Extreme Events: Examining the “Tails” of a Distribution
by
Eric Adams and Samiran Ghosh

PR # 10-03

This manuscript and other can be obtained via the World Wide Web from www.math.iupui.edu

August 09, 2010
Extreme Events: Examining the “Tails” of a Distribution

Eric Adams, Manager, Carrier Co, Syracuse NY

And

Samarin Ghosh, Department of Mathematics, IUPUI, Indianapolis, IN

Abstract

Although our engineering training treats all physics as deterministic, we also know that random variation is a normal part of nature. Strength of parts and loads on parts var. Unusually low strengths and unusually high loads do occur, for example a flood or a hurricane in the case of a building or a bridge, or a slug of liquid refrigerant in the case of a compressor. Accidents can occur when extreme events happen. Failure of a part occurs when the load on the part is greater than the strength.

Extreme events happen much more frequently than predicted by theories based on the normal distribution. Statisticians describe extreme value distributions as “heavy tailed” as a result. In this paper, models of extreme values are discussed for both load and strength. Modeling examples are given for loads, strength of materials, applications to predicting time to failure and maintenance intervals. Extreme values are a part of our normal engineering lives.

Introduction

Most engineering problems are not, by their nature, completely deterministic. While deterministic physics may govern simple electrical circuits via Ohm’s law, neither the applied voltage nor the resistance is completely deterministic. Even the most basic electrical circuit, such as a light bulb, is subject to variation. Small differences in material properties and manufacturing affect the level of resistance of the light, even when the circuit is new. As the circuit ages, corrosion, wear and strain further affect the resistance. Material properties and age affect the voltage delivered by a battery powering the circuit. The result is that a nominally deterministic problem has many features of a problem with random variations.

Human factors are another source of seemingly “random” variations. ASHRAE standard 55 (2010), attempts to define the thermal environmental parameters that lead to comfort for human occupants. This
problem is full of variation. First, in the same room environment, all occupants will have different levels of
clothing, and will have different metabolic rates. Second, experiments (Fanger, 1972) have shown that
people in the same environment, with the same clothing, at nominally the same metabolic rate, still do not
respond identically to the question “are you too hot or too cold?” In a building, there are always multiple
spaces (or zones), and each space is not identical, so there is further variation in the comfort of occupants.

To overcome the problem of variation, engineers use “factors of safety” or other constants in expressions
from “experience”. The notion is that the deterministic expressions are used for design, but then an added
“margin” is given to account for the unknown variations in the load and strength of the structure. Since
failure occurs when load is greater than strength, and the levels of load and strength are not truly
deterministic, the question becomes - what is the probability that load is greater than strength?1 In this
question, the mean load and strength are not as important as the extreme values of load and strength.
Unfortunately, classes in basic statistics focus on statistics for the mean and predicting the main effects for
various factors. The distributions learned in basic statistics, such as the Gaussian or Normal distribution,
that are valuable for predicting main effects are not suitable for predicting extreme values (see O'Connor,
2002).

Consider the thermal comfort problem of an entire building with many spaces. For simplicity, we will
ignore the human factors and state that all people will react identically to the environment. Further, we will
ignore radiation effects from the walls and through the windows. The “strength” variable in this example is
temperature (to include radiation, one might use an operative temperature). The temperature in each zone
will be slightly different - and we will model it as random. The use of zoning, personal control, or other
control strategies will certainly effect the standard deviation of the temperature, but will not affect the basic
fact of variation - all sensors and systems will have variation. The “load” variable is the combination of
clothing and metabolic rate of the occupants that determines if they are comfortable. Consider the case of
“too cold”: a person will be too cold if the combination of “clothing and metabolism” is too small for the

---

1 Is this probability acceptable? Although the consequences of failure will not be discussed here, its
importance cannot be overemphasized. The level of analysis and/or the factor of safety used in design must
be much larger for events that endanger life when compared to events that might make us uncomfortable.
given temperature. The statistical problem is to determine how often someone is too cold in the building. It is important to predict the extremes of the distribution: how many people are wearing very light clothing, and how many rooms are much colder than average.

