STATISTICAL ANALYSIS OF RELIABILITY DATA

Statistical analysis of reliability data can be considered generally as a methodological basis for the development and validation of probabilistic models used in reliability engineering. The statistical methods used for reliability data analysis are very similar to those used in biomedical studies, where lifetime or survivorship models are the subjects of interest; therefore, the term "lifetime" or "survival data analysis" used in many publications includes either type of application—engineering and biomedical sciences.

Any data analysis technique is based on a corresponding probabilistic model. The basic probabilistic models considered in reliability data analysis are:

1. *Time-independent reliability models.* Under these models, strength and/or stress are considered as time-independent (static) random variables (r.v.). The models are widely used in engineering design (see STRESS-STRENGTH MODELS). A special group of the time-independent reliability models constitutes the models associated with the binomial distribution. Typical examples are the models dealing with the probability of failure to start on demand for a redundant unit.
2. *Time-dependent reliability models* can be divided into the following classes:
   a. *Reliability models without explanatory factors.* Under these models, the reliability function is time-dependent. The explanatory or *stress* factors (such as temperature, humidity, voltage, and so on) are considered as constant, having no influence on reliability.
   b. *Reliability models with explanatory factors.* Typical examples of these models are the accelerated life model and proportional hazard (PH) model (see ACCELERATED LIFE TESTING). The reliability models with explanatory factors can be divided into the following groups: reliability models with constant stress factors and reliability models with time-dependent stress factors.

Reliability problems associated with repairable units are considered using special *repair and replacement models* based on the notion of *point process.* The most commonly used models are the homogeneous Poisson process, nonhomogeneous Poisson process, and the renewal process (see REPAIRABLE SYSTEMS).

Problems of statistical data analysis for all the probabilistic models mentioned previously can be, generally speaking, reduced to two types of statistical inferences—estimation and hypothesis testing. The statistical estimation includes distribution estimation and/or the random process estimation. Statistical estimation procedures as well as hypothesis testing can also be divided in two groups: *parametric* and *nonparametric* ones. The hypothesis testing in the context given includes many different hypotheses associated with particular life distributions and/or their general properties (for instance, *aging*) and similar hypotheses related to random processes. In this article, the classical approach to reliability data analysis is considered. The Bayesian data analysis is given in BAYESIAN INFERENCE IN RELIABILITY.

An important characteristic feature of reliability data analysis is associated with the so-called data censoring. Reliability data are very seldom complete samples, that is, the samples are composed of distinct times to failure (TTF) or numbers of cycles to failure. A much more realistic situation is one in which, for example, for a sample of n times to failure only $r \leqslant n$ times to failure are known, while for $n - r$ failure times are known only to be less than a given value.

BASIC DISTRIBUTIONS

In this section we consider some basic lifetime distributions used as probabilistic models for unrepairable units as well as some auxiliary distributions needed for statistical analysis.

Binomial Distribution

Let us consider a random trial having two possible outcomes: a success, with probability p, and a failure with probability $1 - p$. Such trials are known as Bernoulli trials. Consider a sequence of n Bernoulli trials. The distribution of the number of successes, x, in the sequence is known as the binomial distribution. The probability of observing x successes in n Bernoulli trials is known as the binomial probability density function, which is given by

$$f(x;p,n) = \binom{n}{x} p^x (1-p)^{n-x} \tag{1}$$

Because the random variable x can take on only positive integer values, it is obvious that the binomial distribution is a discrete one. The mean and variance of the binomial distribution are

$$\begin{aligned} E(x) &= np \\ \mathrm{Var}(x) &= \sigma^2(x) = np(1-p) \end{aligned} \tag{2}$$

The binomial distribution plays a fundamental role in reliability. Suppose that n identical units are tested (without replacement of the failed units) for a specified time, t, and that the test results in r failures. The number of failures, r, can be considered as a discrete random variable having the binomial distribution with parameters n and $p(t)$, where $p(t)$ is now the probability of failure of a single unit during time t. In other words, $p(t)$ is the probability of success in a Bernoulli trial, which is a test of a single unit during time, t, and "success" is the respective failure. Thus, $p(t)$, in a sense, as a function of time, is the time-to-failure cumulative distribution function, as well as $1 - p(t)$ is the reliability or survivor function. As an example of straightforward application of the binomial distribution, we can mention a model for the number of failures to start on demand for a redundant unit (which is the time-independent reliability model). The probability of failure in this case might be considered as time independent. Thus, we should keep in mind two possible applications of the binomial distribution: (1) the survivor (reliability) function or TTF cumulative distribution function, and (2) the proper binomial distribution.

Poisson Distribution

Another discrete distribution widely used in reliability is the Poisson distribution. Assume that some objects are evenly dispersed at random on a large domain, with some specified density λ (e.g., events of instant duration appearing in a finite time interval). Then the probability to observe k objects (events) in some specified domain, Ω, has the Poisson distribution:

$$P(k; \Lambda) = \frac{\Lambda^k}{k!} e^{-\Lambda} \tag{3}$$

where Λ is the mean number of objects in domain Ω, that is, $\Lambda = \lambda\Omega$. The variance of this distribution is equal to its mean, that is, $\text{var}(k) = \Lambda$. In reliability application, the mean of the Poisson distribution is used in the form of the product λt, having the meaning of a mean number of the events (failures) observed during time t where λ is the failure (hazard) rate. The binomial distribution approaches the Poisson one with $\Lambda = np$ when n is large enough and p (or $1 - p$) approaches zero.

Exponential Distribution

Among the continuous distributions used in reliability the *exponential distribution* can be considered basic. As a time-to-failure distribution, it appears in many reliability problems, some of which are considered below. The exponential distribution is a model of an item subjected to random fatal shocks. If these shocks arrive according to the Poisson process with intensity (failure rate) λ, the item will fail at the moment when a shock occurs. The intervals between these events have the exponential distribution, so the same distribution can be applied to the time to failure of an item.

The probability density function of the exponential distribution is

$$f(t) = \lambda e^{-\lambda t} \tag{4}$$

The corresponding reliability function, $R(t)$, the mean time to failure (MTTF), and variance are

$$R(t) = e^{-\lambda t}, \quad t \ge 0 \tag{5}$$

$$\begin{aligned} \text{MTTF} &= \int_0^\infty t f(t)\, dt = \int_0^\infty R(t)\, dt = \int_0^\infty e^{-\lambda t}\, dt = \frac{1}{\lambda} \\ \text{Var}(t) &= \frac{1}{\lambda^2} \end{aligned} \tag{6}$$

The coefficient of variation (the ratio of a standard deviation to the mean) for the exponential distribution is 1, and it can be used as a quick test for exponentiality.

The unique property of the exponential distribution is that the distribution of the remaining life of a used item (residual TTF) is independent of its initial age. This property is often referred to as the "memoryless" property. This property easily follows from consideration of the conditional reliability function, which is the probability that an item will not fail for time x, if it has survived time t, which is given by:

$$R(x|t) = \frac{R(x+t)}{R(t)} \tag{7}$$

In the case of the exponential distribution, the conditional reliability function obviously is

$$R(x|t) = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} \tag{8}$$

Classes of Distribution Based on Aging. A notion of *aging* in reliability theory is introduced in terms of failure rate. A fundamental study of distributions with a monotonic failure rate can be found in Barlow and Proschan (1).

The conditional reliability function from Eq. (7) can be interpreted as the reliability function of a unit of age t. This is the probability that a unit will not fail during the additional time, x, given that it has not failed by time t. Consider the case where the conditional reliability function is a decreasing function of time. Let $H(t)$ be the cumulative hazard rate, associated with a given TTF distribution, and let $r(t)$ be the respective failure rate. Equation (7) can be written in terms of these functions as:

$$R(X|t) = \frac{\exp[-H(t+x)]}{\exp[-H(t)]} = \exp\left[\int_t^{t+x} r(u)\, du\right] \tag{9}$$

Clearly, if $r(t)$ is increasing, $R(x|t)$ is a decreasing function of t. The distributions having this property constitute a class of increasing failure rate (IFR) distributions. Alternatively, a TTF distribution belongs to a class of decreasing failure rate (DFR) distributions if its conditional reliability, $R(x|t)$, increases with time t for each $x \ge 0$. Note again that the exponential distribution is the only distribution with a constant failure rate, so it is a "boundary" distribution between IFR and DFR classes. Another commonly used class of aging distributions is the class of increasing failure rate average (IFRA) distributions, that is, a class of TTF distributions for which the average failure rate introduced as

$$\langle r(t)\rangle = \frac{1}{t}\int_0^t r(x)\, dx = -\frac{\log R(t)}{t} = \frac{H(t)}{t} \tag{10}$$

is increasing with time. Similarly, the class of decreasing failure rate average (DFRA) is introduced as the class of distributions with a decreasing average failure rate. The properties of these distributions are widely used in reliability data analysis. Many other classes of TTF distributions are also used, such as NBU (new better than used), NBAFR (new better than used average failure rate) and so on (1,2).

