7. Weibull Analysis

Introduction

Among all of the distributions available for reliability calculations, the Weibull distribution is the only one unique to the engineering field. Originally proposed in 1937 by Professor Waloddi Weibull (1887-1979), the Weibull distribution is one of the most widely used distributions for **failure data analysis**, which is also known as **life data analysis** because life span measurements of a component or system are analysed.

A Swedish engineer and mathematician studying metallurgical failures, Professor Weibull pointed out that normal distributions require that initial metallurgical strengths be normally distributed, which is not necessarily the case. He noted the need for a function that could embrace a great variety of distributions, including the normal.

When delivering his hallmark American paper in 1951, *A Statistical Distribution Function of Wide Applicability*, Professor Weibull claimed that life data could select the most appropriate distribution from the broad family of Weibull distributions and then fit the parameters to provide reasonably accurate failure analysis. He used seven vastly different problems to prove that the Weibull distribution could easily be applied to a wide range of problems.

The initial reaction to the Weibull distribution was generally that it was too good to be true. However, pioneers in the field of failure data analysis began applying and improving the technique, which resulted in the U.S. Air Force recognising its merit and funding Professor Weibull's research until 1975.

Today, **Weibull analysis** refers to graphically analysing probability plots to find the distribution that best represents a set of life data for a given failure mode. Although the Weibull distribution is the leading method worldwide for examining life data to determine best-fit distributions, other distributions occasionally used for life data analysis include the exponential, lognormal and normal. By "fitting" a statistical distribution to life data, Weibull analysis provides for making predictions about the life of the products in the population. The parameterised distribution for this representative sample is then used to estimate such important life characteristics of the product as reliability, probability of failure at a specific time, mean life for the product and the failure rate.

Advantages of Weibull Analysis

Weibull analysis is extensively used to study mechanical, chemical, electrical, electronic, material and human failures. The primary advantages of Weibull analysis are its ability to:

- Provide moderately accurate failure analysis and failure forecasts with extremely small data samples, making solutions possible at the earliest indications of a problem.
- Provide simple and useful graphical plots for individual failure modes that can be easily interpreted and understood, even when data inadequacies exist.
- Represent a broad range of distribution shapes so that the distribution with the best fit can be selected.
- Provide physics-of-failure clues based on the slope of the Weibull probability plot.

Although the use of the normal or lognormal distribution generally requires at least 20 failures or knowledge from prior experience, Weibull analysis works extremely well when there are as few as 2 or 3 failures, which is critical when the result of a failure involves safety or extreme costs. **WeiBayes**, a distribution in the Weibull family, can even be used with no failures when prior engineering knowledge is sufficient.

Weibull Probability Plots

Weibull analysis studies the relationship between the life span of a component and its reliability by graphing **life data** for an individual failure mode on a Weibull probability plot. Weibull analysis is most often used to describe the time to failure of parts. These can be light bulbs, ball bearings, capacitors, disk drives, printers or even people. Failure modes include cracks, fractures, deformations or fatigue due to corrosion, excessive physical stress, high temperature, infant mortality, wear-out, etc..

When plotting the time-to-failure data on a Weibull probability plot, engineers prefer using **median rank regression** as the parameter estimation method. Median rank regression finds the best-fit straight line by using least squares regression (curve fitting) to minimise the sum of the squared deviation (regressing X on Y). Median rank regression is considered the standard parameter estimation method because it provides the most accurate results on the majority of data sets.

Typically, the horizontal scale (X-axis) measures the component age, and the vertical scale (Y-axis) measures the cumulative percentage of the components that have failed by the failure mode under consideration.

A Weibull probability plot has a linear/nonlinear time-scale along the abscissa and another nonlinear scale for the distribution function along the ordinate. These nonlinear scales are selected in such a way that the model used for data is an appropriate one. If the scales match the data, the graph turns out to be a straight line. Because of their simplicity and usefulness, probability graphs have been used for many years in statistical analysis. However, it must be noted that the probability plotting methods to derive distribution parameters are independently and identically distributed. This is usually the case for non-repairable components and systems but may not be true with failure data from repairable systems.