Consider a second problem: is the strength of a beam sufficient to hold a given load when both the beam strength and the load are subject to variation. Consider the charts in Figure 1. In Figure 1a, the load is much less than the strength, or using the thermal comfort problem, all people are dressed so that they will not be too cold (ignore the “too hot” problem). In Figure 1a, the probability of the beam breaking is the small gray shaded area where the two distributions intersect. Here, load is greater than strength, even though the average load is much smaller than strength. In the case of Figure 1a, there is a very, very small probability of failure. In Figure 1b, the strength is not sufficiently larger than the load and some fraction of time the load is larger than the strength and failure will occur.

In Figure 1a and 1b, the naïve assumption of normal distributions was assumed for both load and strength. For simplicity, the standard deviation is assumed unity, but this assumption can be relaxed without any change in the conclusions. The normal distribution has the property that the tail of the distribution is very light - that is a very small fraction of the population lies outside 3 standard deviations from the mean. Further, the normal distribution is symmetric, so the probability of an event a certain distance greater than the mean is equal to the probability of an event the same distance less than the mean.

In Figure 1c, the same mean and standard deviation is assumed for both the load and strength, but for this Figure non-normal distributions are used. These distributions have the property that they are skewed rather than symmetric. The load distribution is assumed to be positively skewed. This occurs physically because many loads cannot be negative while in practice there is often little reason why the maximum load is limited, thus the distribution must be right skewed. The strength distribution shown is left-skewed, indicating that some items have a much lower strength than others. This is often the case because flaws will limit the strength of the item, where a completely unflawed (e.g. single crystal) item will have a maximum possible strength. Further, weakness might occur due to the natural effects of aging, which tend
to create an upper limit for the strength value but no lower limit.

The failure rate of the items in Figure 1c is seen to be larger than in Figure 1b. This is the case even though the mean and standard deviations are identical! The distributions in Figure 1c have the property that they have “heavy” tails compared with the normal distribution, which is a larger fraction of the population lies to the extremes relative to the normal distribution. This is quantified in Figure 1d, where the probability of an event larger than a given value is shown for the distributions in Figure 1b and 1c.

Figure 1: Comparison of load and strength curves. a) load greater than strength, low failure rate, b) load only slightly greater than strength, c) load greater than strength with more realistic distribution functions, d) probability of an extreme event for the cases in Figure 1b and 1c.

Figure 1d can be used to understand the error when using the normal distribution to assess the probability of extreme events. The normal distribution can substantially under predict the probability of extreme events. For example, at the odds of an event occurring 1 in 1000 times ($10^{-3}$ on the vertical scale), the
magnitude of the event is about the mean value plus three standard deviations if the normal distribution is used. If the extreme value distribution is used, the magnitude of the event is predicted to be six standard deviations greater than the mean. Alternatively, examine the probability of an event occurring four standard deviations above the mean. For the normal distribution, the probability is less than $10^{-4}$, whereas for the extreme value distribution, the probability is approximately 1:100. Since extreme events (such as failures) often have very negative consequences (in contrast to the example of the fraction of people being too cold) under prediction of the probability of extreme events can lead to financial or, in the case of safety, human catastrophe.

To understand the origin of extreme value theory, consider the problem of “records” - very low or very high values associated with a distribution. If one randomly draws sample load or strength values from the distribution in Figure 1b, and records only the highest values, it can be shown mathematically that the distribution of these extreme values will have the shape given in Figure 1c (see O’Connor, 2002 or Rausand and Hoyland, 2004). This process is exactly the same as the problem of prediction of failure in engineering. During the course of time, loads appear on the part. These loads, while nominally deterministic, have a random component. The only value of importance when considering the failure of the part is the highest (or “record” load in the given time. If one considers equal increments of time, records the highest load in each increment - it is exactly analogous to the statistical problem, and the distribution of those record values is given by Figure 1c.

