New survival distributions that quantify the gain from eliminating flawed components

Rose Baker

School of Business
University of Salford, UK
email r.d.baker@salford.ac.uk

Abstract

A general method for deriving new survival distributions from old is presented. This yields a class of useful mixture distributions. Fitting such distributions to failure-time data allows estimation of the improvement in reliability that could be gained from eliminating 'frail' components. One model parameter is the proportional increase of expected survival time that could be achieved. Some 2 and 3 parameter distributions in this class are described, which are extensions of the Weibull, exponential, gamma and lognormal distributions. The methodology is illustrated by fitting some well-travelled datasets.

Keywords: Weibull distribution, gamma distribution, mixture distribution, hazard function, partial integration, frailty model

1. Introduction

For many years, there has been considerable interest in how components and systems age, with the development of many lifetime distributions and of related concepts such as the hazard function (failure rate) and the mean residual life. Lai and Xie (2006) give an excellent account of the field.

Failure rate can be constant, increasing (IFR), or decreasing (DFR). All this can be modelled using the 2-parameter Weibull distribution (e.g. Johnson *et al* 1995). However, more complicated behaviours are often found in component failure distributions, such as the bathtub shape (BT) and even occasionally the upside-down bathtub (UBT) and the modified bathtub (MTB), where failure rate initially increases before the bathtub shape proper. Roller-coaster shapes are the ultimate in age-related failure-rate fluctuations.

Modelling such failure data has fascinated statisticians and probabilists, and a plethora of failure distributions has been created over the last decades, to add to the longer-known gamma, lognormal and log-logistic distributions.

However, as fascinating as modelling bathtub and roller-coaster hazard functions can be, we must enquire as to the practical purpose of analysing failure data. One aim of course is to obtain an estimate of mean component lifetime and its variability, and here a flexible model that can cope with censored data is certainly useful. Another practical problem with DFR and BT distributions is to decide whether burn-in is needed to reduce the failure rate to an acceptable level, and to find the optimal burn-in period (e.g. Yuan *et al*, 2016). For example, this could be chosen to maximise the mean residual life (MRL). This also usually relies on parametric modelling.

Wang *et al* (2002) discuss the reasons for the bathtub shape in detail, as do Klutke *et al* (2003). There are possible mechanisms that can cause a decreasing failure rate, such as work-hardening, although this and the strengthening of concrete through drying out are the only examples that appear in the literature. Rather, DFR distributions must nearly always be due to a mixture of stronger and weaker components, where the weaker ones fail quickly, and so failure rate progressively decreases, until eventually wearout may cause it to rise again, giving a BT shape. This explanation was first given by Proschan (1963). The discovery of an initial DFR curve might therefore lead to tighter quality-control in manufacturing, or to the identification of the major design flaws and efforts to engineer them out.

Besides the usual modeller's aim of creating new survival distributions to fit failure data better, the methodology presented here aims to estimate the benefit of removing flawed items, and so to provide better decision-support to reliability engineers. A measure ξ can be estimated that gives the potential for reliability growth.

To describe the methodology: techniques originally developed by mathematicians for evaluating integrals have already been used to generate new distributions from old, for example transformation of variables has been widely used. To generate some new survival distributions with the potential to be DFR or BT, the method of integration by parts (IBP) was applied to the pdf of a 'parent' distribution, such as the Weibull. Generating new distributions in this way has apparently not been done before.

This was a mathematical exercise, which although leading to a vast array of new distributions, might have yielded nothing of particular interest for Reliability. However, it led to a useful class of mixture (frailty) distribu-

2

tions, with one parameter (λ) more than the parent distribution. Statistical inference is straightforward, as the parent distribution is nested inside the class of 'descendant' distributions and is regained as $\lambda \to \infty$. Large-sample theory or bootstrap methods then allow one to test the null hypothesis that the extra parameter does not improve the fit to data.

Further, one can reparameterize in terms of a parameter ξ, zero for the parent distribution, and such that ξ is the proportional increase in expected lifetime on eliminating the mixture of flawed components, and retaining only the most durable. Thus one can for example fit the modified Weibull model to failure-time data and have an estimate of the gain from eliminating flawed components, under the assumptions that the parent distribution applies, and that the observed high initial failure rate is caused by flawed components.

It must be stressed that it is not claimed here that every dataset can be fitted better by one of the new distributions to be described than by existing distributions. Rather, a methodology for deriving new distributions is described, and a vast number of new possibilities now exist. Further, potential reliability gain can be quantified.

There are of course many distributions that already allow flexible hazard functions, at the cost of adding one or more additional parameters, and a brief literature review follows. A useful 2-parameter distribution due to Dhillon (1981) has survival function $\bar{F}(t) = \exp\{1 - \exp(\alpha t)^\beta)\}$, and is IFR for $\beta > 1$, or BT if $\beta < 1$. This distribution can never be purely DFR.

A mixture of two Weibull distributions with up to 5 parameters can generate 8 shapes of failure rate, but not a bathtub (Jiang and Murthy, 1998).

Chen (2000) described an intriguing 2-parameter distribution with survival function $\bar{F}(t) = \exp\{\lambda(1 - \exp(t^\beta))\}$ where $\lambda > 0, \beta > 0$, that can have an IFR, DFR or BT failure rate. This distribution would however have 3 parameters if the scale of t is made arbitrary; the author means to say that the distribution has 2 shape parameters. It generalizes Dhillon's distribution.

Yan and Leemis (2008) present the 2-parameter logistic-exponential distribution, where $\bar{F}(t) = \{1 + (\exp(\lambda t) - 1)^\kappa\}^{-1}$, where $\lambda > 0, \kappa > 0$. This interesting distribution can be UBT, BT, or exponential. In the tail, the hazard function tends to a constant, so the 2-parameter BT distribution cannot model wearout; however, on adding more parameters, a more flexible distribution is obtained.