In Figure 7-1, the Weibull probability plot considers the times to failure for a unique failure mode. When a number of parts are tested under normal operating conditions, they do not all fail at the same time for the same cause. The failure times for any one cause tend to concentrate around some average, with fewer observations existing at both shorter and longer times. Because life data is distributed or spread out like this, they are said to follow a distribution. To describe the shape of a distribution, which tends to depend upon what is being studied, statistical methods are used to determine a formula. If the plotted data points fall near the straight line, the Weibull probability plot is considered reasonable.

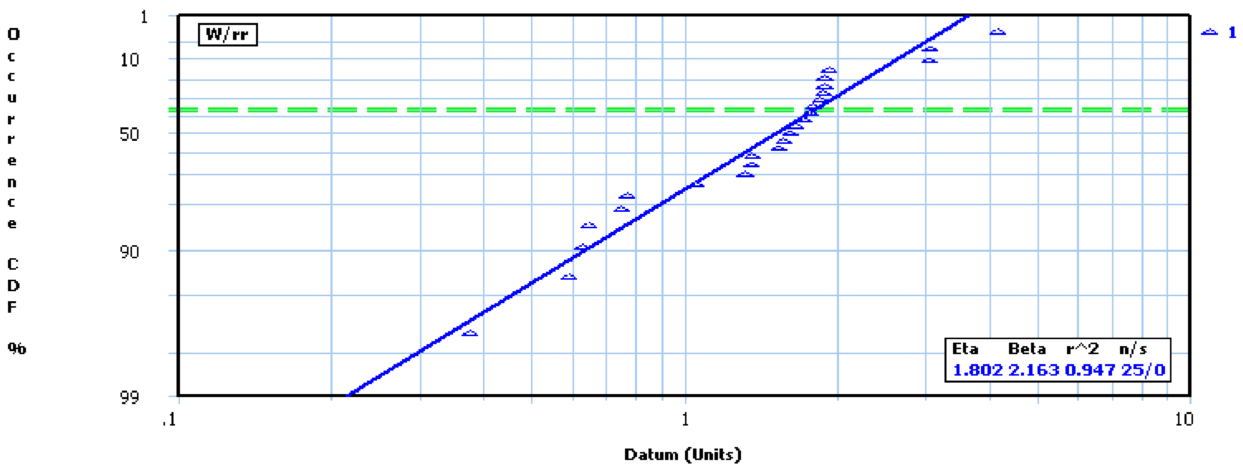

Figure 7-1. Weibull Probability Plot

NOTE Although the Y-axis values are probabilities that go from 1 to 99, the distances between the tick marks on this axis are not uniform. Rather than being based on point changes, the distances between tick marks on both the Y and X axes of the Weibull probability plot are based on percentage changes. Known as a logarithmic scale, the distance from 1 to 2, which is a 100 percent increase, is the same as the distance from 2 to 4, which is another 100 percent increase. A logarithmic scale provides for like-to-like comparisons of several series. In addition to offering more insight into the problem, this visual representation helps to identify the distribution method that best fits a straight line to the data set.

While the previous figure plots occurrences, it is very common to plot the age of components at failure. In these cases:

- The Y-axis is usually $ln\left\{ln\left[\frac{1}{1-F(t)}\right]\right\}$.
- The X-axis is $ln(t)$.
- The Y-axis intercept is $\beta \cdot ln(\eta)$.

Uses for Weibull Analysis

Weibull analysis has traditionally be used for analysing failure data for:

- Development, production and service.
- Quality control and design deficiencies.
- Maintenance planning and replacement strategies.
- Spare parts forecasting.
- Warranty analysis.
- Natural disasters (lightning strikes, storms, high winds, heavy snow, etc.).

New applications of Weibull analysis include medical research, instrument calibration, cost reduction, materials properties and measurement analysis.

Understanding Weibull Analysis

The two-parameter Weibull is by far the most widely used distribution for life data analysis:

$$R(t) = exp\left\{-\left(\frac{t}{\eta}\right)^{\beta}\right\}$$

Where:

$t \geq 0$, $\beta > 0$ and $\eta > 0$. Here, β and η are shape and scale (characteristic life) parameters of the distribution.

