WEIBULL.COM
RELIABILITY ENGINEERING RESOURCES

Reliability HotWire
The eMagazine for the Reliability Professional

Reliability HotWire
Issue 7, September 2001

Reliability Basics

The Reliability Function

The most frequently used function in life data analysis and reliability engineering is the reliability function. This function gives the probability of an item operating for a certain amount of time without failure. As such, the reliability function is a function of time, in that every reliability value has an associated time value. In other words, one must specify a time value with the desired reliability value, *i.e.* 95% reliability at 100 hours. This degree of flexibility makes the reliability function a much better reliability specification than the MTTF, which represents only one point along the entire reliability function. (For more information on the limitations of the MTTF as a reliability specifications, see http://www.reliasoft.com/newsletter/2Q2000/mttf.htm.)

In this article, we will take a look at the reliability function, how it is derived, and an elementary statistical background.

Types of Random Variables
In general, most problems in reliability engineering deal with quantitative measures, such as the time-to-failure of a component or whether the component fails or does not fail. There are two types of random variables that can be used in the analysis of this type of data.

In judging a component to be defective or non-defective, only two outcomes are possible. We can then denote X as representative of these possible outcomes (*i.e.* defective or non-defective). In this case, X is a random variable that can take on only two discreet values (let's say defective = 0 and non-defective = 1), the variable is said to be a *discrete random variable*.

In the case of times-to-failure data, our random variable X can take on the time-to-failure of the product or component and can be in a range from 0 to infinity (since we do not know the exact time *apriori*). The product can be found failed at any time after time 0 (*e.g.* at 12.4 hours or at 100.12 hours and so forth), thus X can take on any value in this range. In this case, our random variable X is said to be a *continuous random variable*.

In this article, we will deal almost exclusively with continuous random variables.

The Probability Density and Cumulative Density Functions
From probability and statistics, given a continuous random variable X, we denote:

- The probability density *pdf*, as $f(x)$.
- The cumulative density *cdf,* as $F(x)$.
- The *pdf* and *cdf* give a complete description of the probability distribution of a random variable.

If X is a continuous random variable, then the probability density function, *pdf*, of X is a function $f(x)$ such that for two numbers, a and b with $a <= b$:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

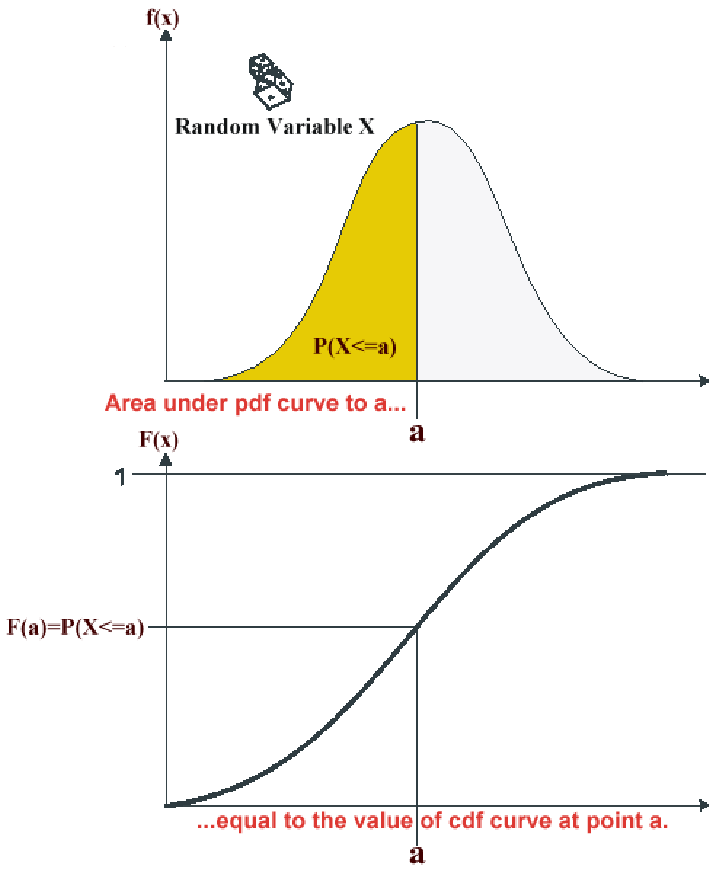
That is, the probability that X takes on a value in the interval $[a,b]$ is the area under the density function from a to b.

The cumulative distribution function, *cdf*, is a function $F(x)$ of a random variable X, and is defined for a number x by:

$$F(x) = P(X \leq x) = \int_{0,-\infty}^{x} f(s)ds$$

That is, for a given value x, $F(x)$ is the probability that the observed value of X will be at most x. Note that depending on the function denoted by $f(x)$, or more specifically the distribution denoted by $f(x)$, the limits will vary depending on the region over which the distribution is defined. For example, for all the distributions considered in this reference, this range would be $[0,+\infty]$, $[-\infty,+\infty]$ or $[\gamma,+\infty]$. In the case of $[\gamma,+\infty]$ we use the constant γ to denote an arbitrary non-zero point or location.

Following is a graphical representation of the relationship between the *pdf* and *cdf*.



