The Logistic–Exponential Survival Distribution

Yingjie Lan, Lawrence M. Leemis

Department of Mathematics, The College of William & Mary, Williamsburg, Virginia 23187–8795

Abstract: For various parameter combinations, the logistic–exponential survival distribution belongs to four common classes of survival distributions: increasing failure rate, decreasing failure rate, bathtub-shaped failure rate, and upside-down bathtub-shaped failure rate. Graphical comparison of this new distribution with other common survival distributions is seen in a plot of the skewness versus the coefficient of variation. The distribution can be used as a survival model or as a device to determine the distribution class from which a particular data set is drawn. As the three-parameter version is less mathematically tractable, our major results concern the two-parameter version. Boundaries for the maximum likelihood estimators of the parameters are derived in this article. Also, a fixed-point method to find the maximum likelihood estimators for complete and censored data sets has been developed. The two-parameter and the three-parameter versions of the logistic–exponential distribution are applied to two real-life data sets. © 2008 Wiley Periodicals, Inc. Naval Research Logistics 55: 252–264, 2008

Keywords: censored data; maximum likelihood estimation; reliability; survival model

1. INTRODUCTION

A unifying survival model that encompasses the four classes of IFR (Increasing Failure Rate), DFR (Decreasing Failure Rate), BT (Bathtub-Shaped Failure Rate), and UBT (Upside-Down Bathtub-Shaped Failure Rate) in a single model would be useful in survival analysis. Such a model would provide considerable flexibility for fitting a wide variety of lifetime data sets. Such a model might also be used to determine the distribution class from which the data is drawn, by establishing confidence regions over its parameters. The distribution presented here, known as the logistic–exponential distribution, satisfies these criteria.

There are several approaches for developing flexible survival models. First, one can create *systems* of distributions, such as the Burr, Johnson, and Pearson systems. These systems are described in Johnson et al. [11, pp. 15–63], along with references to original and subsequent incremental work. Second, competing risks models (David and Moeschberger [7], Crowder [5], Pintlie [16]) can be used to combine popular parametric models by assuming that the lifetime of interest is the minimum of the lifetimes associated with several risks that are competing for the lifetime of the item. This approach can be generalized from a series arrangement of the risks to any arrangement of the risks that is consistent with the arrangement of components in a coherent system. Third, finite mixtures (Everitt and Hand [8]; McLachlan and Peel [13]) can be used to blend several parametric models into a single, more flexible model. Fourth, a new parametric model, such as the IDB (Increasing, Decreasing, Bathtub) model suggested by Hjorth [10] that can belong to the IFR, DFR, and BT distribution classes, can be formulated, and probabilistic and statistical properties of the model can be derived. It is the fourth approach that is taken in this article.

Let T be a positive random variable. Consider first the survivor function for the two-parameter case with a positive shape parameter κ and a positive scale parameter λ:

$$S(t) = P(T \geq t) = \frac{1}{1 + (e^{\lambda t} - 1)^{\kappa}} \quad t \geq 0.$$

This survival function resembles the log logistic survival function with the second term of the denominator being changed in its base to an exponential function, which is why we call it "logistic–exponential."[1] The probability density

---

[1] The survivor function for the log logistic distribution is $S(t) = (1 + (\lambda t))^{-\kappa}$ for $t \geq 0$. Tadikamalla and Johnson [17] describe transformations of a standard log logistic random variable useful for modeling. This survivor function also bears some resemblance to the one-parameter Burr Type II distribution given in Burr [2] which has survivor function $S(t) = 1 - (1 + e^{-t})^{-\kappa}$ for $-\infty < t < \infty$ and the exponentiated Weibull distribution [15] which has survivor function $S(t) = 1 - (1 - e^{-(\lambda t)^{\kappa}})^{\theta}$ for $t \geq 0$.

Correspondence to: Lawrence Leemis (leemis@math.wm.edu)

function for the logistic–exponential distribution is

$$f(t) = -S'(t) = \frac{\lambda\kappa(e^{\lambda t} - 1)^{\kappa-1}e^{\lambda t}}{(1 + (e^{\lambda t} - 1)^\kappa)^2} \quad t \geq 0$$

and the hazard function is

$$h(t) = \frac{f(t)}{S(t)} = \frac{\lambda\kappa e^{\lambda t}(e^{\lambda t} - 1)^{\kappa-1}}{1 + (e^{\lambda t} - 1)^\kappa} \quad t \geq 0,$$

which is used to determine the distribution classes that can be assumed by this parametric model. The distribution reduces to an exponential distribution when $\kappa = 1$, which belongs to both IFR and DFR classes. The distribution is in the BT class when $0 < \kappa < 1$, and in the UBT class when $\kappa > 1$. Interestingly enough, in both the BT and UBT cases, the hazard function achieves its extreme point (minimum in the BT case and maximum in the UBT case) at $t = \ln(x(\kappa) + 1)/\lambda$, where $x(\kappa)$ is the sole positive root of $\kappa x - x^\kappa = 1 - \kappa$. More details concerning this root are given in Section 2. To the best of our knowledge, this survival distribution is unique in that it is the only two-parameter distribution that includes all four distribution classes as special cases.

As shown subsequently, this two-parameter case might not provide enough flexibility for modeling empirical data. Thus, we introduce a third parameter $\theta \geq 0$ (which shifts the distribution to the left, followed by a truncation at zero and a rescaling), making it a full-fledged unifier of the four distribution classes.[2] The survivor function is

$$S(t) = \frac{1 + (e^{\lambda\theta} - 1)^\kappa}{1 + (e^{\lambda(t+\theta)} - 1)^\kappa} \quad t \geq 0; \lambda > 0, \kappa > 0, \theta \geq 0.$$

The third parameter shifts the two-parameter hazard curve to the left, providing a wider variety of models in the IFR, DFR, BT, and UBT classes. Figure 1 illustrates the various cases. The graph shows the classes of logistic–exponential distributions by their three parameters: the horizontal axis is κ, and the vertical axis is $\lambda\theta$; the first quadrant is divided into four regions by a vertical line at $\kappa = 1$ and a monotonically decreasing curve, with each region corresponding to the class as labeled in the graph. The equation of the monotone decreasing curve is $\lambda\theta = \ln(x(\kappa) + 1)$, thus is totally determined by κ, and has a vertical asymptote at zero.

The derivation of the decreasing curve in Fig. 1 is as follows. Since the third parameter θ results in a shifting of the

---

[2] Although the two-parameter case includes all four of the distribution classes, the IFR and DFR classes are restricted to a single value of the shape parameter, $\kappa = 1$, corresponding to a constant hazard function. Therefore a modeler should not use the two-parameter logistic–exponential distribution, for example, to fit a data set known to come from a population with a monotone increasing hazard function. The three-parameter model would be an appropriate candidate model in that case.
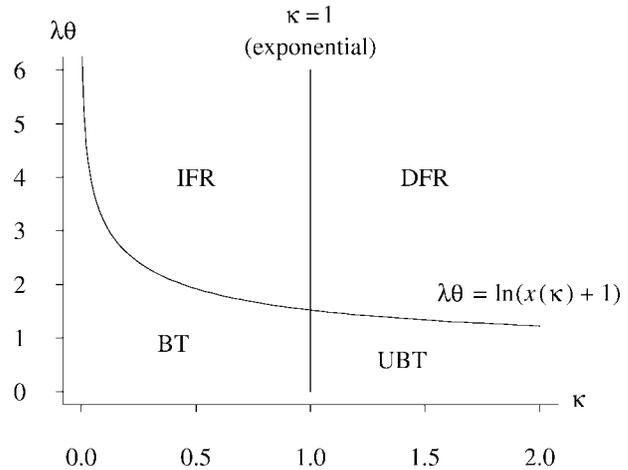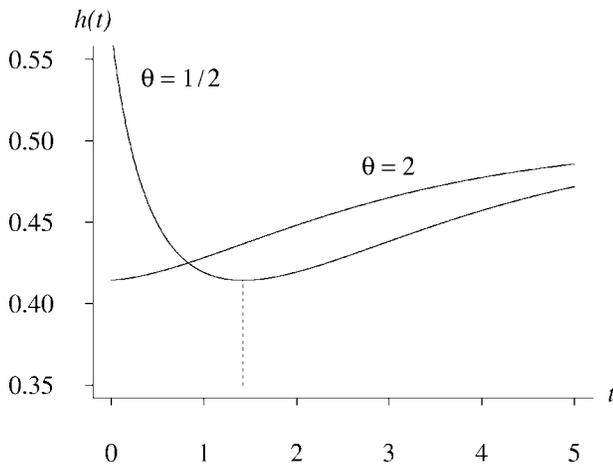

Figure 1. Distribution classes for the three-parameter logistic–exponential distribution illustrated in the parameter space.

hazard function from the two-parameter case, once it passes the extreme point of the hazard function, the UBT shape will only have its decreasing tail remaining, thus becoming a DFR shape, and similarly BT will become an IFR. So consider the hazard function of the two-parameter model that is the special case when the third parameter θ is zero. All that needs to be done is to find out the critical point of the derivative of the hazard function:

$$h'(t) = \frac{\lambda^2\kappa(e^{\lambda t} - 1)^{\kappa-2}(\kappa e^{\lambda t} - 1)e^{\lambda t}}{1 + (e^{\lambda t} - 1)^\kappa}$$
$$- \left(\frac{\lambda\kappa(e^{\lambda t} - 1)^{\kappa-1}e^{\lambda t}}{1 + (e^{\lambda t} - 1)^\kappa}\right)^2 = 0.$$

As expected, there is a unique positive root for the derivative of the hazard function when $\kappa \neq 1$, and that root is found to be $t_c = \ln(x(\kappa) + 1)/\lambda$.