In practice, we usually deal with a more general case: unit's failure rate forms a U-shaped function that is also sometimes called a *life characteristic curve.* Numerous reliability and life data give a similar bathtub shape for a plot of failure rate versus time. This curve is divided into the three parts corresponding to three age periods. The first interval is a period with a decreasing failure rate, known as an infant mortality period. In reliability data analysis, the failures in this period are usually related to manufacturing defects. The period of infant mortality is followed by a period with an approximately constant failure rate. This is the period of catastrophic failures that are due mostly to accidental overloads or shocks. The last age period, characterized by an increasing failure rate, is a period of wearout failures associated with material fatigue or wear.

Cumulative Damage Model Resulting in IFRA Distributions. Consider a unit subjected to shocks occurring randomly in time. Let these shocks arrive according to the Poisson process with constant intensity λ; each ith shock causes a random amount of damage, x_i. All x_1, x_2, . . . are independent and identically distributed random variables, with a distribution function F (a damage distribution function). The unit fails when the accumulated damage exceeds a specified threshold x. It can be shown (1) that for any damage distribution function F, the time-to-failure distribution function is IFRA.

Bounds on Reliability for Aging Distributions. The following simple bounds are given in terms of IFRA (DFRA) and IFR (DFR) distributions (1). Any IFR distribution obviously is a subclass of the class of IFRA distributions.

Bounds Based on a Known Percentile. Let t_p be the 100pth percentile of an IFRA (DFRA) distribution. Then

$$R(t)\begin{cases}\geq(\leq)e^{-\alpha t} & \text{for}\quad 0\leq t\leq t_p\\ \leq(\geq)e^{-\alpha t} & \text{for}\quad t\geq t_p\end{cases} \tag{11}$$

where $\alpha = -(1/t_p)\log(1 - p)$.

Bounds Based on a Known Mean. Let a time-to-failure distribution be IFR with known mean μ_a. Then

$$R(t)\geq\begin{cases}e^{-t/\mu_a} & \text{for}\quad t<\mu_a\\ 0 & \text{for}\quad t\geq\mu_a\end{cases} \tag{12}$$

Inequality for Coefficient of Variation. Let a time-to-failure distribution be IFRA (DFRA) with mean μ_a and variance σ^2. Then, the coefficient of variation is $\sigma/\mu_a \leq (\geq)$ 1. Recall that, for the exponential distribution, the variation coefficient is equal to 1. This criterion is useful for reliability data analysis, but the inequality is only the necessary condition for IFRA (DFRA) distribution.

Weibull Distribution

The Weibull distribution is one of the most popular models for TTF distributions. This distribution was introduced as a model for bearing failures caused by the wearing process. The Weibull distribution can be also obtained as a limit law for the distribution of the smallest-order statistic (the "weakest link" model). It can be also obtained as the TTF distribution for an item subjected to fatal shocks occurring randomly in time in accordance with a time-dependent Poisson process (i.e., with a time-dependent parameter λ). Let the failure rate $r(t)$ be a power function of time t:

$$\begin{aligned}r(t)&=\frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} && t\geq 0,\quad \beta,\alpha>0\\ &=0 && t<0\end{aligned} \tag{13}$$

Using this failure rate and the following basic relationship between the failure rate $r(t)$ and the respective cumulative distribution function (CDF),

$$F(t)=1-\exp\left[-\int_0^t r(\tau)\,d\tau\right] \tag{14}$$

the CDF of the Weibull distribution can be written as

$$F(t)=1-\exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]\qquad t\geq 0 \tag{15}$$

where α is the scale parameter and β is the shape parameter. The scale parameter is also referred to as the "characteristic life," which is $100(1 - e^{-1}) \approx 63.2$th percentile. Note that if t has the two-parameter Weibull distribution, the transformed random variable $x = \ln t$ has the so-called Type I asymptotic distribution of smallest (extreme) values:

$$F(x)=1-\exp\left[-\exp\left(\frac{x-u}{b}\right)\right]\qquad -\infty<x<\infty \tag{16}$$

where $u = \ln \alpha$ and $b = 1/\beta$.

The mean of the two-parameter Weibull distribution is given by

$$\mu=\alpha\Gamma\left(1+\frac{1}{\beta}\right) \tag{17}$$

where

$$\Gamma(c)=\int_0^\infty x^{c-1}e^{-x}\,dx \tag{18}$$

is the gamma function and the variance of the Weibull distribution is

$$\sigma^2=\alpha^2\left[\Gamma\left(1+\frac{2}{\beta}\right)-\Gamma^2\left(1+\frac{1}{\beta}\right)\right] \tag{19}$$

From the expression for the failure rate in Eq. (13), it is obvious that the Weibull distribution is IFR distribution for $\beta >$ 1 and DFR distribution for $0 < \beta < 1$. For $\beta = 1$, the Weibull distribution coincides with the exponential one. Thus, this distribution can be used as a model for any part of the life characteristic curve. This ability of the Weibull distribution to reflect such a wide class of distributions makes it popular for different reliability engineering applications.

Using the properties of the gamma function, it can be shown that if the shape parameter increases, the mean μ approaches the scale parameter α, and the variance σ^2 also approaches zero; so for $\beta > 1$ the coefficient of variation, being less than 1, also approaches zero.

Logarithmic Normal Distribution

Another popular TTF distribution is the logarithmic normal, or lognormal, distribution. There are many reasons for its widespread applicability to reliability. The first is the availability of statistical data analysis based on the normal (Gaussian) origin of the lognormal distribution. The distribution was used as an adequate model in many engineering problems. For example, it was used by many authors as a distribution of particle sizes in naturally occurring or artificially obtained aggregates. As the TTF distribution, the lognormal distribution also arises from simple physical considerations in fracture mechanics problems (3,4).

Let x be a normally distributed random variable, with the mean μ and the standard deviation σ. It is easy to show that the random variable $t = e^x$ has a lognormal distribution with parameters μ and σ, that is, the probability density function (PDF) of the variable t is

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma t} \exp\left[-\frac{1}{2}\left(\frac{\log t - \mu}{\sigma}\right)^2\right] \qquad t \geq 0 \tag{20}$$

where $\mu = E(\log t)$ and $\sigma^2 = \text{Var}(\log t)$ are, respectively, the mean and the variance of $\log t$. The parameters μ and σ are the distribution parameters, but they are not the mean and the variance of the lognormal distribution.

The MTTF and the variance of the lognormal distribution are given by

$$\text{MTTF} = \exp\left[\mu + \frac{\sigma^2}{2}\right] \tag{21}$$

$$\text{Var}(t) = e^{(2\mu+\sigma^2)}(e^{\sigma^2} - 1) = \text{MTTF}^2(e^{\sigma^2} - 1) \tag{22}$$

In practice, almost all lognormal data analysis procedures are based on taking logarithms of all the failure times and analyzing the transformed data as though they were normal distribution data. The lognormal CDF, $F(t)$, is

$$F(t) = \int_0^t \frac{1}{\sqrt{2\pi}\sigma s} \exp\left[\frac{1}{2}\left(\frac{\log s - \mu}{\sigma}\right)^2\right] ds \tag{23}$$

Substituting $s = (\log t - \mu)/\sigma$ and $ds = dt/(\sigma t)$, the CDF can be expressed as

$$F(t) = \int_{-\infty}^{\frac{\log t - \mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-1/2s^2}\, ds = \Phi\left(\frac{\log t - \mu}{\sigma}\right) \tag{24}$$

where $\Phi(\cdot)$ is the standard normal CDF. The respective reliability function can be written as

$$R(t) = 1 - F(t) = \Phi\left(-\frac{\log t - \mu}{\sigma}\right) \tag{25}$$

These expressions show that e^μ is the median of the lognormal distribution. The failure rate function of the lognormal life distribution is given by

$$r(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{\Phi\left(\frac{\log t - \mu}{\sigma}\right)} \tag{26}$$

Note that the failure rate function of the lognormal distribution is not only nonmonotonic but it always has a maximum. The failure rate, $r(t) = 0$ at $t = 0$, increases with age to the maximum and after that it decreases to zero with increasing age. Thus, such a failure rate time dependence does not correspond to any part of the life characteristic curve. Nevertheless, the lognormal distribution (especially its "left tail" with increasing failure rate) fits numerous reliability data.