**Development**

Given the wide range of possible distributions, the problem of describing the extremes of a distribution may appear hopeless. First, it is a relatively simple problem in order statistics to describe the statistics of the minimum or maximum of a random sample of $n$ numbers from a given distribution, $F(x)$ (see Hoog, McKean, and Craig, 2005 or Rausand and Hoyland, 2004)

\[
T_{(1)} = \min\{ T_1, T_2, T_3, \ldots, T_n \} = U_n \\
T_{(n)} = \max\{ T_1, T_2, T_3, \ldots, T_n \} = V_n \\
F_{U_n}(u) = 1 - (1 - F(u))^n \\
F_{V_n}(u) = F(u)^n
\]
Luckily, a number of researchers (e.g. Cramer, (1946), Gumbel (1958), Pickands (1975)) developed Extreme Value Theory\(^2\) that showed that in limit of large samples (\(n\) is large, exactly what we have in engineering where we have a very large number of load and strength values in each increment of time), and under a wide range of conditions, there are only a few models that are needed for describing the statistics of the largest and/or smallest value of a distribution. For details of applying extreme value theory to reliability modeling, see Rausand and Hoyland (2004) and O’Connor (2002). Type I extreme value distributions are used to describe the minimum and maximum for the right and left tails for exponential types of distributions (this includes most standard distributions such as the normal, lognormal, and exponential). These are usually referred to as the Gumbel distribution of the largest or smallest extreme when referring to the left and right tails respectively. It is these distributions which are shown in Figure 1c. This distribution is not limited in magnitude, meaning the smallest extreme can be negative. A Gumbel distribution of the largest extreme would be used to model loads in many situations. The case where the variable may have an upper limit can also be treated by more advanced techniques (see Einmahl and Smeets, 2009). The probability density functions for the extreme value distributions of type I are:

\[
\begin{align*}
\text{For the maximum values:} \\
f(x) &= \frac{1}{\sigma} \exp\left\{-\frac{x - \tilde{\mu}}{\tilde{\sigma}} - \exp\left[-\frac{x - \tilde{\mu}}{\tilde{\sigma}}\right]\right\} \\
\text{For the minimum values:} \\
f(x) &= \frac{1}{\sigma} \exp\left\{\frac{x - \tilde{\mu}}{\tilde{\sigma}} - \exp\left[\frac{x - \tilde{\mu}}{\tilde{\sigma}}\right]\right\}
\end{align*}
\]

\(\tilde{\mu}\) is the mode, \(\tilde{\sigma}\) is the scale parameter. It can be shown that the mode and scale are related to the mean, \(\mu\), and standard deviation, \(\sigma\), of the distribution as:

\[
\begin{align*}
\mu &= \tilde{\mu} + 0.577\tilde{\sigma} \quad \text{For the maximum values} \\
\mu &= \tilde{\mu} - 0.577\tilde{\sigma} \quad \text{For the minimum values} \\
\sigma &= 1.283\tilde{\sigma} \quad \text{For both maximum and minimum values}
\end{align*}
\]

If the log of the load or strength is extreme value distributed, a type II extreme value distribution is used.

\(^2\) Two classes of Extreme Value distributions exist. This paper will cover only the first class, using generalized extreme value distributions. The second class of problems uses generalized Pareto distributions to handle exceedance over threshold problems.
Type III extreme value distributions are the limiting case for the minimum of a set of values which are bounded on the left. The Weibull distribution describes such cases, which are typical of strength of materials (strength must be positive and one is interested in the smallest strength of a collection of bonds making up a material). Another typical example of a Type III extreme value distribution is the smallest time to failure for assemblies made of many possible failure modes. Thus, the Weibull distribution arises naturally in engineering analysis of extreme values. The Weibull distribution (CDF) is:

\[
F(t) = 1 - \exp \left[ - \left( \frac{t}{\eta} \right)^\beta \right]
\]  

(3)

The two constants in the distribution have engineering meaning. \( \eta \), is the time for 62% of the population to fail - a time of little practical interest in reliability because the engineers would all be fired long before that threshold. However it gives a method of comparing failure rates as engineering improvements to the product will increase \( \eta \). Some product, for example bearings, use a different failure percentage, such as 10% to create a B(10) life. The other Weibull parameter, \( \beta \), is the “slope” which describes the rate of failure. All slopes less than one have a decreasing failure rate with time. Slopes greater than one have an increasing failure rate with time. Unity slopes are used in a model with a constant failure rate, where the Weibull distribution reduces to the exponential distribution. Constant failure rate models (Poisson Process) are the simplest form of reliability model.