The q-Weibull distribution with 3 parameters can give IFR, DFR and BT curves (Assis *et al*, 2013). Here the exponential function is replaced by $\exp_q(x) = (1 + (1 - q)x)^{(1-q)^{-1}}$ if $1 + (1 - q)x \geq 0$, else zero. The finite

3

range of $\exp_q(-x)$ for $q < 1$ is the reason why this model can be initially DFR and later IFR; with a finite range, if an event has not yet occurred near the end of the range, it *must* occur in the remaining interval, giving a large hazard. It is also probably the reason for numerical difficulties in fitting the model by maximum-likelihood described by Xu *et al* (2017). The generalized gamma or Stacy distribution (e.g. Cox *et al*, 2007) generalizes the Weibull distribution with an extra parameter, and can have IFR, DFR, BT or UBT failure rates. It includes the gamma, Weibull and lognormal distributions as special cases.

Recently, many more Weibull-like distributions have been produced. Méndez-González *et al* (2017) used the 4-parameter beta-Weibull distribution to model failures of electronic equipment. Bo *et al* (2016) studied the additive modified Weibull distribution, and Prataviera *et al* (2018) studied a 4-parameter extension to the Weibull distribution.

This account is of course incomplete; turning to reviews, Pham and Lai (2007) describe generalizations of the Weibull distribution, Nadarajah (2009) describes 17 distributions with bathtub-shaped hazards, and Almiki and Nadarajah (2014) give a comprehensive account of modifications to the Weibull distribution to achieve a more flexible hazard function. Regarding methodology, Lai (2012) gives a comprehensive account of ways to modify survival distributions.

The methodology for generating new distributions is described next, then some new distributions derived from the Weibull, gamma and lognormal distributions are given, and finally some well-used reliability datasets are fitted to exemplify the methodology. The article ends with some conclusions.

The general method for generating new distributions is described first, followed by a useful special case.

2. A general method for constructing new survival distributions

Let $u(t)$ be a monotonically increasing function of time t, and $v(t)$ a monotonically decreasing one, and write the pdf $f(t)$ as a product $f(t) = -u(t)v'(t)$, where $v(\infty) = 0$, $v'(t) \leq 0$, and $u(0) = 0$, $u'(t) \geq 0$, and where primes denote derivatives. Then by the method of parts,

$$\int_0^\infty f(t)\,\mathrm{d}t = -\int_0^\infty u(t)v'(t)\,\mathrm{d}t = -[uv]_0^\infty + \int_0^\infty v(t)u'(t)\,\mathrm{d}t.$$

4

The first term is zero, and $v(t)u'(t) > 0$, so

$$g(t) = v(t)u'(t) \qquad (1)$$

is a pdf. For example, one could take $u(t) = t^{1/2}, v'(t) = -\exp(-t)/t^{1/2}$, so that $f(t)$ is the exponential pdf, to generate a new distribution.

If the survival function of the original distribution is $\bar{F}(t)$, applying the method of parts to (1), the survival function corresponding to $g(t)$ is

$$\bar{G}(t) = \bar{F}(t) - u(t)v(t), \qquad (2)$$

showing that the probability mass has moved to the left. More precisely, the old distribution stochastically dominates the new one; this is first-order stochastic dominance.

If the original mean is $\mathrm{E}_f(T)$, the mean of the new distribution is

$$\mathrm{E}_g(T) = \int_0^\infty tv(t)u(t)' \, \mathrm{d}t = \mathrm{E}_f(T) - \int_0^\infty u(t)v(t) \, \mathrm{d}t, \qquad (3)$$

from the pdf $g(t)$ on integrating by parts or directly by integrating (2).

One can of course transform the integrand by going L $\rightarrow$ R from the parent distribution instead of R $\rightarrow$ L as shown above, to obtain $f(t)$ from $g(t)$ by integrating u' and differentiating v. In this case, the new R distribution stochastically dominates the parent.

The shifting of probability mass can be better understood by tagging the mass at t_0 using a Dirac delta-function. Then we apply IBP to $-u(x)v'(x)\delta(t-t_0)$, to obtain $-u'(x)v'(t_0)$ for $t \le t_0$, else zero. Hence a unit probability mass at t_0 has been smeared out onto the range $(0, t_0)$ with pdf $u'(t)/u(t_0)$, distribution function $u(t)/u(t_0)$. Going from the L to the R distribution, the probability mass is smeared out over (t_0, ∞) with survival function $v(t)/v(t_0)$.

On letting $R \rightarrow L$, because probability mass has been shifted left, and the lower limit of zero is fixed, applying IBP can give a distribution that is similar in the right tail to the old one, but has much more probability mass at low t. Hence it is easy to generate distributions that are initially DFR and are still IFR/DFR in the tail. From an IFR parent distribution, one can generate distributions that are IFR, BT or modified BT.

The choice of $u(t)$ is wide, and so this is a fertile way of creating new distributions. A function that is zero at $t = 0$ and never decreases is required. Obvious choices are $u = t^\lambda$ for $\lambda > 0$, or $u = \exp(\lambda t) - 1$. Another choice

would be to take $u(t)$ as a monotone function of any distribution function, not necessarily that of the parent distribution. With $u = F(t)^\lambda$ the integral can be done analytically. In this case, the new pdf is given by

$$g(t) = \frac{\lambda}{\lambda - 1}(1 - F(t)^{\lambda-1})f(t) \tag{4}$$

and the distribution function $G(t)$ is given by

$$G(t) = \frac{\lambda F(t) - F^\lambda(t)}{\lambda - 1}. \tag{5}$$

This is a negative mixture of the original pdf with weight $\lambda/(\lambda - 1)$ and the top λ-th order statistic with weight $-1/(\lambda - 1)$. When $\lambda = 1$,

$$G(t) = \{1 - \ln(F(t))\}F(t), \quad g(t) = \ln(1/F(t))f(t). \tag{6}$$

As $\lambda \to \infty$, $G(t) \to F(t)$ because $F^\lambda \to 0$. A tractable example of this case is given briefly in the appendix.