Gamma Distribution

The *gamma,* or *Erlang,* distribution of the nth order is the distribution of the sum of a fixed (nonrandom) number, n, of independent and identically exponentially distributed r.v.'s. For reliability applications, the gamma distribution can be introduced in the following way. If a system consists of n independent components, $n - 1$ of which are redundant, and all the units have the same exponentially distributed TTF with failure rate λ, the TTF of the system is considered as the time to nth failure, that is, it is given by the distribution of the sum of n independent exponentially distributed r.v.'s. The distribution of this sum is the gamma distribution.

It is obvious that the exponential distribution is a particular case of the gamma distribution. If the exponential r.v.'s, composing the gamma distribution are not identical, the corresponding distribution is called a *generalized gamma* distribution [which is considered, for example, in Gnedenko and Ushakov (5)].

As mentioned before, the distribution of intervals in the Poisson process is the exponential one. Let $T(n)$ be the sum of n independent random intervals each exponentially distributed with a failure rate λ. It is obvious that the probability of the event $[T(n) > t]$ is equal to the probability of observing during time t the number of event N less than n. The latter probability is given by the Poisson distribution:

$$R(t) = \sum_{i=0}^{n-1} \frac{(\lambda t)^i e^{-\lambda t}}{i!}, \qquad t > 0 \tag{27}$$

It is clear the function in Eq. (27) is the reliability function of the gamma distribution with positive parameters λ and n, where n is an integer. The corresponding PDF is given by

$$f(t) = \frac{\lambda(\lambda t)^{n-1} e^{-\lambda t}}{(n-1)!}, \qquad t > 0 \tag{28}$$

It is easy to see that the mean and the variance of the gamma distribution are given by

$$\text{MTTF} = \frac{n}{\lambda}, \qquad \text{Var}(t) = \frac{n}{\lambda^2} \tag{29}$$

The particular case of the gamma distribution, when $\lambda = \frac{1}{2}$ and $n = k/2$ where k is a positive integer, is the chi-square distribution with k degrees of freedom [if $k/2$ is not an integer, $(k/2 - 1)!$ must be replaced by $\Gamma(k/2)$]. The chi-square and Poisson distributions are related to each other as (6)

$$\sum_{i=0}^{n-1} \frac{(\lambda t)^i e^{-\lambda t}}{i!} = \text{Prob}(\chi^2_{2n} > 2\lambda t) \tag{30}$$

In general, for any positive η the CDF of the gamma distribution is

$$F(t) = \int_0^t \frac{\lambda(\lambda t)^{\eta-1} e^{-\lambda t}}{\Gamma(\eta)}, \qquad t > 0 \tag{31}$$

The gamma distribution is:

- DFR, if $\eta < 1$
- The exponential distribution, if $\eta = 1$
- IFR, if $\eta > 1$

Besides the previously mentioned obvious applications, the gamma distribution is also widely used as the prior distribution in Bayesian reliability estimation, and it is a popular model in the theory of queuing processes and in the theory of birth-death processes.

BASIC DISTRIBUTION ESTIMATION METHODS

There exist two basic kinds of estimation, *point estimation* and *interval estimation.* Point estimation provides a single number from a set of observational data to represent a parameter of the distribution. Point estimation does not give any information about its accuracy. Interval estimation is expressed in terms of *confidence intervals.* The confidence interval includes the true value of the parameter with a specified confidence probability.

Point Estimation

Estimation of a parameter is necesarily based on a set of sample values, $X_1, \ldots, X_n$. If the sample values are independent and their underlying distribution remains the same from one sample to another, we have a random sample of size n from the distribution of the random variable of interest X. Let the distribution have a parameter θ. A random variable $t(X_1, \ldots, X_n)$, which is a single-valued function of $X_1, \ldots, X_n$, is referred to as a *statistic.* A point estimate is obtained by selecting an appropriate statistic and calculating its value from the sample data. The selected statistic is called an *estimator,* while the value of the statistic is called an *estimate.*

Consider the main properties of estimators. An estimator $t(X_1, \ldots, X_n)$ is said to be an *unbiased* estimator for θ if its expectation $E[t(X_1, \ldots, X_n)] = \theta$ for any value of θ. Another desirable property of an estimator $t(X_1, \ldots, X_n)$ is the property of *consistency.* An estimator t is said to be consistent if, for every $\epsilon > 0$,

$$\lim_{n\to\infty} P(|t(x_1, \ldots, x_n) - \theta| < \epsilon) = 1 \tag{32}$$

This property implies that as the sample size n increases, the estimator $t(X_1, \ldots, X_n)$ gets closer to the true value of θ. In some situations several unbiased estimators can be found. An unbiased estimator t of θ, having minimum variance among all unbiased estimators of θ, is called *efficient.* Another estimation property is *sufficiency.* An estimator $t(X_1, \ldots, X_n)$ of the parameter θ is said to be sufficient if it contains all the information about θ that is in the sample $X_1, \ldots, X_n$.

The most common methods of point estimation are briefly considered below.

Method of Moments. In the framework of this method, the estimators are equated to the corresponding distribution moments. The solutions of the equations obtained provide the estimators of the distribution parameters. For example, as the mean and variance are the expected values of X and $(X - \mu)^2$, respectively, the sample mean and sample variance are defined, as the respective expected values of a sample of size n, as:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{33}$$

and

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \tag{34}$$

where $\overline{X}$ and S^2, respectively, are the point estimates of the distribution mean, μ, and variance, σ^2.

The estimator of variance in Eq. (34) is biased; however, this bias can be removed by multiplying it by $n/(n - 1)$:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \tag{35}$$

Example 1 The life, T, of a component is considered as a random variable having the exponential distribution. The times to failure (in relative units) obtained from the component life test are 3, 8, 12, 35, 42, 42.5, 77, 141, 152.5, and 312. Since the exponential distribution is a one-parameter distribution, only the first moment is used, thus:

$$\bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i = \frac{1}{10}\sum_{i=1}^{10} t_i = 82.5$$

The relationship between the mean and parameter λ for the exponential distribution is $\theta = 1/\lambda$. Therefore, an estimator of λ is $\overline{\lambda} = 1/\bar{t} = 0.0121$.

Method of Maximum Likelihood. The method of maximum likelihood (ML) is one of the most popular methods of estimation. Consider a random variable, X, with probability density function $f(x, \theta_0)$, where θ_0 is the unknown parameter. Using the method of maximum likelihood, one can try to find the value of θ_0 that has the highest (or most likely) probability (or probability density) of producing the given set of observations, $X_1, \ldots, X_n$. The likelihood of obtaining this set of sample values is proportional to the PDF $f(x, \theta_0)$ evaluated at the sample values $X_1, \ldots, X_n$. The likelihood function for a continuous distribution is introduced as

$$L_f(X_1, \ldots, X_n;\ \theta_0) = f(X_1, \theta_0) f(X_2, \theta_0), \ldots, f(X_n, \theta_0) \tag{36}$$

Generally, the definition of the likelihood function is based on the probability (for a discrete random variable) or the PDF (for continuous random variable) of the joint occurrence of n events, $X = X_1, \ldots, X = X_n$. The maximum likelihood estimate, $\hat{\theta}_0$, is the value of θ_0 that maximizes the likelihood function, $L_f(X_1, \ldots, X_n;\ \theta_0)$, with respect to θ_0.