The Weibull model assumes the failed part or system is removed from service at the time of the first failure. Most powerfully, following techniques developed for survival analysis (Kline and Moeschberger, 2002), Weibull analysis allows estimation of the rate of failure when only a few parts have failed and can account for all parts removed from service or which have not reached the end of their service life. These techniques first saw widespread application in the aerospace industry where extremely small data sets (<5 failures, and occasionally only a single failure) are used due to the expense and risk of further testing or accidents (Abernethy, 2000). Techniques have been developed to apply Weibull analysis using graphical techniques (e.g. Abernethy, 2000 and O’Connor (2002)) or least squares analysis (Abernethy, 2000) as well as Maximum Likelihood techniques (Klein and Moeschberger, 2003). Specialty paper is available for
plotting. The vertical scale is the fraction of product that has failed; the horizontal is the time to failure. Some manipulation of equation (3) shows that a linear plot can be found to fit the model when the vertical scale on the paper is log(-log) of the fraction failed and the horizontal scale is log time. When using least squares, it is common practice in Weibull analysis to fit “backwards”, that is it is assumed that the fraction failed are known exactly, where the time to failure is subject to error. As a result, the least squares fit are made on the (log) time. In this case the error is defined as the time difference between the measured time and the time estimated by the model. Maximum Likelihood techniques use the probability distribution itself to determine the values of the parameters, \( \eta \) and \( \beta \), which have the highest probability based on the data. Historically, plotting was used. Today, statistical software can calculate the parameters either by maximum likelihood or least squares fitting.

Weibull analysis strictly applies to a single failure mechanism on a single part. The time scale to be used is that which is significant for the mechanism (cycles, hours of operation, etc.). In practice, warranty data usually contains only calendar data, or in the best situation, run hours. Almost never is information available about the load history which would correlate correctly to the failure mechanism. To understand the effect this has on Weibull analysis in practice, it is useful to think of the units operation. Units are scattered in the field with random run rates, based on climate, building occupancy, and system design. As the number of randomization factors increase, probability of stressful events occurring become almost Poisson distributed. When this occurs, the Weibull slope is almost always nearly unity. The analyst concludes the failures are occurring in the useful life period. The same phenomenon occurs when different mechanisms are grouped together for analysis, for example all compressor failures rather than only compressor bearing failures. At the sub-system and system level, there is more randomization in terms of the part that fails and the system that fails.

Reliability engineers divide the typical life of an engineering system in three regimes: first the “burn in period” or “infant mortality“. Early in the products life, the failure rate may be relatively high due to the effects of “burn in”, misapplication, or manufacturing defect. During this period, the failure rate is decreasing with time. Thus, plotting results from this period on Weibull paper will show a slope < 1. The
regime is the period of useful life. Once “weak” systems are removed, the failure rate is usually
approximately constant over the useful life of the product. In this time period, failure rates are constant.
Mechanisms which cause failure in this period are extreme loading events. It is only during this period that
one can describe the reliability of the system in terms of a Mean Time between Failures (MTBF) or similar
parameter. Finally, there is the “wear-out period”. Near the end of a product’s useful life, environmental
effects such as corrosion, wear, fatigue, and creep lead to a new set of failure mechanisms. These
mechanisms lead to an increasing failure rate with time. Plotting test results on Weibull paper will lead to
slopes greater than unity, typical 2-4.

Reliability engineers attempt to find and test products for the end-of-life mechanisms. The testing is
“accelerated” to simulate the mechanisms postulated with far increased stresses so that results can be
obtained in weeks and months, instead of years. The correct design of such a test is a considerable art, as
both the design and failure mechanism must be well understood so that only the correct stress is accelerated
rather than an unrealistic condition of no value. It should be noted that the Weibull distribution simply
allows the use of these acceleration factors (see Kline, J.P. and Moeschberger, M.L., 2003). Similarly, one
can postulate covariates based on measured field performance and analytically determine the acceleration
caused by the covariate.

It is often assumed that a product which passes its life testing will have a low warranty rate. Note that the
mechanism tested for in life testing are end-of-life mechanisms, which are fundamentally different than the
extreme loading events, and manufacturing defects that lead to early life failures in warranty. One can
simulate these failures and design for them, but it is NOT traditional life testing.