2. PROBABILISTIC PROPERTIES

The logistic–exponential distribution has several useful probabilistic properties for lifetime modeling. Unlike most distributions in the BT and UBT classes, the logistic–exponential distribution enjoys closed-form density, hazard, cumulative hazard, and survival functions. The moments are finite, although they cannot be expressed in closed form. Some asymptotic results on moments are mentioned below and discussed in more detail in Appendix A.

The distribution classes for the three-parameter version of the logistic–exponential distribution and some probabilistic properties are listed below:

- The distribution is in the BT class when $0 < \kappa < 1$ and $\lambda\theta < \ln(x(\kappa) + 1)$. The minimum of the hazard

Figure 2. Hazard functions in the IFR and BT classes with $\kappa = 1/2$ and $\lambda = 1$.

function occurs at $\ln(x(\kappa) + 1)/\lambda - \theta$. This case is illustrated in Fig. 2 for $\kappa = 1/2$, $\lambda = 1$, and $\theta = 1/2$. The minimal hazard is achieved at $t_c - \theta \cong 1.421$.

- The distribution is in the IFR class when $0 < \kappa < 1$ and $\lambda\theta \geq \ln(x(\kappa) + 1)$. This case is illustrated in Fig. 2 for $\kappa = 1/2$, $\lambda = 1$, and $\theta = 2$.
- The distribution is in the UBT class when $\kappa > 1$ and $\lambda\theta < \ln(x(\kappa) + 1)$. The maximum of the hazard function occurs at $\ln(x(\kappa) + 1)/\lambda - \theta$. This case is illustrated in Fig. 3 for $\kappa = 5$, $\lambda = 1$, and $\theta = 0$ (the two-parameter case). The maximal hazard is achieved at $t_c - \theta \cong 0.9748$.
- The distribution is in the DFR class when $\kappa > 1$ and $\lambda\theta \geq \ln(x(\kappa) + 1)$. This case is illustrated in Fig. 3 for $\kappa = 5$, $\lambda = 1$, and $\theta = 1$.
- In the trivial case of $\kappa = 1$, the logistic–exponential distribution collapses to the exponential distribution, which is in both the IFR and DFR classes.
- The distribution has exponential$(\kappa\lambda)$ right-hand tails, i.e., $\lim_{t \to \infty} h(t) = \lambda\kappa$. (The exponential distribution has a constant hazard function.) All hazard functions have horizontal asymptotes at $\lambda\kappa$, which implies that the distribution behaves like an exponential distribution for large t. The only other widely-used survival model with exponential tails is the gamma distribution. On the other hand, when t approaches zero, $e^{\lambda t} - 1 \approx \lambda t$, thus the distribution behaves like a log logistic distribution around $t = 0$. These properties are incorporated into our name for the new distribution.
- The pth fractile of the distribution is:

$$t_p = \frac{1}{\lambda} \ln\left(1 + \left(\frac{(e^{\lambda\theta} - 1)^\kappa + p}{1 - p}\right)^{1/\kappa}\right) - \theta.$$

- Random variates can be generated via inversion by:

$$T \leftarrow \frac{1}{\lambda} \ln\left(1 + \left(\frac{(e^{\lambda\theta} - 1)^\kappa + U}{1 - U}\right)^{1/\kappa}\right) - \theta,$$

where $U \sim U(0, 1)$, yielding an efficient, synchronized, and monotone variate generation algorithm.
- When $0 < \kappa < 1$ and $\theta = 0$, both the probability density function and the hazard function have a vertical asymptote at $t = 0$.
- All moments exist for this distribution, due to its exponential right-hand tail, although moments that can be expressed in closed-form might not be available. Some asymptotic results exist, however, for the two-parameter case:

$$\lim_{\kappa \to \infty} E[T^n] = \left(\frac{\ln 2}{\lambda}\right)^n,$$

$$\lim_{\kappa \downarrow 0} (\lambda\kappa)^n E[T^n] = \int_0^\infty \frac{nz^{n-1}}{e^z + 1} dz,$$

for any positive integer n. These results will be used to construct a plot described next. See Appendix A for more detail on the asymptotic properties of the moments.
- To compare this new distribution with other commonly-known distributions, the chart from Cox and Oakes [4, p. 27] and Meeker and Escobar [14, p. 110] for the two-parameter logistic–exponential is drawn in Fig. 4 for the log normal, log logistic, Weibull, and gamma distributions. The horizontal axis is the coefficient of variation $\gamma_1 = \sigma/\mu$ and the vertical axis is the skewness $\gamma_2 = E[(T - \mu)^3]/\sigma^3$. The
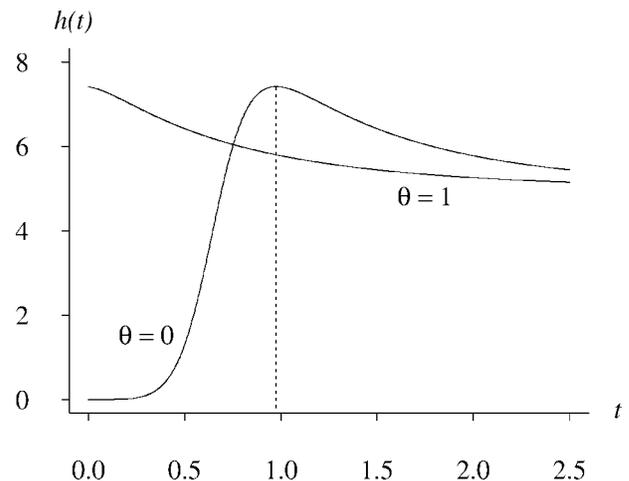

Figure 3. Hazard functions in the DFR and UBT classes with $\kappa = 5$ and $\lambda = 1$.
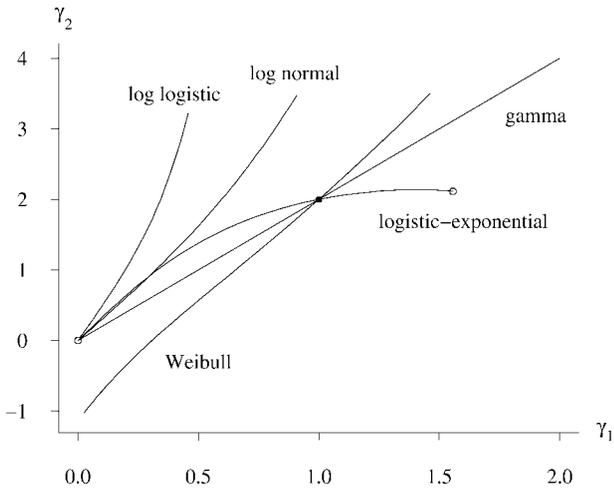
Figure 4. Coefficient of variation γ_1 versus skewness γ_2 for some two-parameter survival models.

curve for the two-parameter logistic–exponential has a few unique characteristics relative to the others plotted in the figure. First, it is the only curve that is bounded. Second, it is the only curve that achieves a maximum in the (γ_1, γ_2) plane, which occurs at $\gamma_1 \cong 1.39$. Beginning at $(0, 0)$ on the graph (the case as $\kappa \to \infty$), the curve closely approximates that of the UBT log normal distribution, although their concavities differ. After passing through the exponential case of $(1, 2)$, which is the IFR/DFR case, the BT region of the curve continues into areas of the (γ_1, γ_2) plane not covered by the other distributions. The curve achieves its maximum of $\gamma_2 \cong 2.14$ at $\gamma_1 \cong 1.39$, then declines and finally ends at $(\gamma_1, \gamma_2) \cong (1.5568, 2.1126)$, which is the limiting case as $\kappa \downarrow 0$. These interesting features of the curve clearly distinguish the logistic–exponential distribution from the other popular survival distributions plotted in Fig. 4. The fact that the curve associated with the logistic–exponential distribution covers territory that is distinct from the other distributions is useful for modeling. Plotting $(\hat{\gamma}_1, \hat{\gamma}_2)$ in Fig. 4 for a particular data set can be used for model discrimination.

- In a similar fashion, the skewness γ_2 can be plotted on the horizontal axis versus the kurtosis $\gamma_3 = E[(T - \mu)^4]/\sigma^4$ on the vertical axis for the same two-parameter survival distributions, as shown in Fig. 5. Beginning at $(0, 4.2)$ on the graph (the case as $\kappa \to \infty$), the curve associated with the UBT class proceeds to the $\kappa = 1$ exponential special case (IFR/DFR) at $(2, 9)$ and ends in the BT class (the case as $\kappa \downarrow 0$) at $(\gamma_2, \gamma_3) \cong (2.1126, 8.6876)$ (see Appendix A.2 for details). Since lower-order moments are more significant, we recommend the use of the (γ_1, γ_2) graph given

in Fig. 4 for model discrimination. Figure 5 is plotted upside down "in accordance to well-established convention" [11, p. 23].

3. STATISTICAL INFERENCE

We fit two data sets that could have come from different classes of distributions. Both complete and right-censored data sets will be considered. We first present a complete data set using a reliability data set that has been suspected to be in the UBT class, then a right-censored data set from survival analysis illustrates the BT case. Comparison with other distributions will be provided when possible. Numerical methods are required to determine the maximum likelihood estimators (MLEs), which is typical of most two-parameter lifetime models.