The usual procedure for maximization with respect to a parameter is to calculate the derivative with respect to this parameter and equate it to zero. This yields the equation

$$\frac{\partial L_f(X_1, \ldots, X_n;\ \theta_0)}{\partial \theta_0} = 0 \tag{37}$$

The solution of the above equation for θ_0 will give $\hat{\theta}_0$, if it can be shown that $\hat{\theta}_0$ does indeed maximize $L_f(X_1, \ldots, X_n;\ \theta_0)$. Because of the multiplicative nature of the likelihood function, it is often more convenient to maximize the logarithm of the likelihood function instead; that is,

$$\frac{\partial \log L_f(X_1, \ldots, X_n;\ \theta_0)}{\partial \theta_0} = 0 \tag{38}$$

Because the logarithm is monotonic transformation, the solution for θ from this equation is the same as that obtained from Eq. (37). For a probability density function with m parameters, the likelihood function becomes

$$L_f(X_1,\ldots,X_n;\ \theta_1,\ldots,\theta_m) = \prod_{i=1}^{n} f(X_i,\theta_1,\ldots,\theta_m) \tag{39}$$

where $\theta_1, \ldots, \theta_m$ are the m parameters to be estimated. In this case, the maximum likelihood estimators can be obtained by solving the following system of m equations:

$$\frac{\partial L_l(X_1,\ldots,X_n;\theta_1,\ldots,\theta_m)}{\partial\theta_j} = 0, \qquad j = 1,\ldots,m \tag{40}$$

Under some general conditions, the obtained *maximum likelihood estimates* are consistent, asymptotically efficient, and asymptotically normal.

Example 2 Let us estimate the parameter p of the binomial distribution. In this case,

$$L_f(m \mid n) = \binom{n}{m} p^m (1-p)^{n-m}, \qquad m = 0,1,\ldots,n$$

Taking the derivative, and equating it to zero

$$\frac{\partial\, \mathrm{Log}\, L_f}{\partial p} = \frac{n}{p(1-p)}\left(\frac{m}{n} - p\right) = 0$$

we find that the maximum likelihood estimator $\hat{p} = m/n$. In general, the maximum likelihood method requires use of numerical optimization methods.

Example 3 For the life-test data given in Example 1, estimate the parameter of the distribution, using the method of maximum likelihood. The maximum likelihood function for this problem is

$$L_f(t_1,\ldots,t_{10},\lambda) = \prod_{i=1}^{10} f(t_i,\lambda) = \lambda^{10} e^{-\lambda}$$

Taking the derivative yields the following equation

$$\frac{dL_f(t_1,\ldots,t_{10},\lambda)}{d\lambda} = \left[10\lambda^9 - \lambda^{10}\sum_{i=1}^{10} t_i\right] = 0$$

which has the following solution

$$\hat{\lambda} = \frac{10}{\sum_{i=1}^{10} t_i} = 0.0121$$

In this example, the estimates by the method of moments and the method of ML coincide.

Interval Estimation

Let $l(X_1, \ldots, X_n)$ and $u(X_1, \ldots, X_n)$ be two statistics, such that the probability that parameter θ_0 lies in an interval $[l, u]$ is

$$P\{l(X_1,\ldots,X_n) < \theta_0 < u(X_1,\ldots,X_n)\} = 1 - \alpha$$

The random interval $[l, u]$ is called a $100(1 - \alpha)\%$ confidence interval for the parameter θ_0. The endpoints l and u are referred to as the $100(1 - \alpha)\%$ confidence limits of θ_0; $(1 - \alpha)$ is called the confidence coefficient or confidence level. The most commonly used values for α are 0.10, 0.05, and 0.01. If $\theta_0 > l$ ($\theta_0 < u$) with probability equal to 1, then $u(l)$ is called one-sided upper (lower) confidence limit for θ_0. A $100(1 - \alpha)\%$ confidence interval for an unknown parameter θ_0 is interpreted as follows: if a series of repetitive experiments yields random samples from the same distribution, and the confidence interval for each sample is calculated, then $100(1 - \alpha)\%$ of the constructed intervals will, in the long run, contain the true value of θ_0.

The following example illustrates the common principle of confidence limits construction. Consider the procedure for constructing confidence intervals for the mean of a normal distribution with known variance. Let $X_1, X_2, \ldots, X_n$, be a random sample from the normal distribution, $N(\mu, \sigma^2)$, in which μ is an unknown parameter, and σ^2 is assumed to be known. It can be shown that the sample mean has the normal distribution $N(\mu, \sigma^2/n)$. Thus, $(X - \mu)\sqrt{n}/\sigma$ has the standard normal distribution. This means that

$$P\left(-z_{1-(\alpha/2)} \le \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \le z_{1-(\alpha/2)}\right) = 1 - \alpha \tag{41}$$

where $z_{1-(\alpha/2)}$ is the $100(1 - \frac{1}{2}\alpha)$th percentile of the standard normal distribution $N(0,1)$. Solving the inequalities inside the parentheses, Eq. (41) can be rewritten as

$$P\left(\overline{X} - z_{1-(\alpha/2)}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z_{1-(\alpha/2)}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha \tag{42}$$

Thus, Eq. (42) gives the symmetric $(1 - \alpha)$ confidence interval for the mean, μ, of a normal distribution with known σ^2. The confidence interval is wider for a higher confidence coefficient $(1 - \alpha)$. As σ decreases, or n increases, the confidence interval becomes smaller for the same confidence coefficient $(1 - \alpha)$.

HYPOTHESIS TESTING

Interval estimation and hypothesis testing are, in a sense, mutually inverse procedures. Consider an r.v. X with a known probability density function $f(x; \theta)$. Using a random sample from this distribution one can obtain a point estimate $\hat{\theta}$ of the parameter θ. Assume that a hypothesized value of θ is θ_0. Under these circumstances, the following question can be raised: Is the estimate value $\hat{\theta}$, in a sense, compatible with the hypothesized value θ_0? In terms of statistical hypothesis testing the statement $\theta = \theta_0$ is called the *null hypothesis,* which is denoted by H_0. The null hypothesis is always tested against an *alternative hypothesis,* denoted by H_1, which for the case considered might be the statement $\theta \neq \theta_0$. The null and alternative hypotheses are also classified as *simple,* or *exact* (when they specify exact parameter values) and *composite,* or *inexact* (when they specify an interval of parameter values). For the example considered, H_0 is simple and H_1 is composite. An example of a simple alternative hypothesis might be H_1: $\theta = \theta^*$.

For testing statistical hypotheses test statistics are used. In many situations the test statistic is the point estimator of the unknown distribution. In this case (as in the case of the interval estimation) one has to obtain the distribution of the test statistic used. Let X_1, X_2, . . ., X_n, be again a random sample from the normal distribution, $N(\mu, \sigma^2)$, in which μ is an unknown parameter, and σ^2 is assumed to be known. One has to test the simple null hypothesis H_0: $\mu = \mu^*$ against the composite alternative H_1: $\mu = \mu^*$. As the test statistic, let us use the same sample mean, $\overline{X}$, which has the normal distribution $N(\mu, \sigma^2/n)$. Having computed the value of the test statistic $\overline{X}$, we can construct the confidence interval Eq. (42) and find out whether the value of μ^* is inside the interval. This is the test of the null hypothesis. If the confidence interval includes μ^*, the null hypothesis is not rejected at significance level α.

In terms of hypothesis testing, the confidence interval considered is called the *acceptance region,* the upper and the lower limits of the acceptance region are called the *critical values,* and the significance level α is referred to as a probability of *type I error.* In deciding whether or not to reject the null hypothesis, it is possible to commit the following errors:

- Reject H_0 when it is true (type I error)
- Not reject H_0 when it is false (type II error—the probability of the type II error is designated by β.)

These situations are traditionally represented by the following table:

Decision	State of Nature	
	H_0 Is True	*H_0 Is False*
Reject H_0	Type I error	No error
Do not reject H_0	No error	Type II error

It is clear that increasing the acceptance region, which results in decreasing α, simultaneously results in increasing β. The traditional approach to this problem is to keep the probability of type I error α at a low level (0.01, 0.05 or 0.10) and to minimize the probability of type II error as much as possible. The probability of not making a type II error is referred to as the *power of the test.*

In reliability data analysis one often needs a statistical procedure to assess the quality of the distribution model fitted for the data given. Such procedures constitute the special class of hypothesis tests known as the *goodness-of-fit tests.* Two of the most commonly used tests, the chi-square and Kolmogorov-Smirnov tests, are briefly discussed below.