In selecting HVAC systems for large facilities, many request for quotes call for failure statistics to be
supplied by OEMs in the form called for by reliability theory, such as the mean time between failures,
MTBF. Manufacturers typically supply MTBF for the useful life period (constant failure rate period) when
the theory applies. In this region, MTBF is very large, and failure rates are very small. The data is not
representative of either the start-up period or end-of-life.
To operational personnel, it is of some importance to determine whether the failure rate of a part is increasing or decreasing with time, or if the product remains in the period of useful life. In this case, the rate of occurrence of failures (ROCOF) and the statistics of the rate is used. There are several methods of practical value. First, the power law model (Crow, 1974) can be used:

$$ROCOF = w(t) = \lambda \cdot \beta \cdot t^{\beta-1}$$

(4)

where $\lambda$ and $\beta$ are parameters. Maximum likelihood methods can be used to fit data of failure times for repairable systems and $\beta$ can be used to determine if the system has a constant, increasing or decreasing failure rate. Alternatively, there are test statistics which can be used to look for an increasing and/or decreasing trend in ROCOF data. Popular statistics given in many software packages are the Laplace test (see Rausand and Hoyland, 2004) and the Military Handbook test (MIL-HDBK-189). For the case of observing a system over a time, $t_0$, the Military Handbook test statistic is:

$$Z = 2 \sum_{i=1}^{n} \ln \frac{t_0}{S_i}$$

(5)

The asymptotic distribution can be shown to be (see Rausand and Hoyland, 2004) $\chi^2$ with $2n$ degrees of freedom. The hypothesis of no trend can be reject for small (increasing failure rate) and large (decreasing failure rate) systems.

**Examples**

The first problem examined is the estimation of the minimum strength of an engineering element using the extreme value distribution of the minimum. Consider the problem of the strength of a wire of a given size (O’Connor, 2002) where a number of samples were cut and tested to failure. A plot of the data using an assumed extreme value distribution is shown in Figure 1a. A good fit to the extreme value distribution is indicated if the data fit a straight line. Using least squares, the line shown is the extreme value distribution of the smallest extreme fit to the data. The fit can be used to estimate the parameters, but the plot itself provides sufficient data for most analysis. While the average strength of the wires was greater than 60 N (~6 lb), extreme value theory predicts that one wire in one hundred would break at 3 N (~3 lb), $\frac{1}{2}$ the
average strength. This is a typical problem in predicting the minimum strength of a part, or the lowest temperature in a building.

Second, consider the problem of modeling large loads. Gumbel’s distribution of the largest extreme (equation 2) would be a natural choice. Figure 2b shows the analysis of Red River flood data for the last century taken at Fargo, ND each spring. Each data point represents the highest level of the Red River during the snow melt period each spring. Only the largest 20 points are shown on the plot. It must be assumed that the data points represent a homogeneous population. Thus, floods due to summer rains must be separated from floods due to melting snow. Further, it is assumed that the basic geological characteristics and climate characteristics are unchanged. At the top of the plot is the “return period”. This return period is the 50% probability point for the occurrence of a flood of a given magnitude. The plot predicts that every 150 years a flood of greater than 30,000 cubic feet per second (850 m$^3$/s). This is equivalent to a cumulative probability of 0.9933.

![Figure 2: Examples of the use of extreme value theory to model a) minimum strengths of a selection of wires (O'Connor, 2002), and b) maximum flood flow at Fargo, ND (USGS, 2010)](image)

Examples can be found applying this theory to such events as: the fastest time humans can run the 100 yard dash (Einmahl, J.H.J. and Smeets, G.W.R. , 2009), insurance losses, equity risks, day to day market fluctuations, size of waves, mutational events in evolution, fires, pipeline and other engineering system failures. Consider the warning of Chavez-Demoulin and Roehrl (2004) “If you look at fat tails, consider
using EVT (Extreme Value Theory), as EVT is too expensive to ignore.”