Going $L \to R$, set $v(t) = \bar{G}^\lambda(t)$, so $u' = g\bar{G}^{-\lambda}$ and $u = \frac{\bar{G}^{\lambda-1}-1}{\lambda-1}$ and $\bar{F} = \frac{\lambda\bar{G}-\bar{G}^\lambda}{\lambda-1}$. This is a negative mixture that is longer-tailed than was the L-distribution. When $\lambda = 1$, $\bar{F} = (1 - \ln(\bar{G}))\bar{G}$.

The case where u is a power of t is particularly tractable and so the method is now specialised to this case.

3. A useful special case

When u is a power of time, much further progress can be made. First, let the original pdf have increasing and decreasing factors $u, v'(t)$ given by $u_0(t), v_0'(t)$. Define $u(t) = t^\lambda u_0(t), v'(t) = v_0'(t)/t^\lambda$ to obtain a general distribution; the new distribution is conveniently named after the 'parent' distribution, e.g. the lambda or $\lambda-$Weibull distribution.

All lifetime distributions must have a scale factor α. Without loss of generality, we set $\alpha = 1$, and put the scale back in later if necessary; this simply means that $t \to \alpha t$, $\mathrm{d}t \to \alpha\,\mathrm{d}t$.

Thus $u = t^{\beta-1+\lambda}$, where $\beta - 1$ is the power of t in an original distribution. The parameter β is positive for Weibull and gamma distributions, and $\beta = 1$ for the lognormal distribution. Then

$$g(t) = (\beta + \lambda - 1)t^{\beta+\lambda-2}v(t). \tag{7}$$

We must have that $\beta + \lambda - 1 > 0$ so that the pdf is positive, hence $\lambda > 1 - \beta$.

3.1. Moments

The moments can now be written more simply than (3).

$$\mathrm{E}_g(T^n) = -(\beta + \lambda - 1) \int_0^\infty t^{\beta - 2 + \lambda + n} v(t) \, \mathrm{d}t,$$

and applying the method of parts,

$$\mathrm{E}_g(T^n) = \frac{\beta - 1 + \lambda}{\beta - 1 + \lambda + n} \int_0^\infty t^{\beta - 1 + \lambda + n} v(t)' \, \mathrm{d}t = \frac{\beta - 1 + \lambda}{\beta - 1 + \lambda + n} \mathrm{E}_f(T^n). \quad (8)$$

The mean residual life (MRL) at time t is given by

$$\mathrm{MRL}_g(t) = \frac{\int_t^\infty g(x)(x - t) \, \mathrm{d}x}{\bar{G}(t)} = \frac{\int_t^\infty \bar{G}(x) \, \mathrm{d}x}{\bar{G}(t)}.$$

From (2) after some manipulation, we have that

$$\mathrm{MRL}_g(t) = \frac{\beta + \lambda - 1}{\beta + \lambda} \left(\frac{\bar{F}(t)}{\bar{G}(t)} \right) \mathrm{MRL}_f(t) - t/(\beta + \lambda). \quad (9)$$

This give the MRL in terms of $\bar{G}$ and quantities pertaining to the parent distribution.

3.2. Status as a mixture distribution

To see that $g(t)$ is the pdf of a a mixture distribution, we can write

$$g(t) = -(\beta + \lambda - 1) t^{\beta + \lambda - 2} \int_t^\infty v_0'(x)/x^\lambda \, \mathrm{d}x,$$

and set $x = t/\omega$, where $\omega < 1$. Then

$$g(t) = -\int_0^1 (\beta + \lambda - 1)(t/\omega)^{\beta - 1} \{v_0'(t/\omega)/\omega\} \omega^{\beta + \lambda - 2} \, \mathrm{d}\omega. \quad (10)$$

This is a mixture of the original pdfs, i.e. $g(t) = \int_0^1 \{f(t/\omega)/\omega\} \, \mathrm{d}F(\omega)$, where $F(\omega) = \omega^{\beta + \lambda - 1}$, so that the mixing distribution is the power-function distribution.

Hence the integration by parts gives a mixture of the original pdfs, where the scale factor α is greater than the original. This is a model for early failures, in that the population of items is heterogeneous, with frail or flawed

7

components. This mixing can be extended to the case where there are n components in a batch with a random frailty, but that is not considered further here. Frailty models are discussed in e.g. Duchateau and Janssen (2008).

This representation allows the generation of random numbers from the pdf $g(t)$. Generate a random value Ω from a uniformly-distributed random number U as $\Omega = U^{1/(\beta+\lambda-1)}$, and a random number X from the parent distribution. Then a r.v. from $g(t)$ is ΩX.

As $\lambda \to \infty$, the mixing distribution tends to a delta function at $\Omega = 1$, and the new distribution reduces to the parent distribution. In the tail for $t \gg 1$, the items with $\omega < 1$ will have failed leaving only the $\omega \simeq 1$ items, and thus the hazard function and MRL should reduce to those of the original distribution in the tail. This will be true if $v'(t)$ goes to zero quickly enough so that $\bar{F}(t;\omega)/\bar{F}(t;1) \to 0$ as $t \to \infty$. It may well happen however that this extreme tail behaviour is irrelevant in practice, because nearly all components will have failed by then.