3.1. Complete Data Sets: Two-Parameter Model

We begin with the case of fitting the two-parameter model to uncensored datasets. Let $t_1, t_2, \ldots, t_n$ denote the failure times. The likelihood function for the two-parameter model is

$$L(\kappa, \lambda) = \prod_{i=1}^{n} f(t_i, \kappa, \lambda) = \prod_{i=1}^{n} \frac{\lambda \kappa (e^{\lambda t_i} - 1)^{\kappa-1} e^{\lambda t_i}}{(1 + (e^{\lambda t_i} - 1)^\kappa)^2}.$$

The log likelihood function is

$$\ln L(\kappa, \lambda) = \sum_{i=1}^{n} (\ln \lambda + \ln \kappa + (\kappa - 1) \ln(e^{\lambda t_i} - 1) + \lambda t_i - 2 \ln(1 + (e^{\lambda t_i} - 1)^\kappa)$$
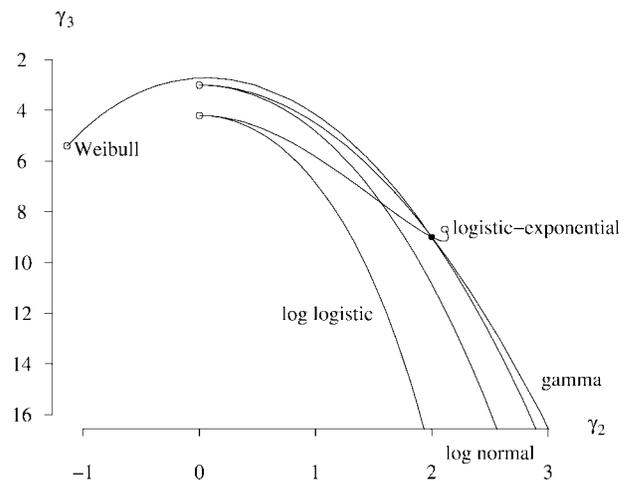

Figure 5. Skewness γ_2 versus kurtosis γ_3 for some two-parameter survival models.

and the partial derivatives of the log likelihood function are

$$\frac{\partial \ln L(\kappa, \lambda)}{\partial \kappa} = \frac{n}{\kappa} + \sum_{i=1}^{n} \ln(e^{\lambda t_i} - 1)$$
$$- 2 \sum_{i=1}^{n} \frac{(e^{\lambda t_i} - 1)^{\kappa} \ln(e^{\lambda t_i} - 1)}{1 + (e^{\lambda t_i} - 1)^{\kappa}}$$

$$\frac{\partial \ln L(\kappa, \lambda)}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=1}^{n} \frac{(\kappa - 1) t_i e^{\lambda t_i}}{e^{\lambda t_i} - 1} + \sum_{i=1}^{n} t_i$$
$$- 2 \sum_{i=1}^{n} \frac{\kappa (e^{\lambda t_i} - 1)^{\kappa - 1} t_i e^{\lambda t_i}}{1 + (e^{\lambda t_i} - 1)^{\kappa}}.$$

Equating these partial derivatives to zero does not yield closed-form solutions for the MLEs and thus a numerical method is used for solving these equations simultaneously. The initial estimates for the numerical method can be acquired through the method of fractiles (Glen and Leemis [9]), as shown subsequently, since the moments of the distribution cannot be expressed in closed form.

The second-order differential (Hessian) matrix of the log likelihood function evaluated at the MLEs is called the *observed information matrix*, denoted as $O(\hat{\kappa}, \hat{\lambda})$, where $(\hat{\kappa}, \hat{\lambda})$ is the MLE vector. For many parametric distributions, it is a consistent estimator of the Fisher information matrix (Cox and Oakes [4]). Information matrices can be used to obtain asymptotic confidence intervals and perform hypothesis tests with respect to the MLEs. In practice, the underlying information matrices are often unavailable, and thus the observed information matrices are used instead:

$$o\left(\hat{\kappa}, \hat{\lambda}\right) = \begin{pmatrix} \frac{-\partial^2 \ln L(\kappa, \lambda)}{\partial \kappa^2} & \frac{-\partial^2 \in L(\kappa, \lambda)}{\partial \kappa \partial \lambda} \\ \frac{-\partial^2 \ln L(\kappa, \lambda)}{\partial \lambda \partial \kappa} & \frac{-\partial^2 \ln L(\kappa, \lambda)}{\partial \lambda^2} \end{pmatrix}_{\kappa = \hat{\kappa}, \lambda = \hat{\lambda}}.$$

We will use the observed information matrix to derive an asymptotic confidence interval for the MLEs of the logistic–exponential distribution.

The first empirical data set is Lieblein and Zelen's [12] $n = 23$ ball bearing failure times (each measurement in 10^6 revolutions) given in Table 1, which has been also considered more recently by Caroni [3]. For this complete data set, Crowder et al. [6, p. 63] conjectured that distributions in the UBT class may fit the ball bearing data better than distributions in the IFR class.

Table 1. The ball bearing failure times.

17.88	28.92	33.00	41.52	42.12	45.60
48.48	51.84	51.96	54.12	55.56	67.80
68.64	68.64	68.88	84.12	93.12	98.64
105.12	105.84	127.92	128.04	173.40	

Experimentation with the three-parameter model has shown that the location parameter θ does not have a statistically significant difference from zero, which indicates that fitting the two-parameter model is reasonable. Thus we conduct the analysis of this data set using the two-parameter model. The three-parameter model will be demonstrated on the second data set for BT distributions, where the location would seldom be zero due to the vertical asymptote on the density and hazard functions at $t = 0$, which is unlikely to occur in real-world survival distributions. However, in the analysis of life times of some electronic parts, a vertical asymptote at zero might be desirable.

As mentioned earlier, the initial estimates for the MLEs could be found by the method of fractiles. From the empirical survivor function for the ball bearing data set, the 5/23rd fractile is 42.12, and the 19/23rd fractile is 105.12. (Choosing fractiles near the 25th and 75th percentiles is a reasonable way to assure the numerical stability of the initial estimates.) Using the closed-form expression for t_p with $\theta = 0$, the pth fractile of the distribution is:

$$t_p = \frac{1}{\lambda} \ln \left(1 + \left(\frac{p}{1 - p} \right)^{1/\kappa} \right).$$

Thus the simultaneous solution of

$$42.12 = \frac{1}{\lambda} \ln \left(1 + \left(\frac{5}{18} \right)^{1/\kappa} \right),$$

$$105.12 = \frac{1}{\lambda} \ln \left(1 + \left(\frac{19}{4} \right)^{1/\kappa} \right),$$

yields the initial estimates. Using the Maple numerical solver fsolve() yields the following initial estimators: $\hat{\kappa}_0 = 2.2023$, $\hat{\lambda}_0 = 0.01054$. These initial values are then fed back into fsolve() to solve for the MLEs:

$$\hat{\kappa} = 2.366 \quad \text{and} \quad \hat{\lambda} = 0.01059.$$

The same MLE values are obtained using the fixed-point method we have developed for determining the MLEs. The calculation of the initial values is straightforward, inspired by the derivation of some bounds on the MLEs. Appendix B contains the derivation of the bounds on the MLEs and Appendix C contains details on the fixed-point method, as well as how the initial values are calculated. The fitted survival curve and the empirical survival function are plotted in Fig. 6, where the fitted survival curve for the Weibull distribution is also provided for the purpose of comparison.

The observed information matrix can be calculated from the MLEs for the ball bearing data set:

$$O(\hat{\kappa}, \hat{\lambda}) = \begin{pmatrix} 5.997 & 362.4 \\ 362.4 & 772500 \end{pmatrix}.$$
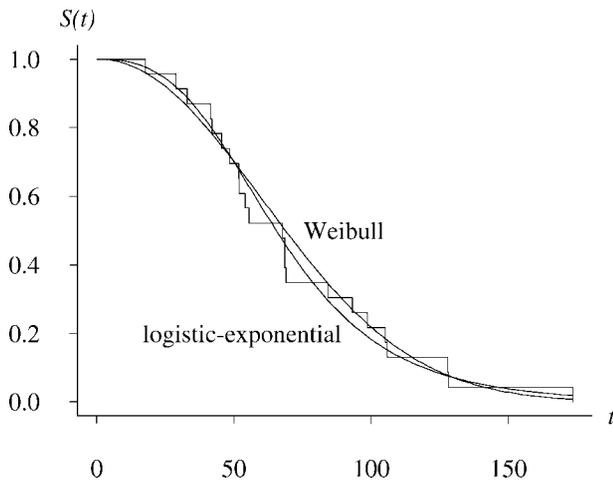
Figure 6. Fitted Weibull, fitted logistic–exponential, and empirical survivor functions for the ball bearing failure times.

Figure 7. Maximum likelihood estimator and 90%, 95%, and 98% confidence regions for the ball bearing failure times.

The inverse of this matrix is an asymptotic estimate of the variance–covariance matrix of the MLEs:

$$O^{-1}(\hat{\kappa},\hat{\lambda}) = \begin{pmatrix} 0.1716152301 & -0.0000805092 \\ -0.0000805092 & 0.0000013323 \end{pmatrix}.$$

Thus, taking the square root of those diagonal entries allows for the asymptotic confidence intervals for κ and λ. The 95% intervals in this case are:

$$1.685 < \kappa < 3.048 \quad \text{and} \quad 0.008696 < \lambda < 0.01249.$$

The asymptotic confidence interval for κ lies wholly above 1, leading to the conclusion that the data set is best fit with a UBT distribution. A more graphical approach is to use the likelihood ratio (Cox and Oakes [4]) to draw the confidence region of the parameters. In Fig. 7, the 90%, 95%, and 98% confidence regions are sketched. It is obvious that all of them lie to the right of the critical line $\kappa = 1$, which strongly supports the use of an UBT survival model. The slight negative correlation between the parameters apparent from the confidence regions is consistent with the correlation between the data values calculated to be -0.168 from the inverse of the observed information matrix.

