Statistical Classification of Major Reliability Event Days in Distribution Systems

Richard D. Christie (M)
University of Washington
christie@ee.washington.edu
August 8, 2001

I. Abstract

This white paper proposes defining a major event day in distribution reliability in terms of average frequency of occurrence. This definition is easy to understand for non-technical people like regulators, seems to be fair to apply to systems of any size, and can be translated into a major event threshold that can be used to classify individual days. An example of this classification is given and practical issues such as the type of probability distribution to use, how to handle zero-outage days, and calculating normal annual reliability with major event days removed are discussed and resolved.

II. Major Event Day Classification

The problem of equitably classifying operating days into normal days and major event days on the basis of distribution reliability is one that is becoming more important as regulators increase scrutiny of, and impose limits on, operating reliability. A method is needed that is based on solid theoretical ground, is equitable when applied to distribution utilities with different sizes and operating territories (and hence different reliability levels), is robust, and is practical for practicing utility engineers to implement. This section starts with some definitions, discusses current methods of classifying days, and their strengths and weaknesses, and then proposes the use of frequency of occurrence for day classification.

A. Major Event Definition

During any given day some portion of the customers on a given distribution system will be out of service for some period of time. Utilities collect this data, and analyze it on a yearly basis. This discussion assumes that outage times for outages (events) that occur (start) on a given day accumulate to that day's reliability statistics. A few days a year have much, much worse reliability than the rest. Major storms, hurricanes, ice storms, earthquakes and other natural disasters are often to blame. Utilities and regulatory agencies want a robust way to distinguish these exceptionally bad days, here called *major event days*, from normal days.

Trial Use Guide P1366 [1] defines "Major Events". This paper uses the term "major event day" to refer to a day on which a P1366 major event occurred. The P1366 definition of a major event is:

> Designates a catastrophic event which exceeds reasonable design or operational
> limits of the electric power system and during which at least 10% of the

customers within an operating area experience a sustained interruption during a 24 hour period.

B. Critique of Existing Definitions

The quantitative criterion in the P1366 definition, outage of 10% of the customers, is unambiguous to apply. In contrast, the definition of "catastrophic event" is very much a matter of opinion, as is the term "reasonable." However, percentage of customers out of service has not proved to be an equitable criterion when applied to different utilities, because it results in significant differences in major event day counts for systems of different size and different average reliability (due to type of service territory) subject to essentially the same weather stresses.

Even without the "reasonable," quantitatively identifying situations that exceed the design and operational limits of the power system is difficult. There is no single clear quantitative measure of stress on the power system, and no one generic "stress" design parameter, so straightforward application of this definition to classify major event days is not feasible.

It is natural to think of a major event day as one in which a catastrophic natural event occurs, such as a storm or earthquake (hence the use of the term "storm reliability" in some discussions). However, a wide variety of natural events can cause major reliability events. Widely different sets of natural events are agreed between utilities and regulators. An all-inclusive natural event-related definition seems unattainable. Furthermore, these severe natural events occur in a continuum of intensity. When is a storm a major storm, and when not?

Finally, the existing definition does not embrace the new reality of widespread (rolling) blackouts due to market shortages, which have been seen in California this year and were at least considered possible in two other states.

C. Proposed New Definition of Major Events

Returning to the concept of a major event as one where the power system is subject to stress that exceeds system design or operational limits, it is common to discuss power system design practices in terms of being able to withstand "once in N year" events like wind storms or ice. If the system is designed to withstand a one in ten year ice storm, then N is 10 for icing, for example, and arguably major events due to icing should occur, on average, less than once every 10 years. Thus the frequency of major event days can be directly related to high level power system design criteria.

Of course, there are a large number of different types of low-frequency events that can occur, and not all of them are as well known or well understood as icing. The combined effects make the practical frequency of major event days on the order of several per year. The exact number can be agreed on between utilities and their regulators, and applied uniformly to all utilities subject to those regulators.

The concept of frequency of major event days (e.g. three a year, six a year, etc.) offers significant advantages compared to other definitions. It is easy for non-technical people like regulators and the general public to comprehend, and provides a common ground for discussion between the utility and interested parties. Costs of more robust design criteria can be brought into the discussion with relative ease.