To see more formally that the new distribution reduces to the parent distribution as $\lambda \to \infty$, from (2) we have that

$$\bar{G}(t) - \bar{F}(t) = t^{\beta+\lambda-1} \int_t^\infty x^{-\lambda} v'(x)\, \mathrm{d}x.$$

Expanding v' in a Taylor series about t, $v'(x) = v'(t) + (x-t)v''(t) + \cdots$, and the integral is

$$\int_t^\infty x^{-\lambda} v'(x)\, \mathrm{d}x = -\frac{t^{1-\lambda} v'(t)}{\lambda - 1} + \frac{v''(t) t^{2-\lambda}}{(\lambda-1)(\lambda-2)} + \cdots$$

Thus

$$\bar{G}(t) - \bar{F}(t) = -t^\beta \{v'(t)/(\lambda-1) + v''(t)/(\lambda-1)(\lambda-2) + \cdots\},$$

from which $\bar{G} \to \bar{F}$ as $\lambda \to \infty$.

The 'descendant' distribution pdf can have the mode at zero or be unimodal if the parent distribution is unimodal. We do not have a proof that it cannot be bimodal, although none of the distributions examined has been.

3.3. Hazard function

The hazard for small t is now discussed. The initial hazard is $\propto t^{\beta+\lambda-2}$ for $\lambda < 1$ if $v'(0)$ is nonzero and finite, because the integral $v(t) = -\int_t^\infty v_0'(x)/x^\lambda\, \mathrm{d}x$

converges. However, for $\lambda > 1$ the integral tends to infinity as $t \to 0$. Expanding v_0 in a Taylor series, for small t, $v(t) \simeq (\lambda-1)^{-1}v_0'(0)t^{1-\lambda}$, and hence the initial hazard is $\propto t^{\beta-1}$ as for the parent distribution, so the descendant distribution is not BT if the parent is not. However, the hazard for small t is still greater than for the parent distribution, being $\simeq -\frac{\beta+\lambda-1}{\lambda-1}t^{\beta-1}v_0'(0)$.

For distributions such as the lognormal, $v'(t)$ tends to zero faster than any power of t. In this case, $v(0)$ is finite, and the initial hazard is $t^{\beta+\lambda-1}$ for all λ. When v is finite at zero time, the initial hazard function goes as $t^{\beta-2+\lambda}$. Thus if $\beta + \lambda < 2$ the distribution is DFR. If the parent distribution was IFR, the new distribution is BT.

4. Bathtub distributions and implications for reliability engineers

With a mixture distribution, equation (8) can yield a formula for the gain in lifetime on eliminating inferior items with mean lifetime reduced to a factor $\omega < 1$ of the lifetime of strong components, where ω is defined in (10).

Define $\xi = 1/(\beta+\lambda-1)$ as the potential for reliability growth. The mean lifetime is $E_g(T) = \frac{\beta-1+\lambda}{\beta+\lambda}E_f(T)$, so the proportional increase in expected lifetime on eliminating inferior components is

$$\mathrm{PI} = \frac{E_f(T) - E_g(T)}{E_g(T)} = \frac{1}{\beta+\lambda-1} = \xi,$$

and the proportional decrease in lifetime caused by inferior items is $1/(\beta+\lambda) = \xi/(\xi+1)$. The lambda-distribution can be fitted to failure-time data, λ and β estimated, and PI and its standard error computed. As λ increases and the parent distribution is regained, PI goes to zero.

A BT failure rate shows that there are indeed flawed components, but for small values of ξ the hazard function can still be IFR. Reliability gain is thus possible not just for BT curves but also for IFR curves, and conversely a curve can be DFR so that burn-in might be done, but if $\xi = 0$ no gain is possible from elimination weak units.

A mixture of flawed units with resulting high 'infant mortality' is thought to occur most often with electronic components, such as micro-electro-mechanical systems. Naturally, this is a problem in areas where technology is advancing rapidly. The reliability engineer must design an optimal burn-in program, perhaps with accelerated or stress testing, that can weed out the weakest components, or may decide that the production line or the product must be

9

re-engineered. Nowadays, degradation-based burn-in is often used because units take too long to fail completely (e.g. Yuan *et al*, 2016).

The first step in solving a problem is often to quantify it, and fitting e.g. the λ-Weibull distribution to failure data allows this to be done. The reliability gain that could be achieved on eliminating weak units can be estimated. The engineer could use the fitted λ-distribution to help design an efficient burn-in program, or might decide that the gain is great enough to justify improving the manufacturing process instead.

When the distribution is IFR, and ξ is small, the model fit may still indicate that there are weak units and reliability growth is possible. Here burn-in will not help, but knowing that there are weak units, the engineer may be able to investigate the cause and improve the manufacturing process. The examples section illustrates these ideas, and some caveats are given in the conclusions.

5. New distributions

The next sections give specific examples of the new distributions. Here the scale factor α has been put back in, to make the results immediately usable. The properties are described only briefly, because moments, MRL and random number generation have already been dealt with for the general case. Note that a λ-GG distribution can be derived, where GG denotes the generalized gamma distribution. This includes gamma, Weibull, and even lognormal, but it is simpler to derive these 3 lambda-distributions separately. The methodology could be applied to any survival distribution, but only the most widely-used are considered here.