Next consider the extreme value distributions type III, the Weibull distribution, to describe the time to failure of a fan, a component typical HVAC components (data from Nelson, 1982). In this case, there were 69 fans considered. During the period of observation, 12 failed - a relatively high failure rate which would be considered unacceptable by most customers today. Service times and failure times for all fans were recorded and analyzed following methods outlined in Abernethy (2000) using commercially available software for Weibull analysis (see Figure 3). The parameter estimates and confidence intervals (95%) generated by the software is given below:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Error</th>
<th>95.0% Normal CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape, ( \beta )</td>
<td>1.05602</td>
<td>0.267365</td>
<td>0.642927 - 1.73452</td>
</tr>
<tr>
<td>Scale, ( \eta )</td>
<td>26038.9</td>
<td>12068.7</td>
<td>10497.9 - 64586.6</td>
</tr>
</tbody>
</table>

The shape factor is near unity, indicating a constant failure rate. The confidence interval includes unity, so there is no proof that a constant failure rate model is not valid. The scale parameter is approximately 26000 hours, or about 3 years of service, when it would be expected that 62% of the population would have failed. Note that this estimate is made when only 1/3 that many fans have actually failed. This is the usefulness of the Weibull estimates - they are estimates of the current failure rate and allow forward looking estimates of failure and associated costs, assuming conditions in the field remain fixed (duty cycles, environment, etc.), all failures are in the constant failure period, and that no end-of-life failure modes appear. Typical data used for Weibull analysis from residential and commercial HVAC manufacturers would have similar number of failed components (5-25), but the population would be many times larger, and the corresponding failure rates many times smaller. Many manufacturers have automated systems that generate hundreds / thousands of such plots in order to analyze the details of designs and look for potential problems well before they become serious field issues.
Next, consider the case of simple springs (data from Cox and Oakes, 1984). In this case, 60 springs were tested by reliability engineers with 6 groups of ten, with each group having a different level of stress, and the number of cycles to failure was recorded. 53 units were tested to failure. 7 units completed the test without failing. Weibull analysis of the data cannot be performed in traditional fashion because the increased stress has accelerated the failure times and the results from the different tests are not directly comparable. It would be possible to compute each of the 6 groups separately and compare the failure rate (Weibull scale factor) among the groups. Alternatively, the Weibull model can be shown to be a proportional hazards model as well as an accelerated life model (See Klein and Moeschberger, 2003). The result is that one can incorporate a linear model for covariates in log time that increase or decrease the time constant, $\eta$, while leaving the slope unchanged. In this case, all 60 data points can be used to generate the regression, greatly increasing the statistical power of the conclusions (roughly, confidence intervals scale on the inverse square root of the number of samples). In this case, using 60 points vs. 10 would lead to confidence intervals smaller by a factory of 2.5). The regression results are shown below:
The model is given as:

$$\ln(t) = ax + b$$

Where $x$ is the level of stress (non-dimensional here) and “$a$” is the predictor coefficient labeled “stress” in the above table, and $b$ is the predictor coefficient labeled “intercept” in the table. This model can be used to estimate the value of $\eta$ for any stress level. The rate of failure at the stress level is then modeled by the Weibull distribution with shape factor, $\beta$, given as 1.62 above, and the appropriate level of $\eta$ for the stress level. Note that this level is significantly greater than one, indicating that a wear out mode for the spring has been activated. Finally, the table above gives statistical evidence that the null hypothesis that spring life is not affected by stress level can be rejected at any meaningful level of confidence (p~0). The author finds that it is better to create a model based on physical principles and then use statistics to demonstrate the validity of the model.

Next, analysis of a complete system was undertaken. Accelerated life test data from Nelson and Hahn (1972) on the failure of motorettes at 4 different temperatures. 40 units were followed. 17 failed. Figure 4 shows a Weibull analysis performed at one of the temperatures. The confidence intervals on the plot are very wide and the parameter estimation is very poor due to only 10 samples being used. Again, regression can be used to build a model of the life parameter. Here, I choose to model directly with temperature rather naively, rather than attempt to use a more realistic temperature acceleration factor. The regression table from the analysis is given below.
Log-Likelihood = -147.365

Temperature is seen as a significant predictor of life. The shape factor is significantly larger than unity indicating that a wear out mode has been found and that the failure rate is increasing with time at all temperature levels. The effect of increasing temperature is to lower the life (coefficient of temperature is negative). The form of the model here can be criticized on several levels, but the aim was to have a linearized model to be used over a limited range. Tableman, and Kim, (2004) also analyzed this data set, but used a model with inverse temperature. Their model showed a somewhat better fit based on log likelihood (or AIC and/or BIC criteria for model building). An inverse temperature model would be more plausible on physical grounds.