The same frequency criteria applied to utilities of different sizes should produce fairer results in terms of major event day classification than a size criteria, even a relative size criteria (such as 10% of the customers out of service). The same frequency criteria translates to different event size criteria for different utilities, as discussed below, yet by definition should yield the same average number of major event days per year.

Once a specific frequency criteria has been agreed upon for major event days, a size threshold for easily classifying specific days as major event days or normal days can be computed by application of probability theory. The computation of the threshold can be done with common spreadsheet programs.

For these reasons, this paper advocates use of a frequency of occurrence criterion for classifying days as major event days or normal operating days.

III. Computing Classification Threshold From Frequency Criteria

A. A Measure of Daily Reliability

While it is simple to say that some days a system has worse reliability than others, the choice of which reliability measure to use is not trivial. The existing major event definition uses the extent of interruptions, but there is a general sense that the duration of the interruptions during the event is a better measure of reliability. The standard reliability index to measure duration index is System Average Interruption Duration Index (SAIDI) [1], which can be computed for each day of the year in minutes or hours. SAIDI in minutes/day is a slightly easier set of values to work with than SAIDI in hours/day.

Other discussions have used customer-minutes interrupted per day (CMI/day). For any given day SAIDI/day and CMI/day differ only by a constant composed of the customer count and possibly conversion from minutes to hours, so the proposed statistical classification method will give the same result for any data set with the same customer count in each day. However, when using data sets where the customer count varies, for example in combining outage data from different years, using CMI/day would cause small errors in the classification. For this reason, SAIDI/day is preferred as the measure of reliability r_i for day i.

B. Basic Method

The basic method to compute major event day reliability threshold R^* from frequency of occurrence criteria requires converting the frequency of occurrence f to a probability of

occurrence p(major), finding the reliability measure's probability distribution from historical data r_i, and finally finding the reliability threshold $R*$ which has the desired probability of occurrence p under the probability density function to the right of the threshold. Then days where impact exceeds the threshold are major event days.

C. Converting Frequency of Occurrence to Probability of Occurrence

The probability that any random day in one year will be a major event day is found by dividing the frequency of occurrence f by the total number of days n. For example:

> The agreed-on frequency of major event days is $f = 3$ per year.
> The probability that a random day will be a major event day is

$$p(\text{major}) = \frac{f}{n} = \frac{3}{365} = 0.00822 = 0.822\% \tag{1}$$

(Purists may want to use 366 days for leap years.)

D. Characterizing the Probability Distribution

Daily reliability data r_i can be used to determine the probability distribution of the data. First the type of probability distribution (normal, log-normal, etc.) must be found. Type determines the formulas that define the probability density function (pdf), $f(x)$, the cumulative probability distribution $F(x)$, and its inverse. (Here x is a generic random variable.) The formulas in turn determine the parameters needed to characterize the distribution, such as mean, standard deviation and, in some cases, one or more shape parameters.

The first step in characterizing any probability distribution is to prepare a histogram from the historical data and look at its shape. A symmetrical bell-shaped curve would suggest using a normal distribution while a skewed curve would suggest using one of many alternatives such as log-normal or Weibull.

The next step is to prepare a probability plot for each of the candidate probability distributions. If n values of historical reliability data r_i are arranged in increasing order, then the expected probability of occurrence p of a value greater than a given sample in the ith ordered position, r_i, is just

$$p(r_i) = \frac{i - 0.5}{n} \tag{2}$$

The value of 0.5 accounts for the discrete nature of the sampled data.

For any probability distribution, such as the normal distribution with a density function shown in Figure 3, this probability p is the area under the curve to the left of a value of the random variable, $x*$,

$$p(x < x*) = \int_{-\infty}^{x*} f(x)dx \tag{3}$$

where $f(x)$ is the probability density function of the distribution and the integral can be solved or computed.

If the historical data has the probability distribution given by $f(x)$, then the ordered sample r_i should have a linear relationship with the corresponding value of $x*$. (The slope and intercept of the line accounts for differences in mean and standard deviation between the historical data and the nominal distribution.) It is common practice to graph the relationship and inspect it visually for linearity. Statistics packages often provide probability plot functions for a variety of distributions. See an advanced probability textbook, for example, [2] or the web textbook at [3], for more discussion of probability plots.

