INSTRUCTIONS FOR WDAS
WEIBULL DISTRIBUTION ANALYSIS SHEET

TABLE OF CONTENTS

1.0 Introduction
2.0 Properties of the Weibull
3.0 Procedure for Plotting Raw Data
   3.1 Grouped Data, \( n \geq 15 \)
   3.2 Individual Data, \( n \leq 14 \)
   3.3 Sudden Death
   3.4 Missing Failure Data
4.0 Procedure for Estimating Distribution Values
   4.1 Estimating Weibull Parameters
   4.2 Reading Distribution Statistics
5.0 Exceptions to General Procedure
   5.1 Multimodal Plot
   5.2 Curved Plot
   5.3 Points Fall Outside of Probability Grid
6.0 Confidence & Statistical Tolerance Limits
   6.1 Confidence Limits on Weibull Plot
   6.2 Statistical Tolerance Limits
7.0 Procedure for Given Parameters
8.0 Acknowledgement
9.0 Copyright Release
10.0 References
11.0 Bibliography

Appendix

WDAS Mathematical Relationships
Table I Median Ranks
Table II Scale Parameter, \( \alpha \)
Table III Confidence Limits
Table IV Statistical Tolerance Limits
Chart I "a" & "b" Nomograph, \( m \leq 1.5 \)
1.0 INTRODUCTION

The Weibull Distribution Analysis Sheet, WDAS (form 8933), is a graphic technique for analyzing data when the underlying probability distribution is the Weibull. It is a refinement of the short method (Reference 1) used in the reliability analysis of life test data. It also has many other applications in the statistical analysis of data when the underlying probability distribution is non-normal, exponential or otherwise skewed. Typical applications include the evaluation of the following:

- Concentricity, TIR (Figure 1)
- Depth of Penetration of Spot welds
- Tensile Strength of Metals
- Yield Point of Metals

From both empirical data and for physical reasons (TIR approaching 0), these applications tend to conform to a non-normal probability distribution.

The WDAS procedure features a technique for grouping raw data, accumulating the frequencies using a special method developed by Arthur Bender, Jr. (Reference 2), plotting points on a Weibull probability grid developed by Professor John H. Kao (Reference 3), and fitting a line to the plotted points. Once the line has been fitted, the three parameters of Scale, \( \alpha \); Shape, \( \beta \); and Location, \( \gamma \); which completely describe the Weibull Distribution, can be estimated.

Using the constructed line, statistics such as the measures of central tendency, the measure of dispersion and the percent of product out of specification can be read from the graph. In reliability analysis the additional statistics of Characteristic Life, Reliable Life, Reliability Function, and the Initial Failure Rate can also be read from the same line. Although not available on the WDAS, statistical confidence and tolerance limits can be obtained by using the tables and the method given in this instruction.

*Available from West Allis Stationary Stores, Dept. 1816.

September, 1968

---

**FIGURE 1 COMPLETED WDAS CONCENTRICITY**

---

qe-202
3.0  PROCEDURE FOR PLOTTING RAW DATA

3.1 Grouped Data, \( n \geq 15 \)

The data shown on the Data Sheet, Figure 5, will be used to demonstrate this technique.

A. Determine cell midpoints and boundaries from the raw data:
1. Caution should be exercised in selecting too few or too many cells. Too few cells result in a large cell interval which reduces the accuracy of the answers because many values are represented by one cell midpoint. They also reduce the number of data points to be plotted. The more points on the graph, the more accurately a line can be fitted to them. Going to the other extreme of having a midpoint represent an individual value, the benefit of grouping data and reducing the total number of points which must be plotted is lost. Cell intervals should be selected so that a minimum of 10 to 15 data points will be plotted on the graph, taking into consideration the range of the raw data.

2. Determine the smallest and largest values (300 and 9.100 respectively).

3. Subtract the smallest from the largest value to obtain the range (9,100-300 = 8,800). Considering the range (8,800), determine the size of the Cell Interval. Usually 1, 2, 5, 10 or multiples thereof are found to be convenient. (Cell Interval of 500 hours selected.)
8. Record cell boundaries in column 1 and midpoints in column 2 of WDAS (Figure 6). (For the sake of simplicity the values of the cell boundaries and the cell midpoints have been coded by dividing by 100 and recorded in columns 1 and 2, cell boundaries: 2.5, 7.5, 12.5, etc. and cell midpoints: 5, 10, 15, etc.)