The mathematical relationship between the *pdf* and *cdf* is given by:

$$F(x) = \int_{-\infty}^{x} f(s)ds$$

where *s* is a dummy integration variable. Conversely:

$$f(x) = -\frac{d(F(x))}{dx}$$

In plain English, the value of the *cdf* at *x* is the area under the probability density function up to *x*, if so chosen. The total area under the *pdf* is always equal to 1, or mathematically,

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Lifetime Distributions
A statistical distribution is fully described by its *pdf* (or probability density function). The functions most commonly used in reliability engineering and life data analysis, namely the reliability function, failure rate function, mean time function and median life function, can be determined directly from the *pdf* definition, or *f(t)*. Different distributions exist, such as the normal, exponential etc., and each one of them has a predefined *f(t)*. These distributions were formulated by statisticians, mathematicians and/or engineers to mathematically model or represent certain behavior. For example, the Weibull distribution was formulated by Walloddi Weibull and thus it bears his name. Some distributions tend to better represent life data and are most commonly referred to as *lifetime distributions*.

The *pdf* of the well-known normal, or Gaussian, distribution is given by:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}$$

In this definition, note that *t* is our random variable which represents time and the Greek letters μ (mu) and σ (sigma) represent what are commonly referred to as the *parameters* of the distribution. Depending on the values of μ and σ, *f(t)* will take on different shapes. The normal distribution is a two-parameter distribution, with two parameters μ and σ.

For any distribution, the parameter or parameters of the distribution are estimated from the data. For example, in the case of the normal distribution, μ, the mean, and σ, the standard deviation, are its parameters. Both of these parameters are estimated from the data, *i.e.* the mean and standard deviation of the data. Once these parameters are estimated, the *pdf* function *f(t)* is fully defined and we can obtain any value for *f(t)* given any value of *t*.

Given the mathematical representation of a distribution, we can also derive all of the functions needed for life data analysis, such as the reliability function. Once again, this will only depend on the value of *t* after the value of the distribution parameter or parameters are estimated from data. (We will discuss methods of parameter estimation in subsequent HotWire articles.)

The Reliability Function

The reliability function can be derived using the previous definition of the cumulative density function. Note that the probability of an event happening by time *t* (based on a continuous distribution given by *f(x)*, or *f(t)* since our random variable of interest in life data analysis is time, or *t*) is given by:

$$F(t) = \int_{0,\gamma}^{t} f(s)ds$$

One could also equate this event to the probability of a unit failing by time *t,* since the event of interest in life data analysis is the failure of an item.

From this fact, the most commonly used function in reliability engineering can then be obtained, the reliability function, which enables the determination of the probability of success of a unit, in undertaking a mission of a prescribed duration.

To mathematically show this, we first define the unreliability function, *Q(t)*, which is the probability of failure, or the probability that our time-to-failure is in the region of 0 (or γ) and *t*. So, from the previous equation, we have:

$$F(t) = Q(t) = \int_{0,\gamma}^{t} f(s)ds$$

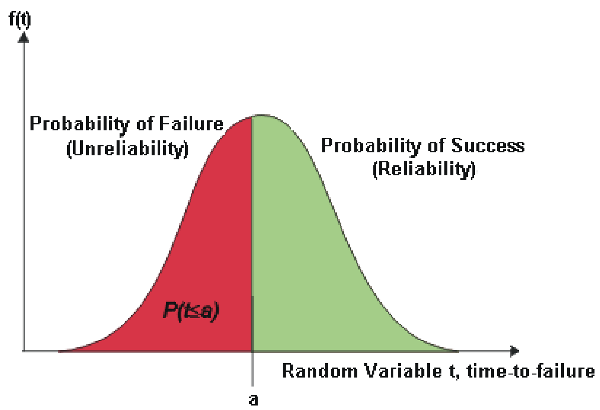
In this situation, there are only two situations that can occur: success or failure. These two states are also mutually exclusive. Since reliability and unreliability are the probabilities of these two mutually exclusive states, the sum of these probabilities is always equal to unity. So then:

$$\begin{aligned} Q(t) + R(t) &= 1 \\ R(t) &= 1 - Q(t) \\ R(t) &= 1 - \int_{0,\gamma}^{t} f(s)ds \\ R(t) &= \int_{t}^{\infty} f(s)ds \end{aligned}$$

Where *R(t)* is the reliability function. Conversely, the *pdf* can be defined in terms of the reliability function as:

$$f(t) = \frac{d(R(t))}{dt}$$

The following figure illustrates the relationship between the reliability function and the *cdf*, or the unreliability function.



We will illustrate the reliability function derivation process with the exponential distribution. The *pdf* of the exponential distribution is given by:

$$f(t) = \lambda e^{-\lambda t}$$

where λ (lambda) is the sole parameter of the distribution. This form of the exponential is a one-parameter distribution. Based on the previous definition of the reliability function, it is a relatively easy matter to derive the reliability function for the exponential distribution:

$$
\begin{aligned}
R(t) &= 1 - \int_0^t \lambda e^{-\lambda s} ds \\
&= 1 - \left[1 - e^{-\lambda \cdot t}\right] \\
&= e^{-\lambda \cdot t}
\end{aligned}
$$

The form of the exponential distribution *pdf* makes such derivations simple (which often leads to inappropriate use of this particular distribution). For the derivation of the reliability functions for other distributions, including the Weibull, normal and lognormal, see *ReliaSoft's Life Data Analysis Online Reference* at http://reliawiki.org/index.php/Life_Data_Analysis_Reference_Book.