Further study for the ball bearing data set has been carried out using various survival models to see how the new distribution performs. Table 2 gives the Kolmogorov–Smirnov (K–S) goodness-of-fit statistic D_{23} at their MLE values for several common survival models, in increasing order. The logistic–exponential distribution fell in the middle, with a goodness-of-fit statistic of 0.109, slightly larger than the Inverse Gaussian distribution. It can be observed that all UBT models fit slightly better than IFR models, further supporting Crowder et al.'s suspicion. Although the two-parameter
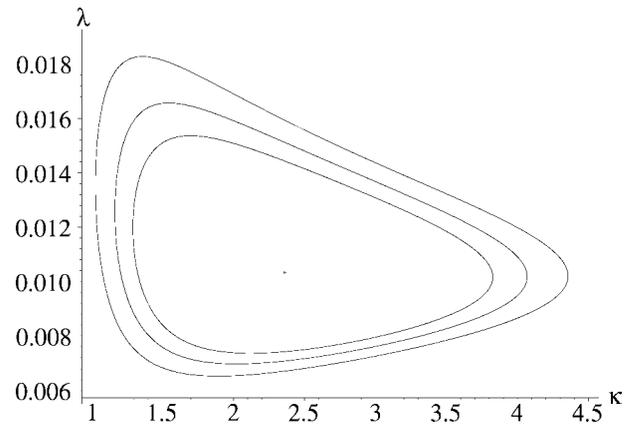
logistic–exponential distribution does not dominate the other UBT models in terms of fit for this particular data set, the territory mapped out by the logistic–exponential distribution in Fig. 4 indicates that it will dominate the others for many lifetime data sets.

It is natural to consider the parameters that minimize the K–S test statistic, which seems to be a more direct approach. Research has been carried out in this direction recently by Weber et al. [18]. Their experiments demonstrated that the effectiveness of this approach is comparable to that of the long-established MLE approach. Table 3 lists another ordering of models fitted to the same ball bearing data this way. Again, we conclude that UBT models fit better than the IFR models, adding further evidence to support Crowder et al.'s conjecture.

3.2. Right-Censored Data Sets: Three-Parameter Model

The second dataset comes from biostatistics (Burns [1]) when experiments were carried out as human subjects were placed in a cubical cabin mounted on a hydraulic piston and

Table 2. K–S goodness-of-fit statistics for the ball bearing failure times at their MLEs.

Distribution	Class	D_{23}
Log normal	UBT	0.090
Arctangent	UBT	0.093
Log logistic	UBT	0.094
Inverse Gaussian	UBT	0.099
Logistic–Exponential	UBT	0.109
Gamma	IFR	0.123
Weibull	IFR	0.152
Exponential	IFR	0.301

Table 3. K–S goodness-of-fit statistics for the ball bearing failure times at minimum K–S parameter estimates.

Model	D_{23}	Scale	Shape
Log Logistic	0.0808	0.0159	3.176
Log normal	0.0813	4.1400	0.517
Logistic–Exponential	0.0885	0.0109	2.284
Gamma	0.0895	17.16	4.032
Weibull	0.0987	0.0132	2.234
Exponential	0.2204	96.10	–

subjected to vertical motion for two hours. Table 4 gives the length of time (in minutes) until each subject first vomited. Censoring can occur in one of two ways: a subject may request an early stop, while several others survived the whole test without vomiting. Although two groups of subjects were put on test under different motion conditions, here we only used the larger group subjected to the more severe conditions. For our analysis here, we assume a random censoring scheme, which may not be reasonable due to the self-censoring present in this data set (an option exercised by the individual at six minutes).

To fit this data to the three-parameter logistic–exponential distribution, let $t_1, t_2, \ldots, t_r$ be the vomiting times and $c_1, c_2, \ldots, c_m$ be the right censoring times. Maximum likelihood estimation is now based on the likelihood function

$$
\begin{aligned}
L(\kappa, \lambda, \theta) &= \prod_{i=1}^{r} f(t_i, \kappa, \lambda, \theta) \prod_{i=1}^{m} S(c_i, \kappa, \lambda, \theta) \\
&= \prod_{i=1}^{r} \frac{\lambda \kappa (1 + (e^{\lambda \theta} - 1)^{\kappa})(e^{\lambda(t_i+\theta)} - 1)^{\kappa - 1} e^{\lambda(t_i+\theta)}}{(1 + (e^{\lambda(t_i+\theta)} - 1)^{\kappa})^2} \\
&\quad \times \prod_{i=1}^{m} \frac{1 + (e^{\lambda \theta} - 1)^{\kappa}}{1 + (e^{\lambda(c_i+\theta)} - 1)^{\kappa}}.
\end{aligned}
$$

To get initial estimates for the three parameters, three fractiles are taken to build three equations by the fractile formula given as a property in Section 2. The same numerical method is used to simultaneously solve for the initial estimates, yielding roughly $\hat{\kappa}_0 = 0.4$, $\hat{\lambda}_0 = 0.01$, and $\hat{\theta}_0 = 6$. Solving for the MLEs of the parameters numerically yields

$$
\hat{\kappa} = 0.133344, \quad \hat{\lambda} = 0.06557, \quad \hat{\theta} = 14.5013,
$$

and a K–S statistic of 0.120. The hazard function at the MLEs is plotted in Fig. 8. The hazard function has the BT shape,

Table 4. Time to vomit data set (in minutes; the *'s denote right-censored observations).

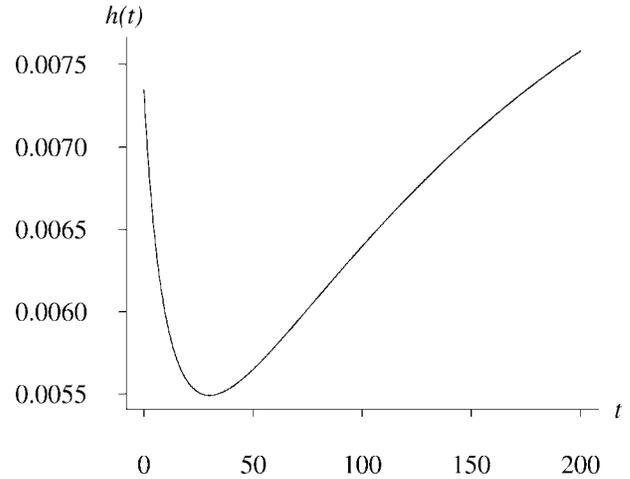| 5	6*	11	11	13	24	63
65	69	69	79	82	82	102
115	120*	120*	120*	120*	120*	120*
120*	120*	120*	120*	120*	120*	120*

Figure 8. Hazard function of the fitted three-parameter logistic–exponential distribution.

indicating that, at the beginning, the human body that survives the first few minutes of vertical motion rapidly adjusts to the cyclical motion conditions, thus lowering the hazard function considerably. As time passes, however, effects accumulate, and the body begins to tire and the hazard function increases.

4. CONCLUSIONS AND FURTHER WORK

The logistic–exponential distribution enjoys the unique property of traversing all of the four commonly seen classes of distributions, i.e., IFR, DFR, BT, and UBT. It has been shown that the survival function and its inverse can be expressed in closed form. Moments of any degree exist, although they cannot be expressed in closed form. This distribution gives reliability engineers and biostatisticians another option for modeling lifetimes.

The distribution can also be used to determine the class from which a particular data set has been drawn. But to automate the use of this model, some specific numerical methods still need to be developed. Although we have proposed a fixed-point method for the two-parameter logistic–exponential distribution, and have successfully applied it to simulated data sets, a rigorous proof of its convergence is still desirable. Also, work still needs to be done to extend the flexibility of the numerical methods for the three-parameter case.

APPENDIX A. ASYMPTOTIC PROPERTIES OF THE MOMENTS

The moments associated with the two-parameter logistic–exponential distribution cannot be expressed in closed form. This appendix investigates the

behavior of these moments as $\kappa \downarrow 0$ and $\kappa \to \infty$. One of the purposes of this investigation is to determine the limiting behavior of the coefficient of variation $\gamma_1 = \sigma/\mu$ and the skewness $\gamma_2 = E[(T - \mu)^3]/\sigma^3$ for plotting the curve in Fig. 4. More specifically, these limits pin down the two ends of the curve for the two-parameter logistic–exponential distribution in Fig. 4.

