^{d+\delta} f(u)du}{\int_{d}^{\infty} f(u)du} = \frac{f(d) \cdot (di)}{\int_{d}^{\infty} f(u)du}
\]

in which \( d \) is the duration to the beginning of the appropriately small duration interval \( di \) of specific interest. This probability expression is typically parsed into two terms: (i) the instantaneous failure rate function IFRF and (ii) the appropriately small duration increment \( di \) of specific interest, where
by definition

\[ \text{IFRF} = \frac{f(d)}{1 - F(d)} \]

Clearly the instantaneous failure rate function IFRF is analogous to a probability density function PDF. The IFRF is also commonly called the hazard rate function HRF, which is sometimes re-expressed as

\[ \text{HRF} = \frac{f(d)}{\text{Reliability}(d)} \]

The replacement strategy for redundant components that fail during a given duration interval \( d_i \) must be explicitly defined so that the duration endured by each replacement component is known. (Remember that the duration required for satisfactory replacement of a component can extend over many appropriately short duration intervals.)

Each respective simulation trial is ended when the duration is reached that pertains to the Type I censoring duration used in conducting the respective component life (endurance) experiment test programs. Once all of the (mutually independent) simulation trials have been conducted, numerous statistics will be available for analysis, e.g., (i) the simulation-based estimate of the reliability at duration \( d \) for a new sub-system (which can be compared to its analytically estimated value), or (ii) the average duration to failure for a new sub-system, or (iii) the simulation-based estimate of the reliability at duration \( d \) for a repaired sub-system, or (iv) the average number of failures for each component in the sub-system (which can be compared to its expected value), or (v) the average number of components that had to be repaired before any duration \( d \) of specific interest, etc. A comparison of the relevant statistics for various alternative designs should aid in selecting a more reliable configuration of components in the sub-system of specific interest.

**Exercise Set 4**

(These exercises are intended to provide further insight into sum of probabilities associated with enumerating all possible mutually exclusive and exhaustive outcomes for two or more random variables whose realizations are dichotomous.)

1. Given the array of all distinct sets of component operational states for two components \( A \) and \( B \), arbitrarily select estimated reliability values for components \( A \) and \( B \) and then demonstrate that the sum of all four estimated probability products is one.

2. Given the array of all distinct sets of component operational states for three components \( A, B \) and \( C \), arbitrarily select estimated reliability values for components \( A, B \) and \( C \) and then demonstrate that the sum of all eight estimated probability products is one.

3. Given the array of all distinct sets of component operational states for four components \( A, B, C \) and \( D \), arbitrarily select estimated reliability values for \( A, B, C \) and \( D \) and then demonstrate that the sum of all sixteen estimated probability products is one.
8.7 CLOSURE

The dominance of quantitative CRD experiment test programs presented in this chapter should cause substantial concern regarding the applicability of these analyses in estimating the reliability of a sub-system (or even a component) in service operation. All CRD experiment test programs presume nominally identical test specimens and test conditions, whereas mechanical components are always produced in batches. Accordingly, it is seldom if ever reasonable to presume that a quantitative CRD test program will generate an unbiased estimate of any conceptual statistical model parameter of specific interest. Moreover, it is likely that vendor-supplied prototype components will markedly excel subsequent vendor-supplied production components. In estimating mechanical reliability, skepticism and investigation are always preferable to credulity and supinit.

All competent mechanical design methodologies attempt to establish adequate sub-system (component) reliability by extrapolating well-documented service-proven reliability performance for similar sub-systems (components) to the sub-system (component) design of specific interest. The statistical and test planning tools presented herein can be used to enhance these mechanical design methodologies. But, regardless of the amount and type of design analysis and prototype testing involved, and regardless of the associated statistical methodology, the actual value for the reliability of a product can only be demonstrated by its service performance.