Probability distributions that fit the historical data better will have more linear probability plots. Probability distributions with shape parameters, such as the Weibull, must be fitted to the data before a probability plot can be drawn. The type of probability distribution to use is usually decided by comparing probability plots and choosing the most linear.

There is some reason to believe that the log-normal distribution is a good default choice for the probability distribution, because daily reliability has a low limit, and variability greater than its average value.

Statistical analysis software can make generating histograms and probability plots easy, but they can also be done, at least for simple distributions like normal and log-normal, with common spreadsheet programs.

E. Finding the Major Event Day Reliability Threshold

Once the probability distribution is identified and characterized, the problem is simply to find the value of a daily reliability threshold $R*$ that leaves the desired probability of a major event day, $p(\text{major})$, under the right hand tail of the probability density function $f(x)$. For some probability distributions $R*$ can be computed directly from p with either a formula or with built-in spreadsheet functions. In others, a guess at $R*$ can be used to compute probability p, and an iterative process will be needed.

F. Classifying Days

Once threshold $R*$ is known, then a day i is a major event day if its reliability value $r_i > R*$.

IV. Detailed Example

Data for this example is taken from the daily reliability data identified as Utility 2, obtained in a spreadsheet from the Distribution Working Group web site [4]. The example focuses on classifying days in year 2000.

A. Data Inspection

The first step is to examine the data. The data for 2000 is in CMI/day. A customer count is given, so the data is easily converted to SAIDI/day, measured in minutes, by dividing by customer count. The customer count is the same for all three years.

Only 361 data points r_i are provided for 2000. This appears to be because four days have no reported outages, and in fact the dates of the omitted days can be found by inspecting the data from [4]. Days with zero outages should be included in the data set. Omitting them will result in a higher threshold for major event days. Four days with SAIDI/day = 0 were added to the data set, making the total number of days $n = 365$.

B. Probability of Occurrence

The discussion in the data spreadsheet [4] heuristically identifies 9 major event days in three years of data, so $f = 3$ major event days/year is taken as the agreed-on standard for major event day frequency. Then

$$p(\text{major}) = \frac{f}{n} = \frac{3}{365} = 0.00822 = 0.822\% \tag{4}$$

C. Histogram

Reference [2] recommends using $\sqrt{n}$ bins in a histogram, where n is the number of samples. The histogram in Figure 1 shows that the distribution has a sharp peak in the first bin (351 of 365 data points) and a long tail to the right. The distribution is clearly not normal, or there would be a correspondingly long tail to the left.

Fig. 1 - Utility 2 2000 Histogram

Rev 2 8/8/01
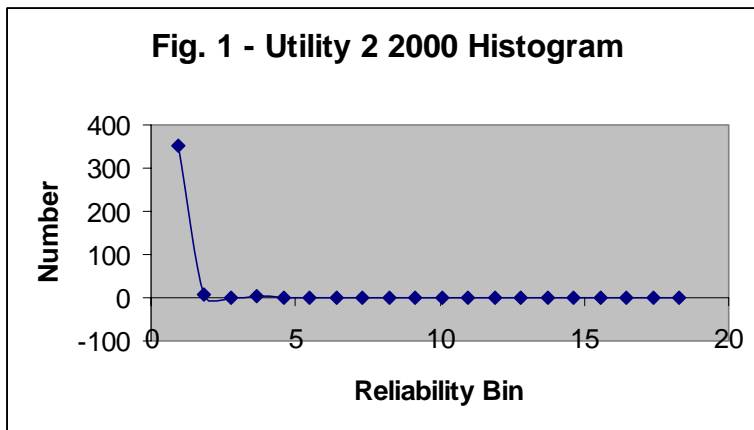
As a matter of interest, an expanded histogram using 20 bins to sort the 351 smallest reliability values was also plotted and is shown in Figure 2. This continues to have most of the data points in the first bin, and shows the start of the long right hand tail in more detail. This histogram emphasizes the non-normal nature of the distribution of the data.