When κ Approaches Infinity

Moments for the logistic–exponential distribution in the two-parameter case are:

$$M_n = E[T^n] = \int_0^\infty t^n f(t) dt = \int_0^\infty S(t)(t^n)' dt$$

$$= \int_0^\infty \frac{nt^{n-1}}{1 + (e^{\lambda t} - 1)^\kappa} dt,$$

via integration by parts. Replacing t with $\ln(x+1)/\lambda$, we have:

$$M_n = \int_0^\infty \frac{n \ln^{n-1}(x+1)}{\lambda^n (x+1)(x^\kappa + 1)} dx.$$

In order to calculate the limit as $\kappa \to \infty$, consider the inequalities below:

$$\lambda^n M_n = \int_0^1 \frac{n \ln^{n-1}(x+1)}{(x+1)(x^\kappa+1)} dx + \int_1^\infty \frac{n \ln^{n-1}(x+1)}{(x+1)(x^\kappa+1)} dx$$

$$< \int_0^1 \frac{n \ln^{n-1}(x+1)}{x+1} dx + \int_1^\infty \frac{nx^{n-1}}{x^\kappa} dx$$

$$= \ln^n(2) + \frac{n}{\kappa - n},$$

for $n - \kappa < 0$ since $\ln(x+1) \le x$ for $x \ge 0$ and the random variable T has positive support (which implies that $x > 0$). On the other hand,

$$\lambda^n M_n > \int_0^1 \frac{n \ln^{n-1}(x+1)}{(x+1)(x^\kappa+1)} dx$$

$$= \int_0^1 \frac{n(1 + x^\kappa - x^\kappa) \ln^{n-1}(x+1)}{(x+1)(x^\kappa+1)} dx$$

$$= \int_0^1 \frac{n \ln^{n-1}(x+1)}{x+1} dx - \int_0^1 \frac{nx^\kappa \ln^{n-1}(x+1)}{(x+1)(x^\kappa+1)} dx$$

$$> \ln^n(2) - n \int_0^1 x^{\kappa+n-1} dx$$

$$= \ln^n(2) - \frac{n}{\kappa + n}$$

The squeeze method is used to determine that the value of the nth moment as κ approaches infinity:

$$\lim_{\kappa \to \infty} M_n = \left(\frac{\ln 2}{\lambda}\right)^n.$$

This conclusion can be used to calculate the limiting coefficient of variation γ_1:

$$\lim_{\kappa \to \infty} \gamma_1 = \lim_{\kappa \to \infty} \frac{\sigma}{\mu} = \lim_{\kappa \to \infty} \frac{\sqrt{M_2 - M_1^2}}{M_1}$$

$$= \frac{\sqrt{\ln^2 2/\lambda^2 - (\ln 2/\lambda)^2}}{\ln 2/\lambda} = 0.$$

For the calculation of the limiting skewness γ_2, the direct application of the results presented here yields an indeterminate form:

$$\lim_{\kappa \to \infty} \gamma_2 = \lim_{\kappa \to \infty} \frac{M_3 - 3M_2 M_1 + 2M_1^3}{\left(M_2 - M_1^2\right)^{3/2}} = \frac{0}{0}.$$

This is consistent with the fact that the limiting distribution of T is degenerate at $\ln 2/\lambda$ as $\kappa \to \infty$ since

$$\lim_{\kappa \to \infty} F(t) = \begin{cases} 0 & t < \ln 2/\lambda \\ 1/2 & t = \ln 2/\lambda \\ 1 & t > \ln 2/\lambda. \end{cases}$$

Numerical analysis, however, reveals that the skewness also approaches 0 for large κ. This establishes $(\gamma_1, \gamma_2) = (0, 0)$ as the limiting point in Fig. 4 as $\kappa \to \infty$.

When κ Approaches Zero

From the previous section:

$$M_n = \int_0^\infty \frac{n \ln^{n-1}(x+1)}{\lambda^n (x+1)(x^\kappa + 1)} dx.$$

Replacing x with y^m, where $m = 1/\kappa$, we have:

$$M_n = \int_0^\infty \frac{mny^m \ln^{n-1}(y^m + 1)}{\lambda^n y(y+1)(y^m + 1)} dy.$$

Taking the limit and simplifying:

$$\lim_{\kappa \downarrow 0} (\lambda \kappa)^n M_n = \lim_{m \to \infty} \int_0^\infty \frac{ny^m \ln^{n-1}(y^m + 1)}{m^{n-1} y(y+1)(y^m + 1)} dy$$

$$= \lim_{m \to \infty} \int_1^\infty \frac{ny^m \ln^{n-1}(y^m + 1)}{m^{n-1} y(y+1)(y^m + 1)} dy$$

$$+ \left[\lim_{m \to \infty} \int_0^1 \frac{ny^m \ln^{n-1}(y^m + 1)}{m^{n-1} y(y+1)(y^m + 1)} dy = 0 \right]$$

$$= \lim_{m \to \infty} \int_1^\infty \frac{n(y^m + 1 - 1) \ln^{n-1}(y^m + 1)}{m^{n-1} y(y+1)(y^m + 1)} dy$$

$$= \lim_{m \to \infty} \int_1^\infty \frac{n \ln^{n-1}(y^m + 1)}{m^{n-1} y(y+1)} dy$$

$$- \left[\lim_{m \to \infty} \int_1^\infty \frac{n \ln^{n-1}(y^m + 1)}{m^{n-1} y(y+1)(y^m + 1)} dy = 0 \right]$$

$$= \lim_{m \to \infty} \int_1^\infty \frac{n(\ln^{n-1}(y^m + 1) - \ln^{n-1}(y^m) + \ln^{n-1}(y^m))}{m^{n-1} y(y+1)} dy$$

$$= \left[\lim_{m \to \infty} \int_1^\infty \frac{n(\ln^{n-1}(y^m + 1) - \ln^{n-1}(y^m))}{m^{n-1} y(y+1)} dy = 0 \right]$$

$$+ \lim_{m \to \infty} \int_1^\infty \frac{n \ln^{n-1}(y)}{y(y+1)} dy$$

$$= \lim_{m \to \infty} \int_1^\infty \frac{n \ln^{n-1}(y)}{y(y+1)} dy.$$

The reason that

$$\lim_{m \to \infty} \int_0^1 \frac{ny^m \ln^{n-1}(y^m + 1)}{m^{n-1} y(y+1)(y^m + 1)} dy = 0$$

is that $0 \leq \ln(y^m + 1) \leq y^m$, so substituting with this inequality, it is obvious that

$$\lim_{m\to\infty} \int_0^1 \frac{ny^{mn}}{m^{n-1}y(y+1)(y^m+1)}\, dy = 0.$$

The squeeze method can be applied successfully here as well. The reason that

$$\lim_{m\to\infty} \int_1^\infty \frac{n\ln^{n-1}(y^m+1)}{m^{n-1}y(y+1)(y^m+1)}\, dy = 0$$

is obvious when $n = 1$. For $n > 1$ it can be reasoned that since

$$\frac{\ln^{n-1}(y^m+1)}{(y^m+1)}$$

has an upper bound $a(n)$ for any integer $n > 1$, the integral also has an upper limit, and since

$$\lim_{m\to\infty} \int_1^\infty \frac{na(n)}{m^{n-1}y(y+1)}\, dy = 0,$$

the squeeze method can be applied to obtain the zero limit. The reason that

$$\lim_{m\to\infty} \int_1^\infty \frac{n(\ln^{n-1}(y^m+1) - \ln^{n-1}(y^m))}{m^{n-1}y(y+1)}\, dy = 0,$$

is also obvious when $n = 1$. For $n > 1$ it can be reasoned that since

$$\ln^{n-1}(y^m+1) - \ln^{n-1}(y^m)$$

has an upper bound $b(n)$ for any given $n > 1$, the integral also has an upper limit, and since

$$\lim_{m\to\infty} \int_1^\infty \frac{nb(n)}{m^{n-1}y(y+1)}\, dy = 0,$$

the squeeze method can be applied to obtain the zero limit. Replacing z with $\ln(y)$, we have:

$$\lim_{\kappa\downarrow 0}(\lambda\kappa)^n M_n = \lim_{m\to\infty} \int_0^\infty \frac{nz^{n-1}}{e^z+1}\, dz = \int_0^\infty \frac{nz^{n-1}}{e^z+1}\, dz.$$

Since this integral exists and depends only on n, denote it as $m(n)$. From this result we can calculate the limiting coefficient of variation $\gamma_1 = \sigma/\mu$ as κ decreases to 0:

$$\lim_{\kappa\downarrow 0}\gamma_1 = \lim_{\kappa\downarrow 0}\frac{\sqrt{M_2 - M_1^2}}{M_1}$$
$$= \lim_{\kappa\downarrow 0}\frac{\sqrt{(\lambda\kappa)^2 M_2 - (\lambda\kappa M_1)^2}}{\lambda\kappa M_1}$$
$$= \frac{\sqrt{m(2) - m(1)^2}}{m(1)}.$$

Likewise, we can express the limiting skewness as:

$$\lim_{\kappa\downarrow 0}\gamma_2 = \lim_{\kappa\downarrow 0}\frac{M_3 - 3M_2M_1 + 2M_1^3}{(M_2 - M_1^2)^{3/2}}$$
$$= \lim_{\kappa\downarrow 0}\frac{(\lambda\kappa)^3 M_3 - 3(\lambda\kappa)^2 M_2(\lambda\kappa)M_1 + 2(\lambda\kappa)^3 M_1^3}{((\lambda\kappa)^2 M_2 - (\lambda\kappa)^2 M_1^2)^{3/2}}$$
$$= \frac{m(3) - 3m(2)m(1) + 2m(1)^3}{(m(2) - m(1)^2)^{3/2}}.$$

Using numerical methods, these values are $\lim_{\kappa\downarrow 0}\gamma_1 \cong 1.5568$ and $\lim_{\kappa\downarrow 0}\gamma_2 \cong 2.1126$.

APPENDIX B. RANGE OF MLES FOR PARAMETERS

In this appendix we provide a rectangular boundary for the MLEs of $\hat{\lambda}$ and $\hat{\kappa}$ for the two-parameter logistic–exponential distribution for a complete data set. If numerical solvers fail to converge to the MLEs, a grid search over this rectangle is appropriate. For the theorems stated below, we assume that $t_1, t_2, \ldots, t_n$ denote the lifetimes and $t_{(1)}, t_{(2)}, \ldots, t_{(n)}$ denote the associated order statistics drawn from a logistic–exponential distribution with survivor function

$$S(t) = P(T \geq t) = \frac{1}{1 + (e^{\lambda t} - 1)^\kappa} \quad t \geq 0.$$

Range of $\hat{\lambda}$ When $\hat{\kappa}$ Is Unknown

This section proves the inequality that provides lower and upper bounds for $\hat{\lambda}$ as:

$$\frac{\ln 2}{t_{(n)}} < \hat{\lambda} < \frac{\ln 2}{t_{(1)}}.$$

Lower Bound for $\hat{\lambda}$

THEOREM: The lower bound for $\hat{\lambda}$ is $\hat{\lambda}_l = \ln 2/t_{(n)}$.

