A Point about Reliability Data

Gain the maximum amount of information from reliability data when only partial information is available.

By Jennifer Brown

Q When performing a reliability assessment, what happens if the exact value of a data point is unknown?

A Reliability is defined as "The probability that a component, device, product, process, or system will function or fulfill a function after a specified duration of time or usage under specified conditions" (from the standard guide for general reliability, E3159). Reliability may also be used more generally as a measure of quality over time or over a usage or demand sequence. In industry, reliability assessments may be needed for the determination of material capability, inspection capability, life cycle, and maintenance requirements, among other applications. Reliability data are often collected either through an experiment or in the field.

Many traditional statistical analysis methods assume that the response variable is an exact value in all cases. In reliability applications, however, the exact value of the response may not be known. For example, when gathering failure-time data in an experimental setting, the experiment might have been stopped before all the units have failed. Just because the unit didn't fail doesn't mean that failure-time information cannot be gathered from that unit. On the contrary, it is known that the failure-time of the unit that "survived" is some time after the point at which the experiment was stopped.

If a statistical method that requires an exact value for each response is used to analyze the data, then the units that survived would have to be either excluded from the analysis or treated as a failed unit, with failure time equal to the time at which the experiment was stopped. However, doing so may lead to seriously misleading results. Fortunately, advanced statistical methods have been developed to deal with this type of data, referred to as censored data.

This article deals primarily with the identification and classification of censored data using examples in material property development, probability of detection, and life-testing/assessment. Only one statistical method is discussed to illustrate how censored data are handled in the analysis.

In general, there are three types of censored data: right-censored, left-censored, and interval-censored. Let y_i represent the response for the i^{th} unit. A right-censored data point is one in which there is a lower bound y_{iL} for the i^{th} response. That is, the exact response value is somewhere in the interval (y_{iL}, ∞). A left-censored data point is one in which there is an upper bound y_{iU} for the i^{th} response. That is, the exact response value is somewhere in the interval $(-\infty, y_{iU}]$. An interval-censored data point is one in which y_i is between some lower bound y_{iL} and some upper bound y_{iU}. That is, the exact response value is somewhere in the interval $(y_{iL}, y_{iU}]$. Censoring can be random or predetermined [1].

Right-censored data are often encountered in material property curve development. When developing a fatigue curve that relates stress and cycles-to-failure for a particular material, for example, data are collected through an experiment. Often the experiment is designed so that each specimen in the experiment is randomly assigned to a given stress level under which it will be subjected up to a maximum number of cycles. In other words, the experiment is designed to stop when either the specimen fails or when the maximum cycles is reached. Note that the censoring in this case is predetermined by the established test conditions.

Suppose that the maximum amount of cycles a single test specimen is exposed to at a given stress level is 100,000. Suppose one specimen that was subjected to the lowest stress level reached the 100,000-cycle test limit without failing. If the test was allowed to continue, the specimen would probably have failed at some point. The information gathered from this specimen that "survived" is that its failure-time is sometime after (or to the right of) 100,000 cycles for the given stress level. Thus, the failure time is recorded as (100,000, ∞).

Left-censored data are common in experiments designed to demonstrate the detection capability of an eddy current inspection system. EC inspection is commonly used to detect surface or near-surface anomalies, such as cracks, in metallic hardware. As the probe scans the surface of the hardware it generates a magnetic field. Anything placed in the magnetic field that has electric and/or magnetic properties (such as a surface or near-surface crack) will change the field and produce an EC signal. The measured EC signal is expected to increase as defect size increases.

Understanding the relationship between crack size and EC signal response is the basis for determining the detection capability (often referred to as probability of detection) of an EC system. Left-censored data occur when the inspection system cannot distinguish the signal generated by a small flaw from inherent noise in the EC system and/or material. In this case the censoring is predetermined by inherent noise. For example, suppose that the noise threshold is known to be 1 division. That is, any signal below 1 division is indistinguishable from noise. If the measured signal from a small flaw falls below 1, the response for that flaw is recorded as (0,1]. In other words, the exact measured signal is some amplitude to the left of 1, or within the noise.

Interval-censored data are often present in data resulting from field inspection of engines at given time points. Suppose that inspection data from 12 engines in the field have been gathered to analyze the time-to-crack initiation for a particular component. In this example, "failure" is defined as a crack initiating at a particular critical location, and the failure mode is assumed to be driven by the number of starts of the engine. When engines are being inspected in the field to gather data on time-to-crack initiation, indication of cracking may not be present at a given inspection point but then is present at the next inspection point. That is, the actual time the initiation occurred is between the two inspection points.

Suppose there are 12 engines in the field and each engine is inspected after six starts, with a life limit of 24 starts. Two engines that were inspected at 18 starts had indications of cracking. However, no indications were found after 12 starts. Thus the time the crack initiated is said to be somewhere between 12 and 18 starts. Thus the time-to-crack initiation for both engines is recorded to be in the interval (12, 18].