Fig. 2 - Expanded Histogram
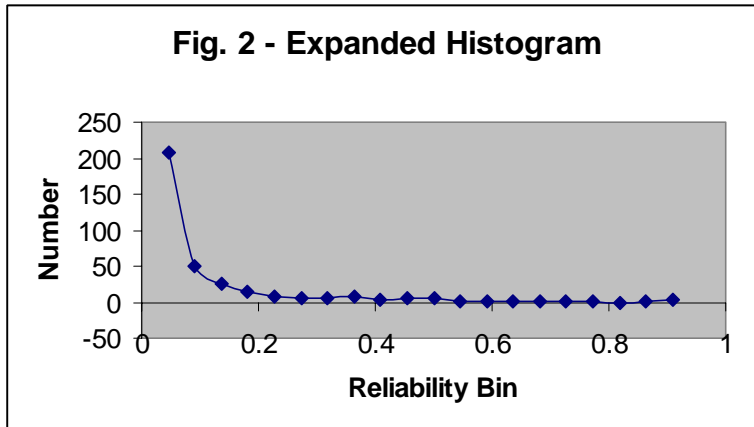
For comparison, a normal reliability distribution is shown in Figure 3. A "typical" log-normal distribution is shown in Figure 4, and a log-normal distribution with the approximate parameters of the historical data is shown in Figure 5.

Fig. 3 - Normal Distrbution

Fig. 4 - Typical Log-Normal Distribution

Fig. 5 - Log-Normal Distribution
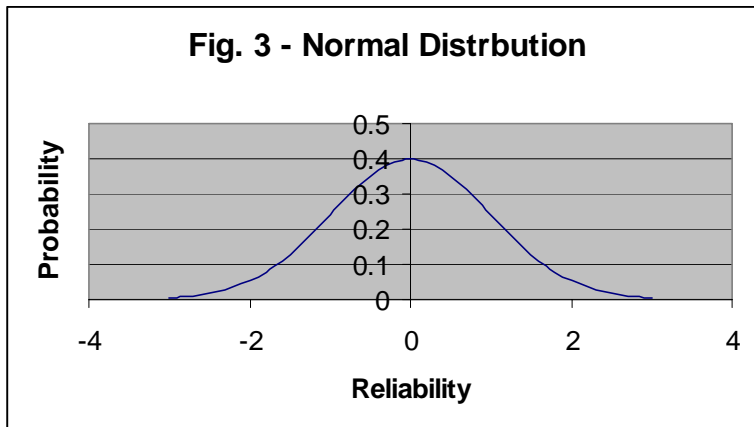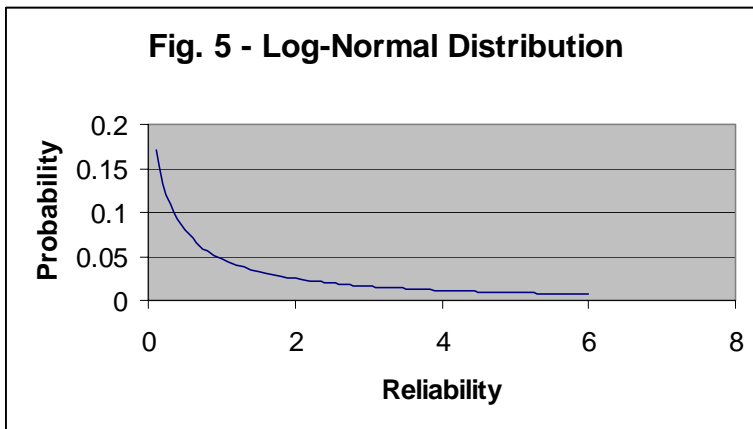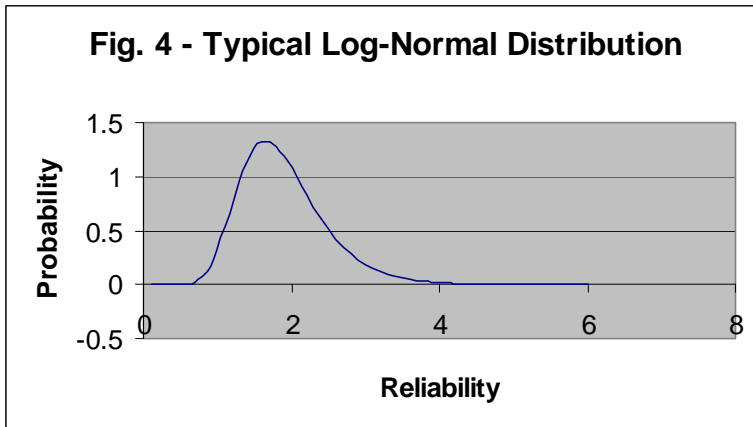
It seems clear that the data distribution is not normally distributed. Determining whether it is log-normally distributed is best done with a probability plot.