Because two-parameter Weibull distribution effectively analyses the life data from burn-in (infant mortality), useful life and wear-out periods, it can be used in increasing, constant and decreasing failure rate situations.

The first parameter defining the Weibull probability plot is the **slope**, beta (β), which is also known as the **shape parameter** because it determines which member of the Weibull family of distributions best fits or describes the data. The second parameter is the **characteristic life**, eta (η),which is also known as the **scale parameter** because it defines where the bulk of the distribution lies. The parameters β and η are estimated from the life data, which are always positive values. After Weibull analysis is completed, the Weibull probability plot visually indicates the slope and the goodness of fit.

NOTE A three-parameter Weibull distribution is also widely used. The third parameter, **location**, is a constant value that is added to or subtracted from the time variable, t. For additional information, refer to page 7-15.

The Weibull hazard function or **failure rate** depends upon the value of β. Because the β value indicates whether newer or older parts are more likely to fail, the Weibull hazard function can represent different parts of the bathtub curve:

- **Infant Mortality.** In electronics and manufacturing, infant mortality refers to a higher probability of failure at the start of the service life. When the β value is less than 1.0, the Weibull probability plot indicates that newer parts are more likely to fail during normal usage, which is known as a **decreasing instantaneous failure rate**. To end infant mortality in electronic and mechanical systems with high failure rates, manufacturers provide production acceptance tests, "burn-in" and environmental stress screenings prior to delivering such systems to customers. Providing that the part survives infant mortality, its failure rate should decrease, and its reliability should increase. In this case, because such parts tend to fail early in life, old parts are considered better than new parts. Overhaul of parts experiencing high infant mortality is generally not appropriate.
- **Random Failures**. Assuming that the Weibull probability plot is based on a single failure mode, a β value of 1.0 indicates that the failure rate is constant or independent of time. This means that of those parts that survive to time *t*, a constant percentage will fail in the next unit of time, which is known as a **constant hazard rate** or **instantaneous failure rate**. This makes the Weibull probability plot identical to the exponential distribution. Because old parts are assumed to be as good as new parts, overhaul is generally not appropriate. The only way to increase reliability for components or systems that experience random failures is by redesigning them.
- **Early Wear-out**. Unexpected failures during the design life are often due to mechanical problems. When the β value is greater than 1.0 but less than 4.0, overhauls or part replacements at low **B-lives** may be cost effective. B-lives indicate the ages at which given percentages of the population are expected to fail. For example, the B-1 life is the age at which 1 percent of the population is expected to fail, and the B-10 life is the age at which 10 percent of the population is expected to fail. Reliability and cost performance for parts experiencing early wear-out may be improved by optimizing the preventative maintenance schedule.
- **Rapid Wear-out**. Although a β value greater than 4.0 within the design life of a part is a major concern, most Weibull probability plots with steep slopes have a safe period within which the probability of failure is negligible, and the onset of failure occurs beyond the design life. The steeper the slope, the smaller variation in the times to failure and the more predictable the results. For parts that have significant failures, overhauls and inspections may be cost effective. Because scheduled maintenance can be costly, it is usually only considered when older

parts are more likely to wear out and fail, which is known as an **increasing instantaneous failure rate**.

Because different slopes imply different failure classes, the Weibull probability plot provides clues about what may be causing the failures. Figure 7-1 lists the failure causes that are most likely for each failure class.

β Value	Class	Description
$\beta < 1.0$	Infant Mortality	When $\beta < 1.0$, failures tend to be due to: • Inadequate burn-in or stress screening. • Quality problems in components. • Quality problems in manufacturing. • Improper installation, setup or use. • Problems in rework/refurbishment.
$\beta = 1.0$	Random Failures	When $\beta = 1.0$, failures tend to be due to: • Human error during maintenance. • Induced rather than inherent failures. • Accidents and natural disasters (foreign objects, lighting strikes, wind damage, etc.).
$\beta > 1.0$ and < 4.0	Early Wear-out	When $\beta > 1.0$ and < 4.0, failures tend to be due to such problems as: • Low cycle fatigue. • Bearing failures. • Corrosion/erosion. • Manufacturing process.
$\beta > 4.0$	Rapid Wear-out	When $\beta > 4.0$, failures tend to be due to rapid wear-out associated with old age or: • Inherent property limitations of materials (such as ceramic being brittle). • Severe problems in manufacturing process. • Minor variability in manufacturing or in material.