In integrating $v'(t)$, the incomplete gamma function will sometimes be needed, defined as

$$\Gamma(a;t) = \int_t^\infty x^{a-1} \exp(-x)\, \mathrm{d}x,$$

where $a > 0$. If $a \leq 0$ $\Gamma(a;t)$ is still defined, but is not computable using software that computes the incomplete gamma function. The gamma function itself is $\Gamma(a) = \Gamma(a;0)$.
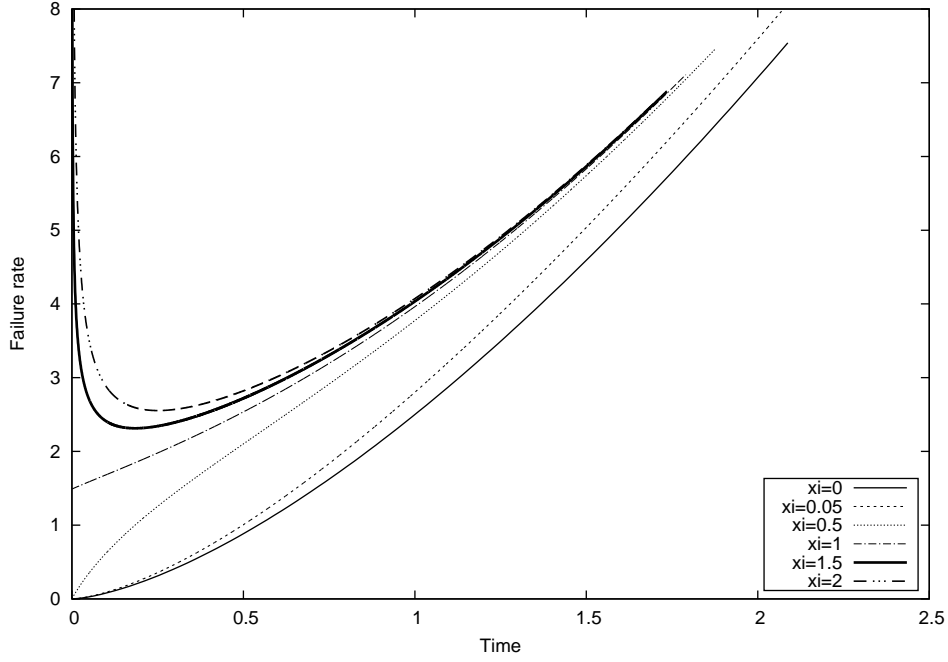
10

Figure 1: Hazard functions for a λ-Weibull distribution with $\beta = 2.5$ for varying values of ξ. The curves end when the survival function becomes less than 0.001.

5.1. The λ-Weibull distribution

Here $v'(t) = -\alpha \exp(-(\alpha t)^\beta)/(\alpha t)^\lambda$ and $u(t) = \beta(\alpha t)^{\beta+\lambda-1}$. Hence on setting $y = (\alpha x)^\beta$,

$$v(t) = \beta^{-1} \int_{(\alpha t)^\beta}^{\infty} y^{\frac{1-\lambda}{\beta}-1} \exp(-y) \, \mathrm{d}y = \beta^{-1} \Gamma(\frac{1-\lambda}{\beta}; (\alpha t)^\beta).$$

The pdf is

$$g(t) = \alpha(\beta + 1 - \lambda)(\alpha t)^{\beta+\lambda-2}\Gamma(\frac{1-\lambda}{\beta}; (\alpha t)^\beta).$$

From (2) the survival function is

$$\bar{G}(t) = \exp(-(\alpha t)^\beta) - (\alpha t)^{\beta+\lambda-1}\Gamma(\frac{1-\lambda}{\beta}; (\alpha t)^\beta).$$

Figure 1 shows the hazard function for $\beta = 2.5$ as ξ increases from zero, with the BT shape for $\xi > 1$.

11

5.2. The λ-exponential distribution

This is the special case of the λ-gamma and λ-Weibull distributions when $\beta = 1$. The parameter $\xi = \lambda^{-1}$. The pdf is $g(t) = \alpha\lambda(\alpha t)^{\lambda-1}\Gamma(1-\lambda;\alpha t)$ and the survival function is

$$\bar{G}(t) = \exp(-\alpha t) - (\alpha t)^\lambda\Gamma(1-\lambda;\alpha t).$$

Since the moments of the exponential distribution are $\mathrm{E}_f(T^n) = \alpha^{-n}n!$, from (8) we have that

$$\mathrm{E}_g(T^n) = \alpha^{-n}\lambda n!/(\lambda + n),$$

so that $\mathrm{E}_g(T) = \alpha^{-1}\lambda/(\lambda + 1)$,

$$\mathrm{var}(T) = \alpha^{-2}\frac{\lambda\{\lambda^2 + 2\lambda + 2\}}{(\lambda + 1)^2(\lambda + 2)}.$$

This simple form means that the parameters α, λ can be estimated from the sample mean and variance by the method of moments. Solving,

$$\hat{\lambda} = \frac{2/(A-1)}{1 + \sqrt{1 + 2/(A-1)}},$$

and

$$\hat{\alpha} = \frac{\hat{\lambda}}{(\hat{\lambda} + 1)\bar{t}},$$

where $A = s^2/\bar{t}^2$, $\bar{t}$ is the sample mean and s^2 the sample variance. If the coefficient of variation $A^{1/2} \leq 1$ then the exponential distribution would be fitted as λ could not be estimated.

Figure 2 shows the hazard function for several values of ξ. The initial hazard is a constant for $\xi < 1$, in fact it is $1/(1 - \xi)$, and the distribution is always DFR. When $\xi = 1$, then $\Gamma(0; \alpha t)$ is the exponential integral, and the initial hazard is infinite. In fact, the hazard $\simeq -\alpha\ln(\alpha t)$ for small t. For $\xi > 1$ the initial hazard is infinite.