D. Probability Plots

Figure 6 shows the probability plot for the example data and the normal distribution. The values on the axes are not important, just the shape of the curve. The relationship is clearly non-linear, and is also clearly not linear with a few outliers. This confirms that the historical data does not have a normal distribution.
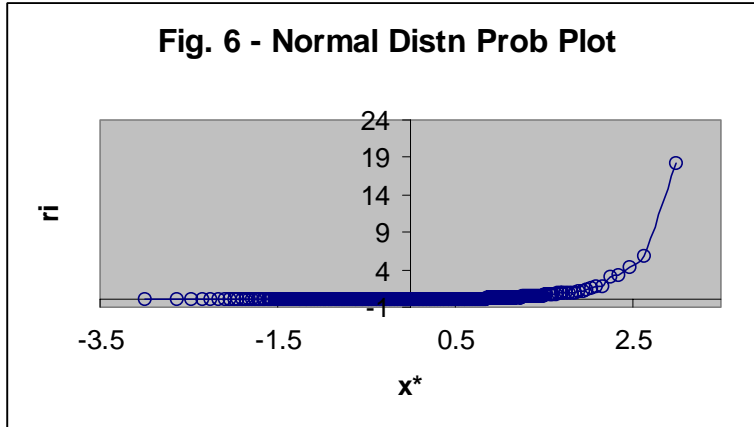
Fig. 6 - Normal Distn Prob Plot

Figure 7 shows a probability plot for the log-normal distribution. In a log-normal distribution the natural log of the random variable is normally distributed, so to create this plot, the natural log of the historical data, $\ln(r_i)$ is graphed against the normal distribution x^*. Of course $\ln(0)$ is undefined, so the days with 0 SAIDI/day are omitted from the plot.

Fig. 7 - Log-Norm Prob Plot
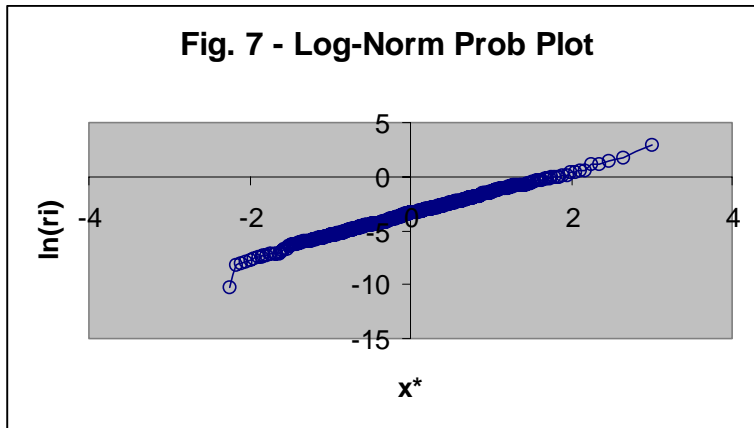
This plot is clearly linear with the exception of one low value of r_i. This low value may be an outlier, produced by a process different than most of the daily values, but low value outliers are not of much interest. Much more interesting is that the highest value in the data set, which is over three times higher than the next highest value, falls comfortably near the straight line, and there are no obvious outliers at the high end of the data.

At this point a complete analysis would prepare probability plots for a number of other asymmetrical probability distributions, such as Weibull, Gamma, Erlang, Beta, etc. Many of these distributions require fitting to the historical data before a probability plot can be drawn.

This paper will stop with the log-normal distribution and adopt it as the distribution of the daily reliability values in the historical data. The reasons for this are:

a. Most importantly, the probability plot shows a very good fit to the log-normal distribution. Other distributions would provide only minor improvements in fit, if any.

b. The log-normal distribution is computationally easy to deal with. Its parameters are easily computed from historical data. Probabilities corresponding to random variable values, *p(x*),* and random variable values corresponding to probabilities, *x*(p)*, are easily computed using the normal distribution functions. From a practical point of view, this computational tractability seems worth any minor increase in accuracy of fit.

c. The final reason is that there are weak theoretical reasons to believe that the distribution of daily reliability has a log-normal distribution. Basically, random processes what would be normally distributed, but have a lower limit, are log-normally distributed. Log-normal distributions become more like normal distributions as the mean becomes large compared to the standard deviation. The process behind distribution reliability values is too complex to make any definitive theoretical statement about its normality, which is why this part of the argument for a log-normal distribution is weak.