Chi-Square Test

Consider a sample of N observed values (measurements) of a random variable. The chi-square goodness-of-fit test compares the observed frequencies (histogram), n_1, n_2, . . ., n_k, of k intervals of the random variable with the corresponding frequencies, e_1, e_2, . . ., e_k, from an assumed theoretical distribution, $F_0(x)$. The basis for this goodness-of-fit testing is the distribution of the statistic

$$\sum_{i=1}^{k} \frac{(n_i - e_i)^2}{e_i} \tag{43}$$

This statistic has an approximate chi-square (χ^2) distribution with $f = k - 1$ degrees of freedom. If the parameters of the theoretical distribution are unknown and are estimated from the same data, the above distribution remains valid, given the number of degrees of freedom is reduced by one for every unknown parameter that must be estimated. Thus, if an assumed distribution yields a result such that

$$\sum_{i=1}^{k} \frac{(n_i - e_i)^2}{e_i} < C_{1-\alpha,f} \tag{44}$$

where the critical value, $C_{1-\alpha,f}$, is the value of the χ^2 corresponding to the cumulative probability $(1 - \alpha)$, then the assumed theoretical distribution is not rejected (i.e., the null hypothesis H_0: $F(x) = F_0(x)$ is not rejected) at significance level α. If the inequality in Eq. (44) is not satisfied, the alternative hypothesis H_1: $F(x) \neq F_0(x)$ is accepted. Employing the χ^2 goodness-of-fit test, it is recommended that at least five intervals be used ($k \geq 5$), with at least five expected observations per interval ($e_i \geq 5$) to obtain satisfactory results. The steps for conducting the χ^2 test are as follows:

- Divide the range of data into intervals (number of intervals > 5), the first and the last being infinite intervals, and count n_i = the number of measurements in each interval.
- Estimate the parameters of the assumed theoretical distribution, $F_0(x)$, and calculate the theoretical quantity of data in each interval, e_i, as follows:

$$e_i = [F_0(x + \Delta x) - F_0(x)] \cdot [\text{sample size}]$$

- Calculate statistic using Eq. (43).
- Choose a specified significance level, α (generally, $1 - \alpha$ = 90 or 95 percent).
- Determine the number of degrees of freedom of the χ^2 distribution:

$$f = k - 1 - [\text{number of parameters } of\ F_0(x)]$$

- Determine $C_{1-\alpha}$ from the table and compare it with the obtained value of Eq. (43). If the inequality of Eq. (44) is satisfied, then the assumed theoretical distribution function, $F_0(x)$, is not rejected.

Example 4 The sea wave loads acting on marine structures are the objects of probabilistic reliability design. The sample of 219 measurements of wave bending moments (in arbitrary units) is given in Table 1.

Table 1. Wave Bending Moments

Interval Number	Interval Start	Number of Measurements in Each Interval
1	0.00	11
2	2526.31	25
3	5052.63	34
4	7578.94	35
5	10105.26	37
6	12631.57	27
7	15157.89	23
8	17684.20	15
9	20210.52	6
10	22763.83	3
11	25263.14	2
12	27789.46	1

For the data given, the Weibull distribution was fitted in the form:

$$F(x) = 1 - \exp(-\lambda x^{\gamma})$$

The obtained estimates of the parameters are: $\hat{\lambda} = 1.018$ 10^{-8} and $\hat{\gamma} = 2.327$. Based on these estimates and the data in the table, the chi-square statistic is 2.33, and it has 9 degrees of freedom. This value of statistic is much less than the corresponding critical value, 14.7, chosen at the robust significance level 0.10. Thus, the hypothesis about the Weibull distribution is not rejected.

Kolmogorov-Smirnov Test

Another widely used goodness-of-fit test is the Kolmogorov-Smirnov (K-S) test. The basic procedure involves comparing the empirical (or sample) cumulative distribution function with an assumed distribution function. If the maximum discrepancy is large compared with what is anticipated from a given sample size, the assumed distribution is rejected.

Consider a sample of n observed values of a continuous random variable. The set of the data is rearranged in increasing order: $t_{(1)} < t_{(2)} < \ldots t_{(n)}$. Using the ordered sample data, the empirical distribution function $S_n(x)$, is defined as follows:

$$S_n(t) = \begin{cases} 0 & -\infty < t < t_{(1)} \\ \dfrac{i}{n} & t_{(i)} \le t < t_{(i+1)} \\ 1 & t_{(n)} \le t < \infty \end{cases} \quad (45)$$

$$i = 1, \ldots, n-1$$

where $t_{(1)}, t_{(2)}, \ldots, t_{(n)}$ are the values of the ordered sample data (the order statistics). It can be shown that the empirical distribution function is a consistent estimator for the corresponding cumulative distribution function.

In the K-S test, the test statistic is the maximum difference between $S_n(t)$ and $F_0(t)$ over the entire range of random variable t. It is clear that the statistic is a measure of the discrepancy between the theoretical model and the empirical distribution function. The K-S statistic is denoted by

$$D_n = \max_x |F_0(t) - S_n(t)| \quad (46)$$

If the null hypothesis is true, the probability distribution of D_n turns out to be the same for every possible continuous $F_0(t)$. Thus, D_n is a random variable whose distribution depends on the sample size, n, only. For a specified significance level, α, the K-S test compares the observed maximum difference with the critical value D_n^{α}, defined by

$$P(D_n \le D_n^{\alpha}) = 1 - \alpha \quad (47)$$

Critical values, D_n^{α}, at various significance levels, α, are tabulated (3). If the observed D_n is less than the critical value D_n^{α}, the proposed distribution is not rejected. The steps for conducting the K-S test are as follows:

- For each sample item datum, calculate the $S_n(t_{(i)})$ ($i = 1, \ldots, n$) according to Eq. (45).
- Estimate the parameters of the assumed theoretical distribution, $F_0(t)$, using another sample or any information not associated with the sample given, and calculate $F_0(t_{(i)})$ from the assumed distribution function. If the parameters of the distribution are estimated using the same data, the special modifications of the test must be used; see Lawless (7).
- Calculate the differences of $S_n(t_i)$ and $F_0(t_{(i)})$ for each sample item, and determine the maximum value of the differences according to Eq. (46).
- Choose a specified significance level, α, and determine D_n^{α} from the appropriate statistical table.
- Compare D_n with D_n^{α}. If $D_n < D_n^{\alpha}$, the assumed distribution function, $F_0(t)$, is not rejected.

CENSORED DATA

As mentioned earlier, reliability data are seldom complete samples, and typically sample data are censored. The likelihood function for a complete sample was introduced before. In this section, some basic types of censored data and the respective the likelihood functions are considered.

Left and Right Censoring

Let N be the number of items in a sample, and assume that all units of the sample are tested simultaneously. If during the test period, T, only r units have failed, their failure times are known, and the failed items are not replaced, the sample is called *singly censored on the right at T*. In this case, the only information we have about $N - r$ unfailed units is that their failure times are greater than the duration of the test, T. Formally, an observation is called "right censored at T" if the exact value of the observation is not known, but it is known that it is greater than or equal to T (7). The left censoring is introduced in an obvious way. This type of censoring practically never appears in reliability data collection practice. It is very important to understand the way in which censored data are obtained. The basic discrimination is associated with random and nonrandom censoring, the simplest cases of which are discussed below.

Type I Censoring

Consider a situation of right censoring. If the test is terminated at a given nonrandom time, T, the number of failures, r, observed during the test period will be a random variable. These censored data are an example of type I or time right singly censored data (sometimes called *time-terminated*). In general, a Type I censoring is considered under the following scheme of observations.

Let a sample of n units be observed during different periods of time $L_1, L_2, \ldots, L_n$. The TTF of an individual unit, t_i, is observed as a distinct value if it is less than the corresponding time period, that is, if $t_i < L_i$. This is the case of Type I multiply censored data; the case considered above is its particular case, when $L_1 = L_2 = \cdots = L_n = T$. Type I multiply censored data are quite common in reliability testing. For example, a test can start with sample size n but at some given times $L_1, L_2, \ldots, L_k$ ($k < n$) the prescribed numbers of units can be removed from the test.