PROOF: Use proof by contradiction. Rewrite the second element of the score vector as:

$$\frac{\partial \ln L(\kappa, \lambda)}{\partial \lambda} = \sum_{i=1}^n \left(\frac{e^{\lambda t_i} - \lambda t_i - 1}{\lambda(e^{\lambda t_i} - 1)} + \frac{1 - (e^{\lambda t_i} - 1)^\kappa}{1 + (e^{\lambda t_i} - 1)^\kappa} \cdot \frac{\kappa t_i e^{\lambda t_i}}{e^{\lambda t_i} - 1} \right).$$

If $0 < \hat{\lambda}t_{(n)} \leq \ln 2$, which implies $0 < \hat{\lambda}t_i \leq \ln 2, i = 1, 2, \ldots, n$, then $0 < e^{\hat{\lambda}t_i} - 1 \leq 1, i = 1, 2, \ldots, n$, so

$$\frac{1 - (e^{\hat{\lambda}t_i} - 1)^{\hat{\kappa}}}{1 + (e^{\hat{\lambda}t_i} - 1)^{\hat{\kappa}}} \cdot \frac{\hat{\kappa} t_i e^{\hat{\lambda}t_i}}{e^{\hat{\lambda}t_i} - 1} \geq 0, \quad i = 1, 2, \ldots, n.$$

Since $\hat{\lambda}t_i > 0 \Rightarrow e^{\hat{\lambda}t_i} - \hat{\lambda}t_i - 1 > 0, i = 1, 2, \ldots, n$, the term below is obviously positive:

$$\frac{e^{\hat{\lambda}t_i} - \hat{\lambda}t_i - 1}{\hat{\lambda}(e^{\hat{\lambda}t_i} - 1)} > 0.$$

Thus each term is positive, so the sum of all these positive terms should also be positive, which means

$$\frac{\partial \ln L(\hat{\kappa}, \hat{\lambda})}{\partial \lambda} > 0.$$

This contradicts the fact that the MLEs satisfy

$$\frac{\partial \ln L(\hat{\kappa}, \hat{\lambda})}{\partial \lambda} = 0.$$

This establishes the proposed lower bound for $\hat{\lambda}$. $\square$

Upper Bound for $\hat{\lambda}$

THEOREM: The upper bound for $\hat{\lambda}$ is $\hat{\lambda}_u = \ln 2 / t_{(1)}$.

PROOF: Again, use proof by contradiction. After simplification,

$$\kappa \frac{\partial \ln L(\kappa, \lambda)}{\partial \kappa} - \lambda \frac{\partial \ln L(\kappa, \lambda)}{\partial \lambda} = \sum_{i=1}^{n} \left(\kappa \frac{1 - (e^{\lambda t_i} - 1)^{\kappa}}{1 + (e^{\lambda t_i} - 1)^{\kappa}} \right.$$
$$\left. \times \left(\ln(e^{\lambda t_i} - 1) - \frac{\lambda t_i e^{\lambda t_i}}{e^{\lambda t_i} - 1} \right) + \frac{\lambda t_i}{e^{\lambda t_i} - 1} \right).$$

If $\hat{\lambda} t_{(1)} \geq \ln 2$, which implies that $\hat{\lambda} t_i \geq \ln 2, i = 1, 2, \ldots, n$, then $e^{\lambda t_i} - 1 \geq 1, i = 1, 2, \ldots, n$. Since additionally, $\hat{\kappa} > 0, 1 - (e^{\hat{\lambda} t_i} - 1)^{\hat{\kappa}} \leq 0$, which implies

$$\hat{\kappa} \frac{1 - (e^{\hat{\lambda} t_i} - 1)^{\hat{\kappa}}}{1 + (e^{\hat{\lambda} t_i} - 1)^{\hat{\kappa}}} \leq 0 \quad i = 1, 2, \ldots, n.$$

Also, the other factor in the summand is always negative, i.e.,

$$\ln(e^{\hat{\lambda} t_i} - 1) - \frac{\hat{\lambda} t_i e^{\hat{\lambda} t_i}}{e^{\hat{\lambda} t_i} - 1} < 0 \quad i = 1, 2, \ldots, n.$$

If we rewrite this factor as

$$\frac{(e^{\hat{\lambda} t_i} - 1) \ln(e^{\hat{\lambda} t_i} - 1) - e^{\hat{\lambda} t_i} \ln(e^{\hat{\lambda} t_i})}{e^{\hat{\lambda} t_i} - 1} < 0 \quad i = 1, 2, \ldots, n,$$

and denote $x_i = e^{\hat{\lambda} t_i} - 1 \geq 1$, we have

$$\frac{x_i \ln(x_i) - (x_i + 1) \ln(x_i + 1)}{x_i} < 0, (x_i \geq 1) \quad i = 1, 2, \ldots, n,$$

since $x_i \ln(x_i)$ is monotonically increasing when $x_i \geq 1$. So

$$\hat{\kappa} \frac{\partial \ln L(\hat{\kappa}, \hat{\lambda})}{\partial \kappa} - \hat{\lambda} \frac{\partial \ln L(\hat{\kappa}, \hat{\lambda})}{\partial \lambda} > 0.$$

which contradicts the first-order MLE assumptions:

$$\frac{\partial \ln L(\hat{\kappa}, \hat{\lambda})}{\partial \kappa} = 0 \quad \text{and} \quad \frac{\partial \ln L(\hat{\kappa}, \hat{\lambda})}{\partial \lambda} = 0.$$

This establishes the upper bound as claimed. In summary, the range of $\hat{\lambda}$ is

$$\frac{\ln 2}{t_{(n)}} < \hat{\lambda} < \frac{\ln 2}{t_{(1)}}. \qquad \square$$

Range of $\hat{\kappa}$ When $\hat{\lambda}$ Is Unknown

This section proves the inequality that provides lower and upper bounds for $\hat{\kappa}$ as:

$$\frac{t_{(1)} p}{t_{(n)} \ln 2} < \hat{\kappa} < \frac{n + 3}{-\ln(2^{t_{(1)}/t_{(n)}} - 1)},$$

where $p \cong 1.58$ is a constant determined numerically.

Lower Bound for $\hat{\kappa}$

THEOREM: The lower bound for $\hat{\kappa}$ when $\hat{\lambda}$ is unknown is

$$\hat{\kappa} > \frac{p}{\ln 2} \cdot \frac{t_{(1)}}{t_{(n)}},$$

where p is the only positive root of $1 - x \tanh(x/2) = 0$, found numerically as $p \cong 1.58$.

PROOF: Consider the ith data value in the second element of the score vector, which is

$$\frac{\partial \ln f(t_i; \kappa, \lambda)}{\partial \kappa} = \ln(e^{\lambda t_i} - 1) + \frac{1}{\kappa} - \frac{2(e^{\lambda t_i} - 1)^{\kappa} \ln(e^{\lambda t_i} - 1)}{1 + (e^{\lambda t_i} - 1)^{\kappa}}$$

Defining $x_i = \kappa \ln(e^{\lambda t_i} - 1)$, we can reduce the expression above to

$$\frac{\partial \ln f(t_i; \kappa, \lambda)}{\partial \kappa} = \frac{1}{\kappa} \left(1 + \frac{x_i(1 - e^{x_i})}{1 + e^{x_i}} \right).$$

Thus, the sign of the partial derivative is determined as a function of x_i. Define

$$g(x) = 1 + \frac{x(1 - e^x)}{1 + e^x} = 1 - x \tanh(x/2),$$

which is an even function, and

$$g'(x) = \frac{1 - e^{2x} - 2x_i e^x}{(1 + e^x)^2},$$

which is an odd function. There are two real roots to $g(x) = 0$, denoted by $p_1 = -p$ and $p_2 = p$. Define the right-hand-side of the inequalities below

$$\kappa \ln(e^{\hat{\lambda} t_{(n)}} - 1) < \kappa \ln(e^{\hat{\lambda}_u t_{(n)}} - 1) = X_u,$$
$$\kappa \ln(e^{\hat{\lambda} t_{(1)}} - 1) < \kappa \ln(e^{\hat{\lambda}_l t_{(1)}} - 1) = X_l.$$

It can be proved that[3] $X_l + X_u \geq 0$, which means if $X_u \leq p$, then $X_l > -X_u \geq -p$. Since $X_l < x_i < X_u \Rightarrow p_1 < x_i < p_2 \Rightarrow g(x_i) > 0, i = 1, 2, \ldots, n,$

$$\frac{\partial \ln f(t_i; \kappa, \lambda)}{\partial \kappa} > 0, i = 1, 2, \ldots, n \quad \Rightarrow \quad \sum_{i=1}^{n} \frac{\partial \ln f(t_i; \kappa, \lambda)}{\partial \kappa} > 0.$$

That contradicts the fact that the partial derivative of the log likelihood function is zero at the MLEs. So $X_u > p$, which yields

$$\kappa \ln(e^{\hat{\lambda}_u t_{(n)}} - 1) > p \quad \Rightarrow \quad \kappa > \frac{p}{\ln(2^{t_{(n)}/t_{(1)}} - 1)} > \frac{t_{(1)} p}{t_{(n)} \ln 2}. \qquad \square$$

---

[3] This is equivalent to proving that $(2^r - 1)(2^{1/r} - 1) \geq 1$, when $r \geq 1$. Since, when $r = 1$, the equality holds, it can be proved if its derivative is non-negative: $2^r (2^{1/r} - 1) - 2^{1/r} (2^r - 1)/r^2 \geq 0$. Rewrite that as $r 2^r / (2^r - 1) \geq (1/r) 2^{1/r} / (2^{1/r} - 1)$, which suggests we can just prove the monotonicity of $r 2^r / (2^r - 1)$, or $x \log_2 x / (x - 1)$ is increasing for $x \geq 2$ after substitution of $x = 2^r$, which can be proved by taking a derivative.