9. The heavy lines in columns 1 and 2 are used to center the data on the form and to locate the extremes.

B. Tally raw data in column 3. (Figure 6)

C. Count tally marks. Record frequencies in column 4.

D. Total column 4. Record sum in box No. 5. This total frequency is the sample size. (35)

E. Double the value of the total in column 5 (35 x 2). Record product in No. 6 (70).

F. For each cell midpoint, accumulate the frequencies in column 4 as follows and record the accumulation in column 7. (Figure 6): 
   1. Consider each block in column 7 as the lower right hand corner of a 4-block matrix in columns 4 and 7.
   2. The number to be recorded in the block in paragraph F above is the sum of the numbers in all other blocks of the matrix.

3. If any block in column 4 is blank, enter a zero.

4. The value in column 7 one line below the last entry in column 4 must be the same as the double total (70) in box 6. This value is a check point. Place parentheses around it to indicate it is not to be used in future operations.

G. For each cell midpoint compute the percent of CDF, sample percent of cumulative distribution function, as follows and record percentages in column 8. (Figure 6)

   1. Divide each value in column 7 by the double total on line 6 (3/70, 11/70, etc.) and multiply each quotient by 100 to obtain the percentage (4.3, 16, 27, etc.).

   2. Exclude those values in column 8 for which there was a zero on the same line in column 4.

   3. Division is simplified if the value in box No. 5, the sample size, is 10, 25, 50, or 100. In these cases the values in column 7 can be multiplied by 5, 2, 1, or \( \frac{1}{2} \) respectively to obtain the percent CDF in column 8.

H. Plot the values in columns 2 and 8 (Figure 6) on the Weibull probability grid using the percent CDF scale, the auxiliary y-axis versus the variable scale, the auxiliary x-axis.

---

**FIGURE 6 PLOTTING GROUPED DATA**

*September, 1968*
3.3 Sudden Death Test Method (Figure 8)

A. Sudden death testing is an economical technique which utilizes a minimum number of samples while providing the effective level of confidence which is associated with larger sample sizes. In a sudden death program the total sample is randomly divided into several groups of equal size. All units in each group are tested at the same time until the first failure occurs. Testing is then terminated on that group. All testing is completed when each group has one failure. This technique features a substantial saving in the overall test time required and a resulting reduction in test costs. Any number of groups and samples within a group can be used; however, a minimum of five groups with at least four samples in each group, a total of 20 samples, constitutes the minimum practical test scheme.

B. The time or cycle of first failure in each of the groups is ranked in ascending numerical order in column 2. (The example shown in Figure 8 consists of 5 groups with 4 samples in each group. The time of the first failure in each group has been recorded in column 2.)

C. From Table I in the Appendix obtain the Median Ranks using the number of groups as the sample size. (Since there are five groups, the Median Ranks for a sample size of 5 is used.)

D. In column 8, record the Median Rank which corresponds to the time or cycle of failure in column 2. (A median rank of 15 is recorded on the first line opposite the first failure of 200 hours, a rank of 31 is recorded opposite the first failure of the second group — 700 hours, etc.)

E. Plot the values in columns 2 and 8 on the Weibull probability grid using the percent CDF, auxiliary y-axis versus the variable scale, auxiliary x-axis.

F. Visually fit a straight line or curve to these points using the method of “least squares” to obtain the best fit. This fitted line or curve is called the Sudden Death Plot. If a curved line is determined to be the best fit, refer to section 5.0 of this instruction for the appropriate procedure.

G. The next step is to construct the Weibull Plot line from the Sudden Death Plot constructed above in paragraph F. This is accomplished by obtaining the point at which the Sudden Death Plot intersects the 50% CDF horizontal line and projecting vertically downward to the point of the first median
Calculating new rank order increments and new median ranks:

\[ R_i = \frac{(N + 1) - (\text{Previous Rank Order No.})}{1 + \text{(No. of items beyond successful sample)}} \]

Modified Median Rank: \[ j' = \frac{.3}{N + .4} \] Where \( N \) = total sample size

Calculations for Figure 9:

<table>
<thead>
<tr>
<th>Sample No.</th>
<th>Indiv. Hours</th>
<th>Mode</th>
<th>( \frac{(N + 1) - P}{1 + n - \text{etc.}} )</th>
<th>( R_i )</th>
<th>New Rank ( j' )</th>
<th>( j' - .3 )</th>
<th>Col. 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>240</td>
<td>Failure</td>
<td>0</td>
<td>1</td>
<td>Appx. Table I 6.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>360</td>
<td>Failure</td>
<td>1</td>
<td>2</td>
<td>Appx. Table I 16%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>559</td>
<td>Success</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>600</td>
<td>Failure</td>
<td>( \frac{10 + 1 - 2^*}{1 + 7} = \frac{9}{8} = 1.125 )</td>
<td>3.125</td>
<td>3.125 - .3 = 27%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>650</td>
<td>Failure</td>
<td>1.125</td>
<td>4.250</td>
<td>4.250 - .3 = 38%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>750</td>
<td>Failure</td>
<td>1.125</td>
<td>5.375</td>
<td>5.375 - .3 = 49%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>835</td>
<td>Success</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>1,100</td>
<td>Failure</td>
<td>( \frac{10 + 1 - 5.375}{1 + 3} = \frac{1.406}{1.4} = 6.781 )</td>
<td>6.781</td>
<td>6.781 - .3 = 62%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1,300</td>
<td>Failure</td>
<td>1.406</td>
<td>8.187</td>
<td>8.187 - .3 = 76%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2,100</td>
<td>Failure</td>
<td>1.406</td>
<td>9.593</td>
<td>9.593 - .3 = 89%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Where \( N \) = Total Sample Size

\( P \) = Previous Rank Order Number

\( * \) Since Sample # 3 did not fail and the rank order of sample ahead is the 2nd failure in order of magnitude of hours of failure \( P = 2 \).

\( n - \text{etc.} \) = Total number of samples (failed and non-failed ones) beyond present non-failure sample.

\( ** \) \( n - \text{etc} = 7 \) since sample numbers 4, 5, 6, 7, 8, 9, & 10 (there are 7 samples) beyond the non-failed sample # 3.

June 1969
D. The data points in columns 2 and 8 are plotted on the Weibull probability grid and a line or curve is fitted to them in the manner described in section 3.2, F through H.

4.0 PROCEDURE FOR ESTIMATING DISTRIBUTION VALUES

This procedure applies to estimating the Weibull Distribution Parameters for Grouped Data, Individual Data, Sudden Death, and Missing Failure Data once the points have been plotted on the Weibull probability grid and a straight line, the Weibull Plot, constructed. Record the values of the distribution statistics in the space provided on the WDAS. Three columns have been provided to accommodate the statistics of more than one Weibull Plot.

4.1 Estimating Weibull Parameters (Figure 10)

A. Location Parameter, \( \hat{\gamma} \)

When the data points follow a linear trend and a straight line is determined to be the best fit, the location parameter, \( \hat{\gamma} \), is estimated to be 0.

If a curve has been fitted to the original data points, use the method described in paragraph 5.2 to estimate \( \gamma \).

B. Shape Parameter, \( \hat{\beta} \)

Construct a straight line parallel to Weibull Plot through the "\( \hat{\beta} \) Estimator Point" (principal axes: 1,0) which will intersect the principal y-axis. From the point of intersection of the constructed line with the principal y-axis project horizontally to the right and read the value of \( m \) from the principal y-axis scale. \( \hat{\beta} \) is the negative value of \( m \). \( \hat{\beta} = -m \). (\( \hat{\beta} = -(-1.6) = 1.6 \))

C. Scale Parameter, \( \hat{\alpha} \)

The y-intercept of the Weibull Plot gives an estimate of \( \hat{\alpha} \) in terms of \( m \). Read the value of \( m \) from the principal y-axis scale and convert to \( \hat{\alpha} \) using Table II in the Appendix. (\( m = 5.45 \), \( \hat{\alpha} = 245 \))

D. Record the values for the Scale, Shape, and Location parameters in the space provided on the WDAS. Three columns are available to provide space for recording the parameters of more than one Weibull Plot.
5.0 EXCEPTIONS TO GENERAL PROCEDURE

5.1 Multimodal Plot (Figure 12)

A. When the data points form 2 distinct straight line trends, the distribution is bimodal. A minimum of 6 data points should be associated with each straight line trend in order to arrive at a statistically valid decision that there is more than one Weibull Plot. Fit two straight lines to the data. Separate the data into two groups and replot each as a separate Weibull Plot. End data points may overlap into each plot. The parameters and statistics for each Weibull Plot are obtained in the same manner as described in Section 4.0. Trimodal distributions are identified by data points forming 3 distinct straight line trends.
increased in magnitude; however, if the curve changes direction, then too high a constant value has been added or subtracted. Lower the magnitude of the constant and replot the points.