![Motorette Failure Data @ 190F](image)

Figure 4: Weibull analysis of motorette failure data at one temperature (see Nelson and Hahn, 1972)

The previous examples were from the point of view of the manufacturer that wishes to predict product life. Similar analysis can occur from the point of view of facilities maintenance and repair. Here it is desired to predict the repair frequency of systems in use. This might be used to predict the need for spares or to understand (and minimize) process down time. The example used is the frequency of repair of AC units. Time between failure data was given by Proschan (1963) in days between failures. 14 time intervals are
given occurring over several years. The view can be of a single AC unit, or a batch of several units, as long as the batch size does not change during the observations. Repair time is not known, and it is only assumed that the repair time is small compared to any of the failure times. A typical desire is to predict the frequency of repair (the extreme event: AC unit failure) and to determine if the system failure rate is increasing, decreasing or constant. A power law model of ROCOF was assumed. The data were fit to the model and the estimation of the parameters is given below. The best estimate of the shape factor is given as 1.21, but note that the 95% confidence interval includes unity. Further, both the MIL-HDBK-189 test and Laplace’s test both fail to reject the hypothesis of a constant failure rate (p values are very high). Since the null hypothesis of constant failure rate cannot be rejected, it makes most sense to use the constant failure rate model to predict the ROCOF. This was done and the MTBF was determined to be 124 days, with a 95% confidence interval of 73 days to 209 days. This rather broad confidence interval is due to the small sample size of the data. Figure 5 presents the cumulative failure plot as a function of time for the data. The linear model is seen to fit the data adequately, in agreement with the Poisson model used to determine the MTBF. The slope of the line in Figure 5 represents the constant ROCOF and is inversely proportional to MTBF.

Model: Power-Law Process
Estimation Method: Maximum Likelihood
Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Standard Estimate</th>
<th>Standard Error</th>
<th>95% Normal CI Lower</th>
<th>95% Normal CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape</td>
<td>1.21747</td>
<td>0.315</td>
<td>0.733539</td>
<td>2.02066</td>
</tr>
<tr>
<td>Scale</td>
<td>198.220</td>
<td>120.674</td>
<td>60.1099</td>
<td>653.653</td>
</tr>
</tbody>
</table>

Trend Tests

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>MIL-Hdbk-189 Test Statistic</th>
<th>Laplace's Test Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-Value</td>
<td>0.734</td>
<td>0.474</td>
</tr>
<tr>
<td>DF</td>
<td>26</td>
<td></td>
</tr>
</tbody>
</table>
Summary

The importance of considering extreme values and the statistics of extreme values has been presented based on simple engineering examples. The extreme value distribution and the Weibull distribution (extreme value distribution type III) are commonly used to model extreme events associated with the failure of engineered systems. Extremely large values of loads and extremely small values of strength are both important when analyzing the design of engineered systems, which can fail due to either cause. Typically, after an initial period of decreasing failure rate due to material defects (smallest extreme value distribution), failures typically occur due to extreme load events (largest extreme value distribution) over the useful life of the system. Finally, at the end of product life, extremely weakened components fail due to environmental effects of time. Weibull analysis can model the time-to-failure events for parts and systems in all three regimes. Deterministic design cannot adequately treat these events, but the language of probability and statistics can be used to estimate the frequency of failure and to estimate the frequency of occurrence of failures in the future, subject to the assumptions that the environment does not change. Furthermore, since Weibull models are both accelerated failure models as well as proportional hazards models, regression modeling can be used to predict the effect of environmental covariates on the life of engineering systems.
The methods outlined here are routinely used by manufacturers to estimate product reliability as well as get early warning of field issues. The methods are also used by operation and maintenance personnel to predict HVAC equipment life, part stocking levels and to anticipate the end of life of products as they wear out.

Several examples of estimation were considered for different systems.

References

Chavez-Demoulin, V. and Roehrl, A. 2004, “Extreme Value Theory can save your neck”, ETHZ publication