Random numbers from this distribution are particularly simple and a random number $T = -\alpha^{-1}U^\xi\ln(V)$, where U, V are uniformly distributed random numbers.

This could be a useful 2-parameter distribution for modelling failure data where there may be some flawed components giving a DFR distribution, but no wearout so the failure rate is expected to be constant in the tail. Unlike
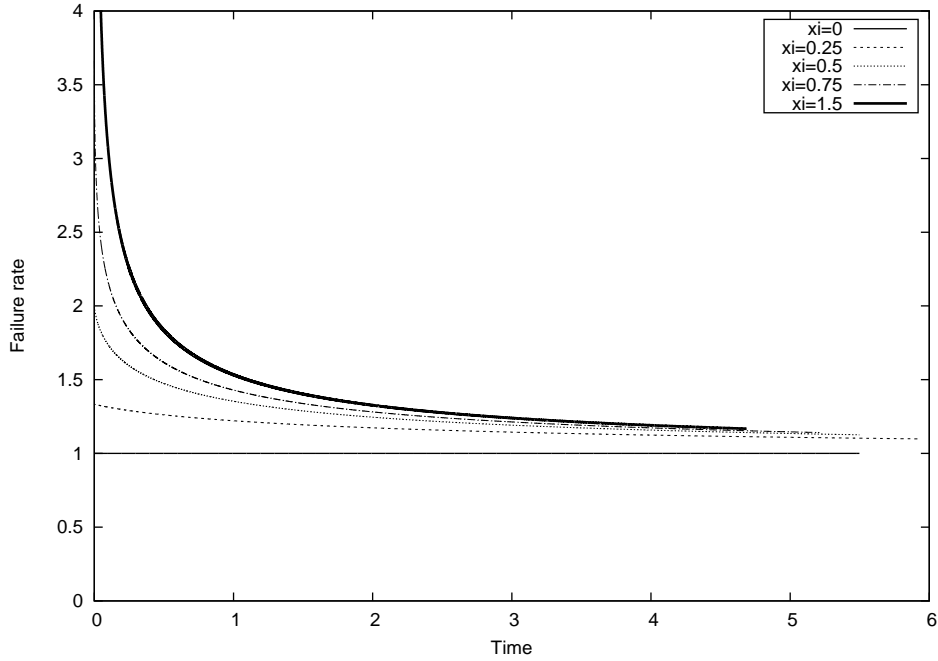
Figure 2: Hazard functions for a λ-exponential distribution for varying values of ξ. The curves end when the survival function becomes less than 0.001.

the gamma distribution, which also has this property, the initial failure rate can be finite for a DFR distribution. This distribution is thus comparable to the exponential-geometric distribution (e.g. Lai and Xie, 2006) which is also a mixture of exponentials. The λ-exponential distribution has more tractable moments but a less tractable hazard function.

An interesting special case occurs when $\lambda = 1/2$. Then $\Gamma(1/2; \alpha t) = 2\sqrt{\pi}\Phi(-\sqrt{2\alpha t})$, where Φ is the normal distribution function, and the pdf is

$$g(t) = \alpha\sqrt{\pi}(\alpha t)^{-1/2}\Phi(-\sqrt{2\alpha t}).$$

5.3. The λ-gamma distribution

Here

$$v(t) = \int_t^\infty \frac{\alpha \exp(-\alpha x)}{\Gamma(\beta)(\alpha x)^\lambda}\,\mathrm{d}x = \Gamma(1 - \lambda; \alpha t)/\Gamma(\beta).$$

Hence the pdf is

$$g(t) = \alpha(\beta + \lambda - 1)(\alpha t)^{\beta + \lambda - 2}\Gamma(1 - \lambda; \alpha t)/\Gamma(\beta),$$
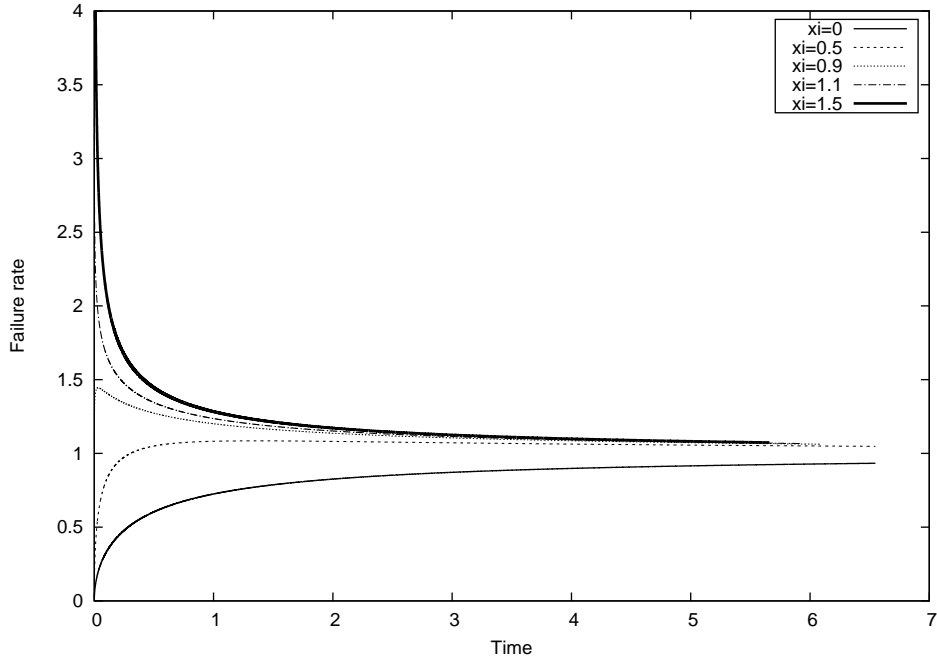
13

Figure 3: Hazard functions for a λ-gamma distribution with $\beta = 1.5$ for varying values of ξ. The curves end when the survival function becomes less than 0.001.

and

$$\bar{G}(t) = \Gamma(\beta; \alpha t)/\Gamma(\beta) - (\alpha t)^{\beta+\lambda-1}\Gamma(1 - \lambda; \alpha t)/\Gamma(\beta).$$

Figure 3 shows the hazard function for $\beta = 1.5$ as ξ increases from zero, with IFR shape for $\xi < 1$, DFR for $\xi > 1$ and a unimodal hazard for $\xi = 0.9$.