Upper bound for $\hat{\kappa}$

THEOREM: The upper bound for $\hat{\kappa}$ when $\hat{\lambda}$ is unknown is

$$\hat{\kappa} < \frac{n+3}{-\ln(2^{t_{(1)}/t_{(n)}} - 1)}.$$

PROOF: The following notation will be used throughout the proof:

$$x_i = \hat{\kappa} \ln(e^{\hat{\lambda} t_i} - 1), \quad i = 1, 2, \ldots, n,$$
$$x_u = \hat{\kappa} \ln(e^{\hat{\lambda} t_{(n)}} - 1) > 0,$$
$$x_l = \hat{\kappa} \ln(e^{\hat{\lambda} t_{(1)}} - 1) < 0.$$

Recall that

$$\frac{\partial \ln L(\kappa, \lambda)}{\partial \kappa} = \kappa \sum_{i=1}^{n} g(x_i) = 0 \quad \Rightarrow \quad \sum_{i=1}^{n} g(x_i) = 0,$$

which implies that (since $g(x) \leq n-1$):

$$\sum_{|x_i|>p} -g(x_i) = \sum_{|x_i|\leq p} g(x_i) \leq n-1.$$

It can be shown that $g(x)$ has the following properties:

$$|x| > p \quad \Rightarrow \quad |x| - p < -g(x).$$

Thus, it follows that

$$(|x_u| - p) + (|x_l| - p) \leq \sum_{|x_i|>p} -g(x_i) \leq n-1$$

since $x_u \geq 0$ and $x_l \leq 0$. Thus we have

$$x_u - x_l \leq n + 2p - 1.$$

Obviously $2p - 1 < 3$, so we have:

$$\hat{\kappa} < \frac{n+3}{\ln(e^{\hat{\lambda} t_{(n)}} - 1) - \ln(e^{\hat{\lambda} t_{(1)}} - 1)}.$$

The right-hand side of the inequality above is monotonically decreasing to $\hat{\lambda}$, so we can use the lower bound of λ here to get an upper bound for $\hat{\kappa}$. In summary, the range of $\hat{\kappa}$ is:

$$\frac{t_{(1)} p}{t_{(n)} \ln 2} < \hat{\kappa} < \frac{n+3}{-\ln(2^{t_{(1)}/t_{(n)}} - 1)}. \qquad \square$$

Range of $\hat{\kappa}$ When $\hat{\lambda}$ Is Known

The boundary for $\hat{\kappa}$ is quite loose, but, if $\hat{\lambda}$ is already known, there is a much narrower range for $\hat{\kappa}$, based on the observation that $g(x) = 1 - x \tanh(x/2)$ and that $1 - |x| \leq g(x) < 1.557 - |x|$, as illustrated in Fig. 9. The value 1.557 is the numerical approximation of $g(1 + \text{LambertW}(1/e)) + 1 + \text{LambertW}(1/e)$, where $\text{LambertW}(1/e)$ is the real root of $xe^x = 1/e$. Since at the MLE it is already known that $\sum_{i=1}^{n} g(x_i) = 0$, summing through all the inequalities on x_i yields

$$n - \sum_{i=1}^{n} |x_i| < 0 < 1.557n - \sum_{i=1}^{n} |x_i|.$$
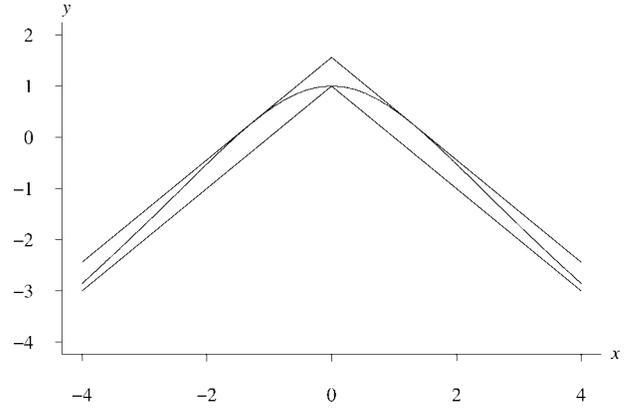
Figure 9. $g(x)$ lies between $1 - |x|$ and $1.557 - |x|$.

Recall that $x_i = \hat{\kappa} \ln(e^{\hat{\lambda} t_i} - 1), i = 1, 2, \ldots, n$, so $\hat{\kappa}$ can be solved from the inequality, producing a much tighter range $\hat{\kappa}_l^t < \hat{\kappa} < \hat{\kappa}_u^t$ than the range in the previous section:

$$\frac{n}{\sum_{i=1}^{n} |\ln(e^{\hat{\lambda} t_i} - 1)|} < \hat{\kappa} < \frac{1.557n}{\sum_{i=1}^{n} |\ln(e^{\hat{\lambda} t_i} - 1)|}.$$

Boundaries Applied to the Ball Bearing Failure Times

As discussed in the statistical inference section, the oft-analyzed ball bearing failure times have been fitted to the logistic–exponential distribution. The boundaries will be calculated for this data set to verify that the MLEs $\hat{\kappa}$ and $\hat{\lambda}$ lie in the rectangular region rendered by the upper and lower boundaries.

First calculate the boundaries for $\hat{\lambda}$ from the ball bearing failure times. The ball bearing data set has $n = 23$ observations, with $t_{(1)} = 17.88, t_{(23)} = 173.40$, thus we have:

$$\frac{\ln 2}{t_{(n)}} < \hat{\lambda} < \frac{\ln 2}{t_{(1)}} \quad \Rightarrow \quad 0.0040 < \hat{\lambda} < 0.0388.$$

Since the upper bound and the lower bounds for $\hat{\lambda}$ are functions of $t_{(1)}$ and $t_{(23)}$ respectively, a natural conjecture about the initial guess for $\hat{\lambda}$ would be something of an average of the function applied to all the observations, which is the harmonic average:

$$\hat{\lambda}_h = \frac{\ln 2}{n} \sum_{i=1}^{n} \frac{1}{t_i} = 0.0126.$$

Still another simple initial guess is to use the arithmetic average:

$$\hat{\lambda}_a = \frac{n \ln 2}{\sum_{i=1}^{n} t_i} = 0.0096.$$

Both of these initial guesses are reasonably close to the MLE $\hat{\lambda} = 0.01059$, and, in general, both will lie in the ranges developed here for distinct data values. It is difficult to tell which initial guess is more accurate; in our fixed-point method, the latter is adopted as the initial guess of $\hat{\lambda}$.
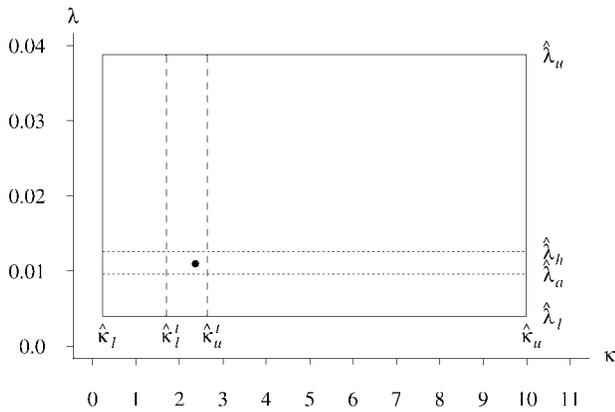
Figure 10. Parameter boundaries for the ball bearing failure times.

Now consider the boundaries for $\hat{\kappa}$, which will be calculated according to the formula developed previously:

$$\frac{t_{(1)} p}{t_{(n)} \ln 2} < \hat{\kappa} < \frac{n+3}{-\ln(2^{t_{(1)}/t_{(n)}} - 1)}$$

or $0.2350 < \hat{\kappa} < 9.9905$. Finally, the tighter boundaries for $\hat{\kappa}$ when $\hat{\lambda}$ is known for a complete data set are

$$\frac{n}{\sum_{i=1}^{n} |\ln(e^{\hat{\lambda} t_i} - 1)|} < \hat{\kappa} < \frac{1.557n}{\sum_{i=1}^{n} |\ln(e^{\hat{\lambda} t_i} - 1)|}$$

or $1.6938 < \hat{\kappa} < 2.6371$.

Figure 10 shows the boundaries, initial parameter guesses, and the MLEs $(\hat{\kappa}, \hat{\lambda})$. The lower and upper boundaries for $\hat{\lambda}$ are labeled $\hat{\lambda}_l$ and $\hat{\lambda}_u$ respectively, and the lower and upper boundaries for $\hat{\kappa}$ when $\hat{\lambda}$ is unknown are labeled $\hat{\kappa}_l$ and $\hat{\kappa}_u$, and the tight boundaries for $\hat{\kappa}$ when $\hat{\lambda}$ is given are labeled $\hat{\kappa}_l^t$ and $\hat{\kappa}_u^t$ respectively. The two initial guesses are labeled $\hat{\lambda}_h$ for using the harmonic average and $\hat{\lambda}_a$ for using the simple arithmetic average. The MLE is also plotted on the graph and lies within the larger rectangular region as expected.