For treating censored data, a special random variable is introduced (7). Suppose again that a sample of n units is observed during different periods of time $L_1, L_2, \ldots, L_n$. The times to failure t_i ($i = 1, 2, \ldots, r$) are considered as independently distributed r.v.'s having the continuous PDF, $f(t)$, and the CDF, $F(t)$. Under these assumptions, the data can be represented by the n pairs of random variables (τ_i, δ_i) that are given by

$$\tau_i = \min(t_i, L_i)$$
$$\delta_i = \begin{cases} 1 & \text{if } t_i < L_i \\ 0 & \text{if } t_i \geq L_i \end{cases}$$

where δ_i indicates whether the time to failure t_i is censored or not, while τ_i is simply the time to failure, if it is observed, or the time to censoring, if the failure of ith unit is not observed. Note that τ_i is a *mixed* r.v. having a continuous component (t_i) and a discrete component (L_i). It can be shown that the joint PDF of τ and δ is

$$f(\tau_i, \delta_i) = f(\tau_i)^{\delta_i}(1 - F(L_i))^{1-\delta_i}$$

so that the corresponding likelihood function, Lh, is given by

$$Lh = \prod_{i=1}^{n} f(\tau_i)^{\delta_i} S(L_i)^{1-\delta_i} \tag{48}$$

This last equation can be rewritten in a more tractable form as

$$L = \prod_{i \in U} f(t_i) \prod_{i \in C} S(L_i)$$

where U is the set containing the indexes of the items that failed during the test (uncensored observations) and C is the set containing the indexes of the items that did not fail during the test (censored observations). For the simple case above, when the simultaneous testing (without replacement) of N units is terminated at a given nonrandom time, T, the corresponding likelihood function is

$$L_I = \prod_{i=1}^{r} f(t_i)(S(T))^{N-r} \tag{49}$$

Type II Censoring

A test can also be terminated when a previously specified nonrandom number of failures (say r), have been observed. In this case, the duration of the test is a random variable. This is known as *type II right censoring* and the individual test is sometimes called *failure terminated.* It is clear that under Type II censoring only the r smallest times to failure $t_{(1)} < t_{(2)} < \cdots < t_{(r)}$ out of sample of N times to failure are observed as distinct ones. The times to failure $t_{(i)}$ ($i = 1, 2, \ldots, r$) are considered (as in the previous case of Type I censoring) as identically distributed r.v.'s having the continuous PDF $f(t)$ and the CDF, $F(t)$. It can be shown that the joint probability density function of the times to failure $t_{(1)}, t_{(2)}, \ldots, t_{(r)}$ is given by

$$\frac{N!}{(N-r)} f(t_{(1)}) f(t_{(2)}) \cdots f(t_{(r)}) [S(t_{(r)})]^{N-r}$$

and the corresponding likelihood function obviously is

$$L_{II} = \prod_{i=1}^{r} f(t_i)(S(t_{(r)}))^{N-r} \tag{50}$$

Note that the likelihood function of Eq. (50) has the same functional form as the likelihood function of Eq. (49).

Random Censoring

The random censoring turns out to be typical in reliability data analysis when there are several failure modes that must be estimated separately. The times to failure due to each failure mode are considered, in this case, as r.v.'s having different distributions, while the object on the whole is considered as a competing risks (or series) system. A simple random censoring process usually considered in life data analysis is the situation in which each item in a sample is assumed to have a time to failure t and a censoring time L, which are continuous independent variables with PDFs $f(t)$ and $g(L)$, and CDFs $F(t)$ and $G(L)$.

Designate the reliability (survivor) functions corresponding to the CDFs $F(t)$ and $G(L)$ by S_F and S_G. Let our data be represented by the same pairs of r.v.'s, (τ_i, δ_i), $i = 1, 2, \ldots, n$, as in the case of Type I censoring. It can be shown that the likelihood function for these data is given by Lawless (7):

$$\prod_{i=1}^{n} [f(\tau_i) S_G(\tau_i)]^{\delta_i} [g(\tau_i) S_F(\tau_i)]^{1-\delta_i} = \left[\prod_{i=1}^{n} S_G(\tau_i)^{\delta_i} g(\tau_i)^{1-\delta_i}\right] \left[\prod_{i=1}^{n} f(\tau_i)^{\delta_i} S_F(\tau_i)^{1-\delta_i}\right]$$

If we are not interested in estimation of the censoring time distribution, the above function is reduced to

$$L_{RC} = \prod_{i=1}^{n} f(\tau_i)^{\delta_i} S(\tau_i)^{1-\delta_i}$$

which has exactly the same form as in the case Type I censoring [Eq. (48)]. From a practical point, the random censoring is usually combined with Type I censoring because of, for example, limited test or observation time. In this case, if the matter is time-to-failure distribution, all the censoring can be expressed also in the framework of the Type I case.

PARAMETRIC DISTRIBUTION ESTIMATION

In this section we consider the estimation of some time-to-failure distribution based on maximum likelihood approach.

Exponential Distribution

The exponential distribution, historically, was the first life distribution model for which statistical methods were extensively developed (7). It is still the most important component reliability model for complex system reliability estimation.

Type II Censored Data

Rewrite the PDF of the exponential distribution of (4) in the form:

$$f(t,\theta) = \frac{1}{\theta}\, e^{-t/\theta}, \quad t \geq 0 \tag{51}$$

Under Type II right censoring only the r smallest times to failure $t_{(1)} < t_{(2)} < \ldots < t_{(r)}$ (order statistics) out of sample of n times to failure are observed as distinct ones. Using the corresponding likelihood function of Eq. (50) for the exponential distribution considered, we can write the likelihood function as

$$L_{II} = \prod_{i=1}^{r} f(t_i)\big(S(t_{(r)})\big)^{n-r} = \frac{1}{\theta^r}\, e^{-T_{II}/\theta} \tag{52}$$

where

$$T_{II} = \sum_{i=1}^{r} t_{(i)} + (n-r)t_{(r)}$$

It is easy to show that

$$\hat{\theta} = \frac{T_{II}}{r} \tag{53}$$

is the maximum likelihood estimate (MLE) for the case considered. It can be shown that $2T_{II}/\theta$ has the chi-square distribution with $2r$ degrees of freedom. Using this fact we can construct a confidence interval. Because the uncensored data are the particular case of the Type II right censored data (when $r = n$), we need to consider only the Type II case in order to treat complete samples. Using the distribution of $2T_{II}/\theta$, one can write

$$Pr\left(\chi^2_{\alpha/2}(2r) \leq \frac{2T}{\theta} \leq \chi^2_{1-\alpha/2}(2r)\right) = 1-\alpha \tag{54}$$

where $\chi^2_{\beta}(2k)$ is the βth quantile (100βth percentile) of the chi-square distribution with $2k$ degrees of freedom. The relationship of Eq. (54) gives the following two-sided confidence interval for θ:

$$\frac{2T_{II}}{\chi^2_{1-\alpha/2}(2r)} \leq \theta \leq \frac{2T_{II}}{\chi^2_{\alpha/2}(2r)} \tag{55}$$

Having the point and interval (confidence) estimates for the MTTF, θ, it is easy to construct the estimates for other reliability measures for the exponential distribution. For example, the point estimate of the reliability (survivor) function is

$$\hat{R}(t) = e^{-t/\hat{\theta}}, \quad t \geq 0 \tag{56}$$

and the upper $(1-\alpha)$ confidence limit is given by

$$R_u(t) = e^{-t/\theta_l}, \quad t \geq 0 \tag{57}$$

where θ_l is the $(1-\alpha)$ lower confidence limit for θ.

Example 5 A sample of 20 identical items was placed on a life test. The test was terminated just after the fourth failure had been observed. The times to failure (in hours) recorded are 322, 612, 685, and 775. Assuming that the TTF distribution is exponential, find the lower 90% confidence limit for the MTTF, θ.

Calculate the total time on test, T_{II}:

$$T_{II} = \sum_{i=1}^{4} t_{(i)} + (20-4)t_{(4)} = 2394 + 16\,X\,775 = 14794.$$

Using Eq. (55) with $\alpha = 0.1$ find the lower limit of interest as

$$\theta_l = \frac{2T_{II}}{\chi^2_{0.9}(8)} = \frac{29588}{13.36} = 2214.7\ \text{h}$$

Type I Censoring without Replacement

Under Type I right censoring without replacement, a test is terminated at a given nonrandom time, T, and the number of failures, r, observed during the test period is random. Recalling the corresponding likelihood function of Eq. (49), we can write the respective likelihood function as

$$L_I = \prod_{i=1}^{r} f(t_i)(S(T))^{n-r} = \frac{1}{\theta^r}\, e^{-T_I/\theta} \tag{58}$$

where

$$T_I = \sum_{i=1}^{r} t_{(i)} + (n-r)T$$

Similarly to the previous case the MLE of θ is given by

$$\hat{\theta} = \frac{T_I}{r} \tag{59}$$

The estimate of Eq. (59) can be generalized for the case of multiple nonrandom right censoring. Let $t_{c1}, t_{c2}, \ldots, t_{c(N-r)}$ be nonrandom times to censoring. In the case considered, the MLE of θ can be obtained replacing T_I by

$$T_{Imc} = \sum_{i=1}^{r} t_{(i)} + \sum_{i=1}^{N-r} t_{ci}$$

On the one hand, the estimate of Eq. (59) looks similar to the estimate of Eq. (53). On the other hand, in the case of Type I censoring, the number of failures observed, r, is random, so that T_I and r are considered as joint sufficient statistics for a single parameter, θ, (8), which results in the absence of an exact confidence estimation for the situation given.