F. Estimates of other parameters and distribution statistics are obtained as explained in Section 4, except \( \hat{\alpha} \) is added or subtracted in reverse of whether the constant value was added or subtracted to straighten the curved plot:

- **Weibull Central Tendency or Characteristic Life**
  \[ \hat{\alpha} = Y_o \pm \hat{\gamma} \]

- **Mean**
  \[ \hat{\mu} = a Y_o \pm \hat{\gamma} \]

- **Median**
  \[ M_e = \hat{\xi} \pm \hat{\gamma} \]

- **Reliable Life at**
  \[ \hat{\beta} = (\text{Variable } x) \pm \hat{\gamma} \]

- **Reliability Function**
  \[ R = \text{Variable } (x \pm \gamma) \]

5.3 Points Fall Outside of Probability Grid

A. When the y-intercept of the Weibull Plot with the principal y-axis falls outside of the grid, follow this procedure. From the \( Y_o \) point on the principal y-axis scale project horizontally to the Weibull plot, then vertically up to the principal x-axis and read the value. Multiply this value by \(-\hat{\beta}\).

This product is the value of \( m \), \( a \) can be estimated by using Table II of the Appendix.

B. When the x and y intercepts fall outside of the grid, modify the scale of the auxiliary x-axis by a factor of \( 10 \pm \hat{\gamma} \) so that intercepts of the plot are brought within the limits of the grid.

C. When the slope \( m \) is less than 1.5, estimate the shape parameter, \( \hat{\beta} \), \( a \) and \( b \), as follows. Obtain the y-intercept of the Weibull Plot with the principal y-axis and read the value of \( m \). Select the value of any convenient variable, \( i \), on the principal x-axis, project vertically down to the Weibull Plot and horizontally to the read the value of \( m \) on the principal y-axis. Then

\[ \hat{\beta} = \frac{m - m_i}{i} \]

Obtain the values of \( a \) and \( b \) by entering the nomograph in Chart I of the Appendix with the value of \( m \), project horizontally to the left and read the values of \( a \) and \( b \).

6.0 CONFIDENCE AND STATISTICAL TOLERANCE LIMITS

The methods given in section 4 provide a point estimate (because only one numerical value is obtained of the true population value of the Weibull distribution such as the mean- \( \hat{\mu} \), standard deviation- \( \hat{\sigma} \), characteristic life- \( \hat{\eta} \), etc. Since these estimates of the population values were read from the Weibull Plot which was constructed using median ranks (median ranks are derived such that in a large number of tests 50% of the values would exceed the median rank value while 50% would be less than that value), there is a 50% probability or chance that the value obtained is greater than or less than the true value of the population. In order to increase the probability of correctly estimating the true value of the population value of interest, confidence limits are used which establish a range within which the true value has a greater chance of being contained. The Confidence Coefficient, the degree of confidence of being correct, should be selected in advance. For example, if from a sample of 50 units the average is determined to be 21.57 we can determine through the use of confidence limits that there is a 95% chance that the true average of the population is between 21.20 and 21.90. Statistical Tolerance Limits differ from confidence limits inasmuch as they define the limits between which a stated percent of the population will fall with a given degree of confidence. For example from a sample of 50 units, if we determine the average to be 21.57 and the standard deviation equal to 1.5, we can then determine the statistical tolerance limit which will give us 95% confidence that 95% of the population (or individual values) will fall between 8.40 and 24.70. One or two-sided confidence or statistical tolerance limits can be constructed depending on our concern of a value exceeding (or being less than) a limit or being between two limits.
having the same but unknown underlying cumulative probability distribution and that the process from which the data is taken be in statistical control. With such a few restrictions imposed one must be willing to sacrifice accuracy. This results in a wider range or band for the statistical tolerance limits. To obtain the same accuracy (the same width of the statistical tolerance limit) as when a known distribution method is used, a larger sample size for nonparametric methods is required.

A. The data must first be arranged in ascending numerical order as in column 2 of the WDAS (Figure 1).

B. Select the confidence coefficient, percent γ, and the reliability or population coefficient, %R or %P, determine if one or two-sided statistical tolerance limits are needed, and enter the appropriate table in the Appendix, Table IV. Tables IVa and IVb are used for a one-sided STL (either an upper or a lower STL). The difference between these two tables is that Table IVa can be used with a sample size of 1 to 1,000 and the value of the limit is the largest or the smallest data point observed; whereas, Table IVb is for sample sizes of 50 to 1,000 and the value of the limit is the nth largest or smallest (n is contained in the body of the table) data point observed. Table IVc is for a two-sided limit and the rank order numbers of the rth smallest and the s th largest of the data points observed which constitute the values of the limits. The paired values r and s in the table can be interchanged. Table IVd is the relationship between confidence coefficients (γ %) and coefficient of population or reliability (R % or P %). The limits are determined by the largest and smallest values observed in the sample.