This paper's assumption of the log-normal distribution is not a statement that all distribution daily reliability values are log-normally distributed. As just discussed, a theoretically supported statement of this type is not possible. From a practical point of view, the author has investigated data from only one utility. However, all three years of that data is reasonably log-normal, as can be seen from the Figure 8, the probability plot of all three years of Utility 2 data against the log-normal distribution.

This suggests that the hypothesis that all daily reliability data is log-normal is at least worth investigation.

Fig. 8 - Log-Norm Prob Plot, 3 Yrs
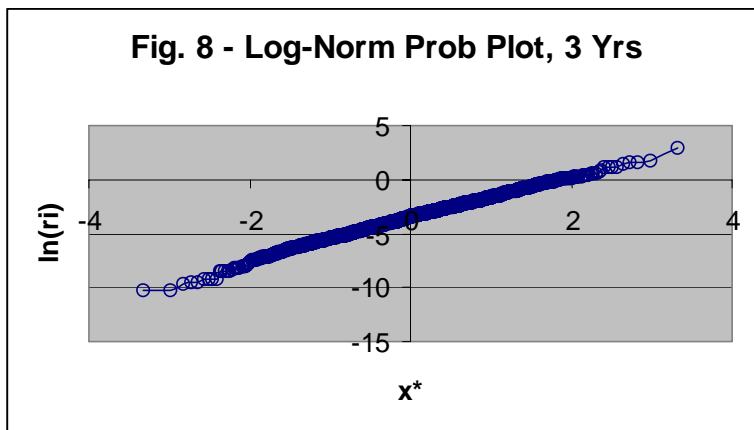
E. Calculating Parameters of the Log-Normal Distribution

The normal distribution is characterized by its mean and standard deviation. If a set of samples is drawn from a normal distribution, the mean and standard deviation of the samples are estimates of the mean and standard deviation of the actual distribution.

For the log-normal distribution, a similar set of parameters can be calculated from the natural logs of the samples. α is the mean of the natural log of the historical data, and β is the standard deviation of the natural log.

$$\alpha = \frac{1}{n}\sum_{i=1}^{n} \ln(r_i) \tag{5}$$

$$\beta = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\alpha - \ln(r_i))^2} \tag{6}$$

For the historical data of Utility 2 for 2000, these values are $\alpha = -3.41$, $\beta = 1.96$. As with the probability plot, the four days with SAIDI/day of zero are omitted from the calculation.

F. Computing the Threshold

The threshold is the value of R* such that the probability under the right hand tail of the log-normal distribution with $\alpha = -3.41$ and $\beta = 1.96$ is

$$p(\text{major}) = p(r > R*) = 1 - p(r \le R*) = 0.822\% \tag{7}$$

One way to find this is to solve the integral of the log-normal distribution's probability density function for *R**.

$$1 - p(r \le R*) = 0.822 = 1 - \int_0^{R*} f(x)dx = 1 - \int_0^{R*}\frac{1}{\sqrt{2\pi}\beta x} e^{-\frac{(\ln x - \alpha)^2}{2\beta^2}} dx \tag{8}$$

This can be more easily evaluated by using the solution for the normal distribution.

$$p(\text{major}) = 1 - p(r \le R*) = 0.822\% = 1 - \Phi\left(\frac{\ln R* - \alpha}{\beta}\right) \tag{9}$$

hence

$$R* = e^{\beta\Phi^{-1}(1 - p(\text{major})) + \alpha} \tag{10}$$

Here Φ is the cumulative probability function for the normal distribution. Φ^{-1} is its inverse, which is a common spreadsheet function. For the Utility 2 year 2000 data, this gives a threshold value of R* = 3.644, which classifies the three highest days of the 2000 data as major event days.

However, this calculation has ignored the four days with $r_i = 0$. These can't be included in the log-normal distribution because the value of ln(0) is undefined. These days must be treated as a lumped probability.