Table 7-1. Failure Classes and Likely Causes by Slope Values

Statisticians, mathematicians and engineers have formulated statistical distributions to mathematically model or represent certain behaviours. Compared to other statistical distributions, the Weibull distribution fits a much broader range of life data. The Weibull **probability density function** (pdf) is the mathematical function that describes the fitted curve over the data. The pdf is represented either mathematically or on a plot where the X-axis represents times. Different members of the Weibull family have widely different shaped pdfs. The **cumulative density function** (cdf) is the area under the curve of the pdf. The cdf for the Weibull distribution is given by:

Equation

$$F(t) = 1 - exp\left\{-\left[\frac{t}{n}\right]^{\beta}\right\} \qquad (7.1)$$

Where:

η represents the characteristic life (scale parameter).

β represents the slope (shape parameter).

The cdf gives the probability of failure within time, t. The parameters η and β are estimated from the failure times. If the failure data comes from a Weibull distribution, the values of η and β can be plugged in the cdf formula to find the fraction of parts expected to fail within a certain time.

Characteristic life, η, and the Mean Time To Failure (MTTF) are related. The characteristic life shows the point in the life of the part or system where the failure probability is independent of the parameters of the failure distribution. For all Weibull distributions, η is defined as the age at which 63.2 percent of the units can be expected to have failed.

For $\beta = 1$, MTTF and η are equal. The relationship between MTTF and η is gamma function:

Equation

$$MTTF = \eta \cdot \Gamma\left[1 + \frac{1}{\beta}\right] \qquad (7.2)$$

When $\beta < 1$, MTTF $> \eta$.

When $\beta = 0.5$, MTTF $= 2\eta$.

When $\beta = 1$, MTTF $= \eta$, the exponential distribution.

When $\beta > 1$, MTTF $< \eta$.

Although Professor Weibull originally proposed using the mean or average value to plot MTTF values on the Y-axis of Weibull probability plots, the standard engineering method is now to rank the life data by the median value of the failure times. Table 7-2 displays a Median Ranks table (50%) for a sample size of 10, which was generated using Leonard Johnson's Rank formula.

Because non-symmetrical distributions are so common in life data, median rank values are slightly more accurate than mean values. Once β and η are known, the probability of failure at any time can easily be calculated.

Rank Order	1	2	3	4	5	6	7	8	9	10
1	50.00	29.29	20.63	15.91	12.94	10.91	9.43	8.30	7.41	6.70
2		70.71	50.00	38.57	31.38	26.44	22.85	20.11	17.96	16.23
3			79.37	61.43	50.00	42.14	36.41	32.05	28.62	25.86
4				84.09	68.62	57.86	50.00	44.02	39.31	35.51
5					87.06	73.56	63.59	55.98	50.00	45.17
6						89.09	77.15	67.95	60.69	54.83
7							90.57	79.89	71.38	64.49
8								91.70	82.04	74.14
9									92.59	83.77
10										93.30

Table 7-2. Median Ranks (50%)

Performing Weibull Analysis

In addition to indicating whether newer or older parts are more likely to fail, the Weibull distribution can be applied to a number of different analyses, including reliability and maintenance analysis, probabilistic design, distribution analysis, cost reduction and design comparison. Weibull software, which is any program capable of using the Weibull distribution to calculate the reliability of a component or system in the future based on its past performance, analyses field or laboratory data. Using Weibull software to predict reliability basically consists of six steps:

1. Gather “good” life data.
2. Select the distribution type.
3. Specify the estimation method.
4. Indicate the confidence values.
5. Generate the analysis.
6. Interpret the results.