A particular data set may contain one, two, or all three types of censored data. The engine inspection data serve as one such example that contains all three types of censoring. In fact, in this example, all the data are censored since inspections are being done at fixed points, equally spaced out in time. Hence, exact information over time will never be known.

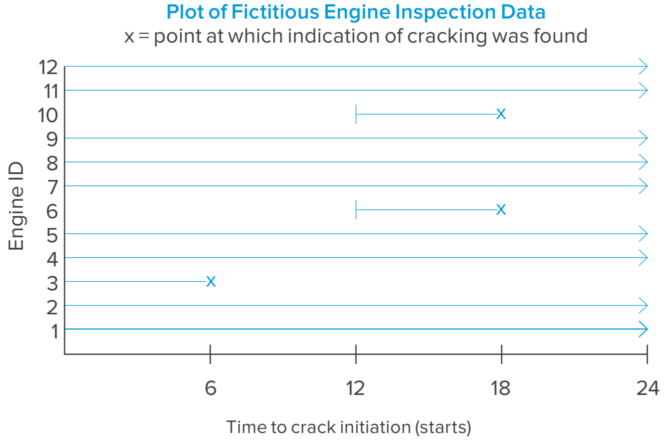
An example of two engines with interval-censored data has already been illustrated for the engine inspection data. Suppose one engine had indications of cracking during the first inspection at six starts. All that is known is that the crack initiated sometime before (or to the left of) six starts. Thus, the time the crack initiated is recorded as (0, 6]. Suppose the remaining nine engines made it to the 24 start life limit without any indication of cracking. Thus, the time a crack might initiate is after (or to the right of) 24. Thus, the response for the remaining engines is recorded as (24, ∞). Table 1 shows the fictional engine inspection data. Figure 1 shows a plot of the right-censored, left-censored, and interval-censored data.

(CONTINUED ON PAGE 44)

Table 1 — Engine Inspection Data

Ficticious Engine Inspection Data	
Time to crack initiation (starts)	Number of Engines
(0,6]	1
(12,18]	2
(24, ∞)	9

Figure 1 — Plot of Engine Inspection Data



"Failure to treat a censored data correctly can have a significant impact on a regression model, resulting in a poor model and misleading predictions."

For the material property curve and EC POD examples, regression analysis is typically used to analyze the data. Traditional regression methods use the method of least squares, which assumes all response values are known exactly, to estimate regression model parameters. If traditional regression methods are used, then the censored data in the material property curve and EC POD examples would have to either be excluded from the analysis entirely or assigned some value (for example, the censored data points in the material property curve example would be assigned the maximum number of cycles allowed in the experiment). Treating censored data in this way may lead to seriously misleading results. Because censored data are commonly encountered in these applications, the method of maximum likelihood is used to estimate the regression model parameters.

The method of maximum likelihood is a more advanced regression technique for estimating the model parameters and is required for the proper treatment of censored data. In general, the method of maximum likelihood is viewed as a more versatile method for fitting a model to data since it can be applied to a wide variety of data types, including censored data, as well as to a wide variety of statistical models [1]. It should be noted that the values for the model parameters using the method of maximum likelihood are the same as those obtained using the method of least squares when no censored data are present [2].

Failure to treat censored data correctly can have a significant impact on a regression model, resulting in a poor model and misleading predictions. (For more information about POD analysis for eddy current when censored data are present, refer to the practice for probability of detection analysis for â versus a data, E3023, and to Reference 3.)

In summary, if censored data are present, it is critical to properly identify and classify the censored data points to gain the maximum about of information out of the data and to avoid erroneous and misleading analysis results. Selecting the appropriate statistical analysis method that can properly handle censored data is also key.

Jennifer Brown worked as a statistician in the aerospace industry for over 12 years. An ASTM International member since 2010, she is a member of the committee on quality and statistics (E11), chair of its subcommittee on terminology (E11.70), and a member of its subcommittee on specialized for NDT methods (E07.10).

John Carson, Ph.D., a senior statistician for Neptune and Co., is the Data Points column coordinator. He is chairman of the subcommittee on statistical quality control (E11.30), part of the committee on quality and statistics (E11), and a member of the committee on petroleum products, liquid fuels, and lubricants (D02), air quality (D22), cannabis (D37), and environmental assessment, risk management, and corrective action (E50).

REFERENCES

1. Meeker, W. Q., and Escobar, L. A., *Statistical Methods for Reliability Data*, John Wiley and Sons, New York, 1998.
2. Neter, John, Kutner, M.H., Wasserman, William, and Nachtsheim, C. J., *Applied Linear Statistical Models*, 4th Ed., The McGraw-Hill Companies Inc., 1996.
3. Herberich Brown, J., "Probability of Detection Analysis for Eddy Current Inspection Systems," ASNT Fall Conference and Quality Testing Show, Columbus, Ohio, 2009.