5.4. The λ-lognormal distribution

The lognormal pdf is

$$f(t) = \frac{1}{x}\frac{\exp(-(\ln(t) - \mu)^2/2\sigma^2)}{\sqrt{2\pi\sigma^2}}.$$

Setting $u(t) = t^\lambda$ and

$$v(t) = \int_t^\infty \frac{1}{x^{\lambda+1}}\frac{\exp(-(\ln(t) - \mu)^2/2\sigma^2)}{\sqrt{2\pi\sigma^2}}\,\mathrm{d}x,$$

and on changing variable to $y = \exp(x)$,

$$v(t) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\ln(t)}^\infty \exp(-(y - \mu)^2/2\sigma^2 - \lambda y)\,\mathrm{d}y.$$

14

Finally, on completing the square in the exponent,

$$v(t) = \exp(-\mu\lambda + \lambda^2\sigma^2/2)\Phi(-(\ln(t) - \mu + \lambda\sigma^2)/\sigma),$$

where Φ is the normal distribution function.

This yields the new pdf

$$g(t) = \lambda t^{\lambda-1}\exp(-\mu\lambda + \lambda^2\sigma^2/2)\Phi(-(\ln(t) - \mu + \lambda\sigma^2)/\sigma).$$

From (2) the survival function is

$$\bar{G}(t) = \Phi(-(\ln(t) - \mu)/\sigma) - t^\lambda\exp(-\mu\lambda + \lambda^2\sigma^2/2)\Phi(-(\ln(t) - \mu + \lambda\sigma^2)/\sigma),$$

or more elegantly

$$\bar{G}(t) = \Phi(-(\ln(t)-\mu)/\sigma) - \exp(\lambda(\ln(t)-\mu+\lambda\sigma^2/2))\Phi(-(\ln(t)-\mu+\lambda\sigma^2)/\sigma).$$

The distribution is initially either IFR like its parent if $\xi = 1/\lambda \leq 1$, so is UBT, and is DFR if $\xi > 1$. When $\lambda = 1$ it is UBT, with initial hazard function $\exp(-\mu + \sigma^2/2)$. Hence it can be a useful long-tailed DFR distribution. The failure rate is shown in figure 4.

5.5. Discrete distributions

Discrete distributions are sometimes used for failure data. Here summation by parts can yield new distributions in the same way that integration by parts did for continuous distributions. The mathematics is not very tractable, but when $\lambda = 1$ the class of Bissinger distributions (Johnson *et al* 2005, p. 509) is obtained. Here given a parent distribution with probability mass function (pmf) p_i, the probabilities q_i are

$$q_i = (1 - p_0)^{-1}\sum_{j=i+1}^{\infty} p_j/j.$$

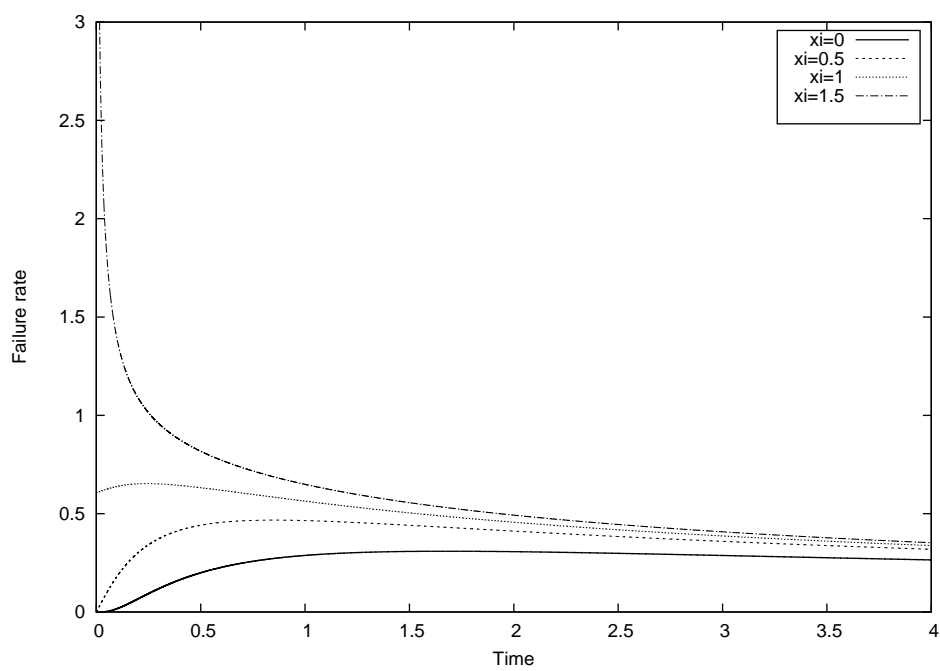The distributions clearly have their mode at zero, and have moments related to those of the parent distribution.

Figure 4: Hazard functions for a λ-lognormal distribution for varying values of ξ.