The most widely used practical approach (approximation) to the confidence estimation for the Type I censoring is based on the assumption that the quantity $2T_I/\theta$ has the chi-square distribution with $2r + 1$ degrees of freedom, which results in the following two-sided confidence interval for θ:

$$\frac{2T_I}{\chi^2_{1-\alpha/2}(2r+1)} \leq \theta \leq \frac{2T_I}{\chi^2_{\alpha/2}(2r+1)} \tag{60}$$

Type I Censoring with Replacement

Consider a situation when n units are placed on test and each failed item is replaced instantly upon a failure. The relation between the exponential and the Poisson distribution was mentioned earlier. Let's now consider this relationship a little more closely. Let T be a fixed test duration, and let the number of failures, N, for a unit during this time have the Poisson distribution with an intensity rate, λ, that is,

$$Pr(N|\lambda t) = \frac{(\lambda T)^N e^{-\lambda T}}{N!}$$

Consider a time interval (0, t], where $t < T$. The probability that TTF $> t$ [which is the reliability function $R(t)$] is the probability that no failure occurs in the interval (0, t], so, it is given by the above formula with $N = 0$ and $T = t$. Thus, $R(t) = e^{-\lambda 1}$, which is the reliability function of the exponential distribution. Using the general form of likelihood function for Type I censoring in Eq. (48), one can find the following likelihood function for the case considered:

$$L_{Ir} = \frac{1}{\theta^r} e^{-\sum t_i/\theta} \tag{61}$$

where r is the observed number of failures during the test duration T, and $\sum t_i$ is the total time on test. It is clear that $\sum t_i = nT$, so that r is sufficient for θ. It is also obvious that r has the Poisson distribution with the mean equal to nT/θ. Finally, using Eq. (61) the MLE of θ can be written as

$$\hat{\theta} = \frac{nT}{r} \tag{62}$$

Using the Poisson distribution of the number of failures, r, and the relationship between the chi-square and Poisson distributions, the following two-sided confidence interval for θ can be obtained in terms of chi-square distribution as

$$\frac{2nT}{\chi^2_{1-\alpha/2}(2r+2)} \le \theta \le \frac{2nT}{\chi^2_{\alpha/2}(2r)} \tag{63}$$

The corresponding hypothesis testing is considered in Lawless (7).

Type II Censoring with Replacement

This case can be reduced to the corresponding case without replacement if the total time on test T_{II} is replaced by $nt_{(r)}$.

Weibull Distribution

Let's consider the right censored data, for example, a test of n units that results in r distinct times to failure $t_{(1)} < t_{(2)} < \cdots < t_{(r)}$ and $(n - r)$ times to censoring $t_{c1}, t_{c2}, \ldots, t_{c(n-r)}$. Using the likelihood function in the form

$$L = \prod_{i=1}^{r} f(t_{(i)}) \prod_{j=1}^{n-r} R(t_{cj})$$

one can write the corresponding log-likelihood function for the Weibull distribution with the scale parameter α and the shape parameter β, as

$$\log L(\alpha, \beta) = r\log\beta - \beta r \log\alpha + (\beta - 1)\sum_{i=1}^{r} \log t_{(i)} - \alpha^{-\beta} T \tag{64}$$

where

$$T = \sum_{i=1}^{r} t_{(i)}^{\beta} + \sum_{i=1}^{n-r} t_{ci}^{\beta}$$

The ML estimates of the parameters α and β can be found as a straightforward solution of the maximization (of the likelihood function) problem under the restrictions $\alpha > 0$ and $\beta > 0$, or using the first-order conditions, which result in solution of a system of two nonlinear equations. In any case, using a numerical method is a must.

Example 6 A sample of 10 identical components was placed on a life test. The times to failure (in hours) recorded are 459, 462, 780, 1408, 1735, 1771, 1967, 2105, 2860, and 3441. Assuming that the component TTF distribution is Weibull, find the point estimates of the distribution parameters α and β.

Using an appropriate numerical procedure for maximizing likelihood function [Eq. (64)] with respect to α and β, find the following estimates: $\hat{\beta} = 1.91$ and $\hat{\alpha} = 1916.17$ h. To get a feeling about the accuracy of the estimates obtained, mention that the data were generated from the Weibull distribution with $\hat{\beta} = 2$ and $\hat{\alpha} = 2000$ h.

NONPARAMETRIC DISTRIBUTION ESTIMATION

The estimation and hypothesis testing procedures previously discussed involved special assumptions. For instance, it could be assumed that TTF has the exponential, or Weibull, distribution, and it was necessary to test the goodness of fit to verify the assumption. However, even though the goodness of fit is validated by hypothesis testing, the hypothesis remains a hypothesis. There are, also, special statistical methods that do not require knowledge of the underlying distribution. In some situations, it is enough to assume that a sample belongs to the class of all continuous or discrete distributions. The statistical procedures based on such assumptions are known as *nonparametric* or *distribution-free* procedures. The nonparametric procedures used in reliability and life data analysis are also constructed for the special classes of distribution functions related to concepts of aging discussed previously.

Cumulative Distribution Function and Reliability Function Estimation

Any random variable is completely described by its CDF, so the problem of CDF estimation is of great importance. The estimate of CDF is the *empirical* (or *sample*) distribution function (EDF) for uncensored data was given by Eq. (45). The respective estimate of the reliability (survivor) function is called the *empirical* (or *sample*) reliability function (ERF). It can be written for a sample of size n as

$$R_n(t) = \begin{cases} 1 & 0 < t < t_{(1)} \\ 1 - \dfrac{i}{n} & t_{(i)} \le t < t_{(i+1)} \\ 0 & t_{(n)} \le t < \infty \end{cases} \qquad i = 1, \ldots, n-1 \tag{65}$$

where $t_{(1)}, t_{(2)}, \ldots, t_{(n)}$ are the ordered sample data (order statistics). The construction of an EDF requires a complete sample. The EDF can be obtained also for the right censored samples for the times less than the last TTF observed ($t < t_{(r)}$). The empirical distribution function is a random function, since it depends on the sample units. For any given point t, the EDF, $S_n(t)$, is the fraction of sample items failed before t.

Thus, the EDF is the estimate of the probability of a success (in this context, "success" means "failure"), p, in a Bernoulli trial, and this probability is $p = F(t)$. Note that the maximum likelihood estimator of the binomial parameter p (see Example 2) coincides with $S_n(t)$. It can be shown that EDF $S_n(t)$ is a consistent estimator of the CDF, $F(t)$. It is clear that the mean number of failures observed during time t is $E(r) = pn = F(t)n$, so that the mean value of the fraction of sample items failed before t is $E(r/n) = p = F(t)$ and the variance of this fraction is given by

$$\mathrm{Var}\left(\frac{r}{n}\right) = \frac{p(1-p)}{n} = \frac{F(t)(1-F(t))}{n} \tag{66}$$

For some practical problems in which the estimate of the variance (66) is required, the formula above is used with replacement $F(t)$ by $S(t)$. For example, it is known that as the sample size, n, increases, the binomial distribution can be approximated by a normal distribution with the same mean and variance ($\mu = np$, $\sigma^2 = np(1 - p)$), which gives reasonable results if np and $n(1 - p)$ are both ≥ 5. Basing on this approximation, the following approximate $100(1 - \alpha)\%$ confidence interval for the unknown CDF, $F(t)$, at any given point t can be constructed:

$$S_n(t) - z_{\alpha/2}\left(\frac{S_n(t)(1-S_n(t))}{n}\right)^{1/2} \leq F(t) \leq S_n(t) + z_{\alpha/2}\left(\frac{S_n(t)(1-S_n(t))}{n}\right)^{1/2} \tag{67}$$

where z_α is the quantile of level α of the standard [N(0,1)] normal distribution.