APPENDIX C. FIXED-POINT METHOD FOR MLE

In this appendix we develop a fixed-point method for the numerical calculation of the MLEs for the two-parameter logistic–exponential distribution for a complete data set. A rigorous proof of this method is not provided, but we include Maple code for estimating MLEs. Our sense is that the procedure given here works with field data, and would only fail when the data set is ill conditioned, i.e., the ratio of the largest observation to the smallest observation is extremely large (e.g., in 10^8 when you are planning to have eight effective digits of accuracy for the MLEs). The term "ill conditioned" is not in an absolute sense though, as it actually depends on what precision one expects from the fixed-point method. Generally speaking, the more ill conditioned, the fewer effective digits one can get. On the other hand, if we have infinite-precision computation facility, this is not an issue at all.

The fixed-point method we have developed and experimented with is quite simple. We need two functions that calculate new κ and λ from old ones as well as all the observations. These are the fixed-point functions we have implemented:

$$\kappa_{k+1} = n \left(\sum_{i=1}^{n} \ln(e^{\lambda_k t_i} - 1) \tanh(\kappa_k \ln(e^{\lambda_k t_i} - 1)/2) \right)^{-1},$$

$$\lambda_{k+1} = \lambda_k - \frac{\partial \ln L(\lambda_k, \kappa_{k+1})}{\partial \lambda} \bigg/ \frac{\partial^2 \ln L(\lambda_k, \kappa_{k+1})}{\partial \lambda^2}.$$

The fixed-point equation for κ is just a straightforward transformation of the first-order condition for κ

$$\frac{\partial \ln L(\lambda, \kappa)}{\partial \kappa} = \frac{n}{\kappa} - \sum_{i=1}^{n} \ln(e^{\lambda t_i} - 1) \tanh(\kappa \ln(e^{\lambda t_i} - 1)/2) = 0$$

and the fixed-point equation for λ is simply Newton's method, which is a special case of fixed-point method, applied to the first-order condition for λ

$$\frac{\partial \ln L(\lambda, \kappa)}{\partial \lambda} = 0$$

with a little variation in that we used the new κ as soon as it is available, which speeds convergence.

Another important issue in numerical methods is how to determine a good initial guess. Fortunately, this can be readily computed from our previous work on the boundaries. We chose the following initial guesses:

$$\lambda_0 = \frac{n \ln 2}{\sum_{i=1}^{n} t_i}, \quad \kappa_0 = \frac{n}{\sum_{i=1}^{n} |\ln(e^{\lambda_0 t_i} - 1)|}.$$

Once the initial values are calculated, they can be fed into the fixed-point method and iterations are carried out until it converges to the desired accuracy. The Maple procedure `lgxFixedPoint` given at the end of this appendix follows the considerations above: it first checks if the data is ill conditioned, and eventually returns the number of iterations and the results from the last iteration. Convergence is achieved if the number of iterations is less than the prescribed maximal number of iterations.

This procedure works well and converges with an average of about 20 iterations for simulated logistic–exponential data for $n = 5, 10, 25$, and 50, and for λ ranging from 0.1 to 100 and κ ranging from 0.5 to 10. Our results show that except for those ill-conditioned data sets, the fixed point method provides MLEs with the desired accuracy, which demonstrates that the fixed-point method is quite reliable.

Some experiments on how to extend this method to data sets with censored data have been carried out with satisfactory results. The initial guess and the fixed-point formula changed slightly. We have implemented the iterative method in Maple. Similar testing procedures are also developed, the testing results are quite good. Below we give the mathematical formulas:

$$\kappa_{k+1} = r \left(\sum_{i=1}^{r} \ln(e^{\lambda_k t_i} - 1) \tanh(\kappa_k \ln(e^{\lambda_k t_i} - 1)/2) \right.$$
$$\left. + \sum_{j=1}^{s} \frac{\ln(e^{\lambda_k c_j} - 1)}{1 + (e^{\lambda_k c_j} - 1)^{-\kappa_k}} \right)^{-1},$$

$$\lambda_{k+1} = \lambda_k - \frac{\partial \ln L(\lambda_k, \kappa_{k+1})}{\partial \lambda} \bigg/ \frac{\partial^2 \ln L(\lambda_k, \kappa_{k+1})}{\partial \lambda^2},$$

where the $t_i, i = 1, 2, \ldots, r$ are observed lifetimes, $c_j, j = 1, 2, \ldots, s$ are

censoring times. The initial guesses are given by

$$\lambda_0 = \frac{r \ln 2}{\sum_{i=1}^{r} t_i + \sum_{j=1}^{s} c_i} \quad \text{and}$$

$$\kappa_0 = \frac{r}{\sum_{i=1}^{r} |\ln(e^{\lambda_0 t_i} - 1)| + \sum_{j=1}^{s} |\ln(e^{\lambda_0 c_j} - 1)|}.$$

```
#
# procedure name:   lgxFixedPoint
# argument:         t (a list of lifetimes)
# returned values:  number of iterations, MLEs
#
lgxFixedPoint := proc(t)
local newk, newl, logf, dlogf, ddlogf, lgxeps,
      big, sml, nobs, lamb, kapp, iterat, delta,
      converged;
nobs := nops(t);
lgxeps := 0.000000001;
sml := t[1];
big := t[1];
for iterat from 1 to nobs do
  big := max(t[iterat], big);
  sml := min(t[iterat], sml);
end do;
if (sml < 100 * big * lgxeps) then
  return [0, "Ill-conditioned data!"];
end if;
newk := (kappa, lambda, t, n) -> n /
      sum(ln(exp(lambda * t[j]) - 1)
      * tanh(kappa * ln(exp(lambda * t[j])
      - 1) / 2), j = 1 .. n);
logf := (kappa - 1) * ln(exp(lambda * nt[i])
      - 1) + ln(kappa) + ln(lambda)
      + lambda * nt[i] - 2 * ln(1
      + (exp(lambda * nt[i]) - 1)^ kappa);
dlogf := diff(logf, lambda);
ddlogf := diff(dlogf, lambda);
newl := (k, l, t, n) -> l - sum(subs (kappa = k,
      lambda = l, nt = t, dlogf), i = 1 .. n)
      / sum(subs(kappa = k, lambda = l, nt = t,
      ddlogf), i = 1 .. n);
lamb := evalf(nobs * log(2) / sum(t[i],
      i = 1 .. nobs));
kapp := nobs / sum(abs(ln(exp(lamb *
      t[i]) - 1)), i = 1 .. nobs);
for iterat from 1 to 90 do
  converged := 0;
  delta := kapp;
  kapp := newk(kapp, lamb, t, nobs);
  delta:= abs(kapp - delta) / kapp;
  if (delta < lgxeps) then converged := 1 end if;
  delta := lamb;
  lamb := newl(kapp, lamb, t, nobs);
  delta := abs(lamb - delta) / lamb;
  if (delta < lgxeps) then
    converged := converged + 1
  end if;
  if (converged = 2) then break end if;
end do;
return [iterat, [kapp, lamb]];
end proc;
```

REFERENCES

[1] K.C. Burns, Motion sickness incidence: Distribution of time to first emesis and comparison of some complex motion conditions, Aviation Space Environ Med 56 (1984), 521–527.
[2] I.W. Burr, Cumulative frequency functions, Ann Math Statist 13 (1942), 215–232.
[3] C. Caroni, The correct "ball bearings" data, Lifetime Data Anal 8 (2002), 395–399.
[4] D.R. Cox and D. Oakes, Analysis of survival data, Chapman and Hall, London, England, 1984.
[5] M.J. Crowder, Classical competing risks, Chapman and Hall/CRC Press, Boca Raton, FL, 2001.
[6] M.J. Crowder, A.C. Kimber, R.L. Smith, and T.J. Sweeting, Statistical analysis of reliability data, Chapman and Hall, London, England, 1991.
[7] H.A. David and M.L. Moeschberger, The theory of competing risks, Macmillan, New York, 1978.
[8] B.S. Everitt and D.J. Hand, Finite mixture distributions, Chapman and Hall, New York, 1981.
[9] A.G. Glen and L.M. Leemis, The arctangent survival distribution, J Qual Technol 29 (1994), 205–210.
[10] U. Hjorth, A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates, Technometrics 22 (1980), 99–107.
[11] N.L. Johnson, S. Kotz, and N. Balakrishnan, Continuous univariate distributions, 2nd ed., Vol. 1, Wiley, New York, 1994.
[12] J. Lieblein and M. Zelen, Statistical investigation of the fatigue life of deep-groove ball bearings, J Res Nat Bureau Stand 57 (1956), 273–316.
[13] G. McLachlan and D. Peel, Finite mixture models, Wiley, New York, 2000.
[14] W.Q. Meeker and L.A. Escobar, Statistical methods for reliability data, Wiley, New York, 1998.
[15] G.S. Mudholkar, D.K. Srivastava, and M. Freimer, The exponentiated Weibull family: A reanalysis of the bus-motor-failure data, Technometrics 37 (1995), 436–445.
[16] M. Pintlie, Competing risks: A practical perspective, Wiley, Hoboken, NJ, 2006.
[17] P.R. Tadikamalla and N.L. Johnson, Systems of frequency curves generated by transformations of logistic random variables, Biometrika 69 (1982), 461–465.
[18] M.D. Weber, L.M. Leemis, and R.K. Kincaid, Minimum Kolmogorov–Smirnov test statistic parameter estimates, J Statist Comp Simulat 76 (2006), 195–206.