If there were no zero values, the area of 1.0 under the log-normal probability density function ($f(x)$ in equation (8)) would represent 100% of the probability. The zero values represent probability that is lumped at zero, and is not under the log-normal curve. Four zeros in 365 days gives

$$p(0) = \frac{4}{365} = 1.01\% \tag{11}$$

Thus the area of 1.0 under the log-normal curve represents 100% - 1.01% = 98.9% of the probability. The correct area under the curve above the threshold, $\hat{p}$, must be the desired probability p increased by the ratio of area to non-zero probability,

$$\hat{p} = \frac{1}{1 - p(0)} p = \frac{1}{0.989} 0.00822 = 0.00831 \tag{12}$$

The threshold found from equation (10) using the corrected value $\hat{p}$ in place of p(major) is R* = 3.616. This does not affect the number of major event days in 2000. The effect on the threshold value would increase as the number of days with $r_i = 0$ increased.

Threshold values were computed using data for 1998 and 1999, and then using a combination of all three years of data. Results are presented in Table 1.

Table 1 - Major event days for different data sets

Data	R*	# Major Event Days, 1 yr of data	# Major Event Days, 3 yrs of data
1998	3.281	2	2
1999	2.616	2	1
2000	3.616	3	5
3 years	3.148		8

The results give a sense of the variability in the classification process. The analysis with three years of data gives a better estimate of the probability distribution, so these results will be more accurate. The variation in the number of major event days based on the data used for the analysis is to be expected. Therefore, when applying a uniform evaluation process, the number of years of data used to determine R* should be specified. See section VI for more discussion on this point.

Table 2 gives the number of major event days in each year using the three years of data, for different average annual frequencies of occurrence of major event days. As the desired frequency increases, the classification process generates a smaller than expected number of major event days.

Table 2 - Major event day results for different frequencies

Freq f MED/yr	p	$\hat{p}$	R*	1998 days	1999 days	2000 days	Total days
3	0.00822	0.00831	3.148	2	1	5	8
4	0.01096	0.01104	2.552	2	2	5	9
5	0.01370	0.01380	2.157	3	2	5	10
6	0.01644	0.01656	1.873	3	3	5	11

V. Computing Normal Reliability

Utilities and regulators are interested in computing reliability statistics for normal days, with major days omitted. Annual SAIDI can be computed by adding up SAIDI/day. When doing so, the annual SAIDI value should be corrected to account for the number of major event days that are omitted. In theory, a major event day that is omitted would have had average normal reliability if the major event had not happened.

For example, using the 2000 data with 5 major event days determined from three years of data, the sum of the normal days' SAIDI is 45.8 minutes. Ratioing up by 365/360 gives a corrected annual value of 46.5. Failing to correct for omitted major event days understates normal annual reliability. Table 3 gives annual reliability with and without major event days for the three years of data.

Table 3 - Annual SAIDI values, minutes/year.

Year	Total SAIDI	Uncorrected Normal SAIDI	Major Event Days	Corrected Normal SAIDI
1998	51.7	43.5	2	43.7
1999	49.7	44.6	1	44.7
2000	80.8	45.8	5	46.5

VI. A Simpler Classification Method

The simplest method of classifying major event days based on an N event per year frequency of occurrence criterion would be to classify the N least reliable days in each year as major event days. This approach is clearly too simple - it does not account for the variation in number of major event days that is expected to occur from year to year.

A better, but still quite simple method is to combine multiple years of data before picking off the worst days. For M years of data, the M·N worst days would be classified as major event days. For the example data, Table 4 gives the major event days (MEDs) for different frequencies. For the nominal three days/year value, the classification differs by one day from the statistical method, but as the frequency increases, the statistical method classified fewer days as major events. The classification threshold (R*) using this method is somewhere between the SAIDI/day of the highest non-major event day and the lowest

major event day. This range can be compared to the statistically derived value of R* for a given frequency. The difference is probably due to random variation in the data, and illustrates the range of variability to be expected in major event day classification.

Table 4 - Major event day classification using the simple method

Freq f MED/yr	1998 days	1999 days	2000 days	R* high	R* low	R* stat
3	2	2	5	3.00	2.19	3.148
4	3	3	6	1.81	1.73	2.552
5	4	4	7	1.60	1.56	2.157
6	4	5	9	1.42	1.25	1.837

The remaining question is then what is the best value of M, the number of years of data to use. Practical limitations may at first limit M to the amount of data on hand. To take an obvious case, the example data used in this paper has only three years of data available.