Example Using Data in Figure 1:

- Sample Size, n = 56, however the smallest sample size tabulated in the table is used, i.e., n = 55 in Table IVb in place of n = 56.
- From Table IVa, n = 56. One-Sided Upper Limit 90% Confidence and 96% Reliability or Product - 90% γ x 96% P < .008.
- From Table IVb, n = 55. One-Sided Limit 75% Confidence and 75% reliability, 12th 75% γ x 75% P < .004
- From Table IVc, n = 55. Two-Sided Limit 75% Confidence and 75% Reliability or Product, r = 6, s = 6 .001 < 75% γ x 75% P < .006
- From Table IVd, n = 50, P = .95 72% γ x 95% P < .008

7.0 PROCEDURE FOR GIVEN PARAMETERS AND STATISTICS.

Determination of the Weibull Plot when the Shape, Parameter, standard deviation or mean and the number of individual values in the sample are known.

A. Determine the value of β by selecting the desired distribution, i.e., β < 1 decreasing failure rate, β = 1 constant failure rate (exponential), β > 1 increasing failure rate, β = 3.25 Normal Probability Distribution, etc. on the Weibull probability grid, β is the negative value of the y-intercept with the principal ordinate.

B. Construct a straight line through the β estimator point (1,0) which has a y intercept with the principal y-axis of β. This straight line represents the slope.

C. From the scales on the extreme right side of the Weibull probability grid determine the values m, a or b.

D. Determine the magnitude of the variable scale, auxiliary x-axis.

E. Using the given value of β or δ, determine y₀, x-intercept with the principal y-axis from the following equations:

- \( y₀ = \frac{A}{\delta} \) when β is known
- \( y₀ = \frac{A}{\delta} \) when δ is known

F. Plot y₀, construct a straight line through y₀ and parallel to the straight line representing the slope (refer to paragraph B above). This straight line is the Weibull Plot of the selected distribution.

G. From this Weibull Plot other statistics can be obtained by using the methods given in paragraph 4.2.

8.0 ACKNOWLEDGMENT

Allis-Chalmers expresses its appreciation to Mr. Arthur Bender, Jr. of General Motors Corporation, Delco-Remy Division, for permission to use the special cumulative method for handling group data.


STATISTICAL CONCEPTS


Computer Methods for Estimating Weibull Parameters in Reliability Studies, John H. K. Kao, Assistant Professor, Department of Industrial and Engineering Administration, Cornell University, Ithica, New York.


"Sampling Plans Based on the Weibull Distribution," A. P. Goode and J. H. K. Kao, Proceedings, 7th National Sym-


APPENDIX

WDAS MATHEMATICAL RELATIONSHIPS

1. Bender's Special Frequency Accumulation

\[ P_i = 100 \left( \frac{\sum_{j=1}^{i-1} f_j + f_i}{2n} \right) \]

Where \( P_i \) is Percentage of values above \( i \)th cell midpoint
\( f_i \) is Frequency in \( i \)th cell
\( f_j \) is Frequency in \( j \)th cell
\( n \) is Sample size

2. Weibull Cumulative Distribution Function (CDF)

\[ F(x) = 1 - \exp \left[ \frac{-(x - \gamma) \beta}{\eta \beta} \right] \quad x \geq \gamma \]

\[ = 0 \quad \text{elsewhere} \]

OR

\[ F(x) = 1 - \exp \left[ \frac{-(x - \gamma) \beta}{\alpha} \right] \quad x > \gamma \]

\[ = 0 \quad \text{elsewhere} \]

Where \( \exp \) denotes the exponent to the base \( e, e = 2.718 \)
\( x \) is the random independent continuous variable
\( \gamma \) is the location parameter
\( \alpha \) is the quasi-scale parameter
\( \beta \) is the shape parameter
\( \eta \) is the true scale parameter and Weibull Central Tendency or Characteristic Life also \( \eta = \alpha / \beta \)
5. Characteristic Life, $\eta$

As $\beta$ changes the mean varies in relation to the mode and the median.

$$
cdf \quad F(x) = 1 - \exp \left[-\left(\frac{x}{\eta}\right)^\beta\right]$$

$x \geq \gamma, \gamma = 0$

when $x = \eta$

$$
F(\eta) = 1 - \exp \left[-\left(\frac{\eta}{\eta}\right)^\beta\right]
= 1 - \exp \left[-(1)^\beta\right] = 1 - \exp \left[-1\right]
= 1 - \frac{1}{e} = 1 - .368 = .632
$$

$F(x)$ is constant for any distribution (any $\beta$) and the % CDF is 63.2%.

Therefore, $\eta$ becomes a useful measure of the central tendency of the Weibull.

6. Reference