The corresponding estimates for the reliability (survivor) function can be obtained using the obvious relationship $R_n(t) = 1 - S_n(t)$.

Confidence Intervals for Unknown Cumulative Distribution Function

Using a complete or right censored sample from an unknown CDF, one can get the strict confidence intervals for the unknown CDF, $F(t)$. This can be done using the Clopper-Pearson procedure for constructing the confidence intervals for a binomial parameter p: The lower confidence limit, $F_l(t)$, at the point t where $S_n(t) = r/n$ ($r = 0, 1, 2, \ldots, n$), is the largest value of p that satisfies the equation:

$$I_p(r, n-r+1) \leq \alpha/2 \tag{68}$$

and the respective upper confidence limit, $F_u(t)$, is the smallest p that satisfies the equation:

$$I_{1-p}(n-r, r+1) \leq \alpha/2 \tag{69}$$

where $I_x(a, b)$ is the incomplete beta function given by

$$I_p(a,b) = \frac{\Gamma(\alpha+\beta)^p}{\Gamma(\alpha)\Gamma(\beta)} \int_{0^{x^{\alpha-1}}}^{p} (1-x)^{\beta-1}\,dx, \quad 0 \leq p \leq 1,\ \alpha > 0 \tag{70}$$

Kaplan-Meier (Product-Limit) Estimate. The point and confidence estimation considered are not applicable to multiply censored data. For such samples, the product-limit estimate, which is the MLE of the CDF, can be applied.

Assume a sample of n items, among which only k failure times are known exactly. Denote these ordered times as: $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(k)}$, and let $t_{(0)}$ be identically equal to zero, $t_{(0)} \equiv 0$. Denote by n_j the number of items under observation just before $t_{(j)}$. Assume that the CDF is continuous, so that there is only one failure at every $t_{(i)}$. Then, $n_{j+1} = n_j - 1$. Under these conditions, the product limit estimate is given by:

$$S_n(t) = 1 - R_n(t) = \begin{cases} 0 & 0 \leq t < t_{(1)} \\ 1 - \prod_{j=1}^{i} \dfrac{n_j - 1}{n_j} & t_{(i)} \leq t < t_{(i+1)},\ i = 1, \ldots, \\ 1 & t \geq t_{(E)} \end{cases} \tag{71}$$

where $E = k$, if $k < n$, and $E = n$, if $k = n$. Clearly, for uncensored (complete) samples, the product limit estimate coincides with the EDF [Eq. (45)]. In the general case (including descrete distribution, censored or grouped data), the Kaplan-Meier estimate is given by

$$S_n(t) = 1 - R_n(t) = \begin{cases} 0 & 0 \leq t < t_{(1)} \\ 1 - \prod_{j=1}^{i} \dfrac{n_j - d_j}{n_j} & t_{(j)} \leq t < t_{(j+1)},\ i = 1, \ldots, E \\ 1 & t \geq t_{(E)} \end{cases} \tag{72}$$

where d_j is the number of failures at $t_{(j)}$. For estimation of variance of S_n (or R_n), Greenwood's formula is used:

$$\hat{\mathrm{Var}}[S_n(t)] = \hat{\mathrm{Var}}[R_n(t)] = \sum_{j: t_{(j)} < t} \frac{d_j}{n_j(n_j - d_j)} \tag{73}$$

Example 7 Table 2 is a typical example of censored data presentation (the left columns) suitable for different statistical software tools. The EDF values estimated using Eq. (72) are given in the right column.

Percentile Life Estimation for Continuous Distributions

For reliability applications, the lower confidence limit of the $100p$th percentile (or the quantile of level p) of time to failure

Table 2. Failure Time Sample and Respective Product-Limit Estimate

i	Ordered Failure or Censoring Times, $t_{(i)}$ or $t^*_{(i)}$	$S_n(t_{(i)})$
0	0	0
1	707	0.067
2	728	0.133
3	950*	0.133
4	972	0.206
5	1017	0.278
6	1100*	0.278
7	1260	0.358
8	1494	0.438
9	1500*	0.438
10	1586	0.532
11	1697	0.626
12	1742	0.719
13	1794	0.813
14	1968	0.906
15	2000*	0.906

is one of the most important interval estimates. The random variable $T(\alpha, p)$ is called the lower limit of the 100pth percentile t_p, which corresponds to the confidence probability, γ ($\gamma = 1 - \alpha$), if these quantities satisfy the relationship

$$Pr\left\{\int_{T(\gamma,p)}^{\infty} dF \geq 1-p\right\} = \gamma$$

where $F(t)$ is the CDF of TTF. Let $t_{(r)}$ be the TTF of the rth failure obtained from a sample of size n from $F(t)$. The TTF $t_{(r)}$ is the lower γ-confidence limit of the 100pth percentile, t_p, if its number, r, satisfies the inequality:

$$I_p(r, n-r+1) \geq \gamma \tag{74}$$

where $I_p(a, b)$ is the above-mentioned incomplete beta function, Eq. (70).

It should be noted that for a given γ and a given p, this confidence limit does not exist for any value of sample size n. For a given γ and p, there is a minimum necessary sample size $n_m(p, \gamma)$, for which $t_{(1)}$ (the time moment of occurrence of the first failure) is the lower γ-confidence limit of the percentile t_p; in other words, $n_m(p, \gamma)$ is a solution of Eq. (74) with respect to n, when $r = 1$.

The procedure for constructing a lower γ confidence limit of the 100pth percentile t_p does not require very large sample sizes. For example, for $\gamma = (1 - p) = 0.9$, $n_m(0.1, 0.9) = 22$.

Percentile Life Estimation for Aging Distributions. When constructing confidence limits in the class of continuous distributions, a basic limitation of the procedure is the size of the minimum necessary sample, n_m. This limitation has stimulated interest in obtaining a solution for the narrower reliability class of aging, that is, for IFR distributions. The lower γ-confidence limit of the 100pth percentile, t_p, for IFR distribution $t_p(\gamma, p, r)$ is given by Barlow and Proschan (9):

$$t(\gamma, p, r) = T_s(t_{(r)}) \cdot \min\left(\frac{2\ln(1/(1-p))}{\chi^2_\gamma(2r)}, \frac{1}{n}\right) \tag{75}$$

where

$$T_s(t_{(r)}) = \sum_{i=1}^{r} t_{(i)} + (n-r)t_{(r)}$$

is the total time on test. This formula gives the lower confidence limit for any sample size. It should be mentioned that if (for a given n, p, and γ) $t_{(r)}$ is the confidence limit for the class of continuous distribution, it always has a larger mean value than the mean of the limit given by the IFR procedure.

BIBLIOGRAPHY

1. R. E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing: Probability Model,* Silver Spring, MD: To Begin With, 1981.
2. A. M. Abouammoh, On the criteria of the mean remaining life, *Stat. and Prob. Lett.,* **6** (4): 205–211, 1988.
3. N. Mann et al., *Methods for Statistical Analysis of Reliability and Life Data,* New York: Wiley, 1974.
4. J. W. Provan, Probabilistic Approaches to the Material-Related Reliability of Fracture-Sensitive Structures, in J. W. Provan (ed.), *Probabilistic Fracture Mechanics and Reliability,* Dordrecht: Martinus Nijhoff, 1987.
5. B. V. Gnedenko and I. A. Ushakov, *Probabilistic Reliability Engineering,* New York: Wiley, 1995.
6. D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events,* New York: Wiley, 1968.
7. J. F. Lawless, *Statistical Models and Methods for Lifetime Data,* New York: Wiley, 1982.
8. L. J. Bain and M. Engelhardt, *Statistical Analysis of Reliability and Life-testing Models,* New York: Marcel Dekker, 1991.
9. R. E. Barlow and F. Proschan, Tolerance and confidence limits for classes of distributions based on failure rates, *Ann. Math. Stat.,* **37** (6): 1966.

Reading List

H. Martz and R. Waller, *Bayesian Reliability Analysis,* New York: Wiley, 1982.

W. Nelson, *Applied Life Data Analysis,* New York: Wiley, 1982.

MARK KAMINSKIY
University of Maryland