From a statistical point of view, the higher the value of M, the better. A result from order statistics, presented in [2], gives the probability that the kth largest value in m samples will be exceeded r times in n future samples. It is

$$p_{r|m,k,n} = \frac{k}{n+k-r} \frac{\binom{m}{k}\binom{n}{r}}{\binom{n+m}{n+k-r}}$$

(13)

(Interestingly, this result is independent of the probability distribution of the reliability.)

For example, the three years of Utility 2 data comprise $m = 1095$ samples. The largest non-major event day is the $k = 1095 - 9 = 1086^{th}$ sample. The probability of $r = 3$ days in the next year of n = 365 samples exceeding the size of the largest non-major event day is found from equation (13) to be 0.194 (19.4%). If 10 years of data were used, the corresponding probability of having three days exceeding the largest non-major event day in the historical data would be 0.214. As the number of samples m increases, the probability increases asymptotically to a limit of about 0.22. See Table 5. From the table it is clear that most of the change takes place in the first five years.

Table 5 - Probability of r major event days in the next year based on years of historical data

Years of data	p r = 3	p r = 5	p r = 10
1	0.157	0.124	0.089
2	0.183	0.144	0.103
3	0.194	0.153	0.110
4	0.201	0.158	0.113
5	0.205	0.161	0.116
10	0.214	0.168	0.121
15	0.218	0.171	overflow

A second consideration is that the process producing the daily reliability statistics also changes. In particular, if a utility significantly improves reliability, especially with respect to major events, major event days classified by either the simple or the statistical method will cluster in the early years of the historical data. If reliability declines, major event days will cluster in the later years. This consideration motivates use of as little historical data as possible to classify major event days. Three to five years of data is probably a reasonable compromise.

VII. Conclusion

This paper has presented two statistical methods - one simple, one more detailed - of classifying operating days as major event days or normal days based on average frequency of occurrence, and argued that this is a better classification method than those currently in use. The paper has provided detailed examples of how to perform the classification. In the process, the probability distribution of daily reliability for the available sample data was shown to be approximately log-normal, and it was argued that the computational convenience of the log-normal distribution is significant enough to justify forgoing any minor improvements in fit that may result from using more complex probability distributions. Finally, discussion of the amount of data to use in the analysis and computation of normal reliability with major event days removed was provided.

One note of caution is that the statistical method presented here relies on evaluating small probability values in the extended tails of probability distributions. In the statistical method, the tail is determined by the fit of the entire distribution. A good fit for the distribution may not be all that good for the tail. The major reasons to believe that the fit to the right hand tail is good are (a) the good match to linearity in the probability plots of the sample data of the high historical reliability days and (b) the intuitively satisfactory classification results obtained using the method. Both of these depend on the nature of the historical data. Data from a larger number and a wide variety of utilities should be examined before drawing final conclusions about the proposed classification method.

VIII. References

[1] IEEE PES Working Group on System Design, IEEE P1366 Trial Use Guide for Electric Power Distribution Reliability Indices, IEEE, New York, 2000. (Draft 4, 1/31/01).

[2] N. T. Kottegoda and R. Rosso, *Statistics, Probability, and Reliability for Civil and Environmental Engineers*, McGraw-Hill, New York, 1997.

[3] http://www.itl.nist.gov/div898/handbook/, NIST/SEMATECH Engineering Statistics Internet Handbook, 2001.

[4] http://grouper.ieee.org/groups/td/dist/sd/utility2.xls, downloaded July, 2001.

VII. Author

Richard D. Christie (M) is an Associate Professor of Electrical Engineering at the University of Washington in Seattle. He obtained the BS and MS in Electric Power Engineering from Rensselaer Polytechnic Institute and the PhD from Carnegie Mellon University. He has been active in developing predictive distribution reliability assessment methods, and is a member of the Working Group on System Design of the Distribution Subcommittee of the IEEE. His email is christie@ee.washington.edu.

VIII. Spreadsheet

A spreadsheet with the example data, illustrating the specific computations of the proposed classification method, can be found on the web at
www.ee.washington.edu/people/faculty/christie

IX. Acknowledgments