6. Model fitting and inference

Prototype fortran programs were written to check the mathematics shown here, and to fit the models to data. It is easiest but not essential to use a function minimiser that does not require derivatives to maximise the log-likelihood function. From the estimated Hessian matrix, large-sample standard errors on fitted model parameters and 95% confidence intervals were calculated. The computation of the incomplete gamma function was done using a standard function from the NAG (Numerical Algorithms Group) library. However, for non-positive arguments numerical quadrature was used. Currently, there is not much software available to evaluate these functions explicitly.

The gain from re-engineering out inferior items is $(\beta + \lambda - 1)^{-1}$. To estimate this, one can reparameterize the model so that the parameter λ is replaced by $\xi = (\beta + \lambda - 1)^{-1}$. Then simply $\lambda = 1 - \beta + \xi^{-1}$. The likelihood-maximisation routine will deliver an estimate of ξ and its standard error. When $\beta > 1$ the IFR distribution will be initially DFR if $\beta + \lambda - 2 < 0$ or $\xi > 1$.

Maximum-likelihood estimation can give biased estimates for small samples. To correct the estimate of ξ, the parametric bootstrap can be used to resample from the fitted distribution. If the average of the bootstrapped replicate estimates is $\hat{\bar{\xi}}$, the correction is $\hat{\xi} \to \hat{\xi} - (\hat{\bar{\xi}} - \hat{\xi})$ or $\hat{\xi} \to 2\hat{\xi} - \hat{\bar{\xi}}$. The nonparametric bootstrap can be used to generate resampled values, which are then corrected, and so to generate standard errors and a 95% confidence interval.

This can be used for a statistical test, using the equivalence between tests and confidence intervals, but a perhaps better method to test for the presence of a mixture of components is as follows: fit a parent-distribution model, e.g. Weibull, and record the increase $\Delta\ell$ in log-likelihood on floating the ξ parameter from zero. Then use the fitted Weibull model in the parametric bootstrap to generate samples conforming with H_0. The distribution of $\Delta\ell$ is found by floating ξ for these parametric-bootstrap resamples. The p-value of the test is the proportion of resamples for which $\Delta\ell$ exceeds the observed value.

7. Examples

The lambda-Weibull model was fitted to 3 datasets listed in Bebbington *et al* (2006). The first comprised 60 failure times (thousands of cycles) for electrical appliances. The data are originally from Lawless (2003). The second is time to first failure in thousands of hours for 36 500 MW generators, originally from Dhillon (1981). Finally, the third comprises 128 interfailure times (in thousands of hours) for a 180-ton rear dump truck, originally from Coetzee (1996). One should by the way probably not be analysing this dataset ignoring the age of the truck at failure.

The model was fitted by maximum-likelihood for the scale and shape parameters and ξ, with standard errors computed from the Hessian matrix, and naïve confidence intervals for ξ from assuming that $\hat{\xi}$ has a log-normal distribution. This means that $\ln \hat{\xi}$, which can take any value on the real line, is normally distributed. The parametric bootstrap was run 1000 times to generate an accurate bias correction for $\hat{\xi}$, and the 95% confidence interval found by generating the nonparametric (regular) bootstrap 1000 times. The bias correction was applied to the confidence interval limits.

Table 1 shows the results. All three datasets show that the bootstrap confidence interval is considerably smaller than the naïve interval, so using the bootstrap to improve inference is worthwhile.

The electrical appliance dataset has an estimated ξ value that suggests a large improvement in expected lifetime if the weaker components could be removed. The lower confidence limit exceeds unity, so a test of the BT property would reject H_0 at 2.5%. The distribution is clearly IFR in the tail, so is clearly BT, although our pragmatic approach focuses more on the gain that could be achieved by 'getting out of the bathtub'.

The generator dataset again shows a high value for $\hat{\xi}$, in fact higher than the previous example, but here the smaller sample size of 36 means that one cannot reject H_0 that the distribution is not BT.

Finally, the dump truck dataset is DFR in the tail, although the upper 95% confidence limit (not shown) for β is 1.116, so the distribution could be IFR in the tail. The distribution must be initially DFR. It may or may not be BT, depending on whether it is IFR in the tail or not. However, although it cannot be claimed as a BT distribution, from the bias-corrected estimate $.788 - .264 = .524$, the gain from removing weaker components ξ looks to be at least 50% increase in expected lifetime.

Comparing with the generalized gamma model, the λ-Weibull fitted slightly

better for the generator data, and slightly worse for the electrical components and the dump truck data, as shown by the log-likelihood function; the improvements in log-likelihood were 0.09. -0.8 and -0.505 respectively. The important message is that the fits are comparable to other models, but here there is also information to help in the quest for reliability growth.

To see what happens when the model is fitted to an IFR distribution, it was fitted to the dataset of failure times of 101 aluminium coupons given in Lie and Xie (2006), and originally from Birnbaum and Saunders (1958). The Weibull shape parameter was $\hat{\beta} = 4.047$, showing a strongly IFR distribution of failure times. On fitting the λ-Weibull model, $\hat{\xi}$ fitted as 0.0018 and the log-likelihood was only increased by a tiny amount (the same to 5 decimal places) on adding the third parameter. Thus, this model did not find a mixture distribution when it was absent in this example.

In general, large-sample inference for ξ must respect the fact that this is a parameter that can only be non-negative, and is at the edge of its range under the null hypothesis that it is zero. How to do this is discussed in e.g. Feng and McCulloch (1992). A bootstrap procedure was given in the previous section.

Figure 5 shows the fitted hazard functions for the 3 datasets. The scale factors have all been set to 1 to put the curves onto the same timescale. This plot confirms what was deduced from the fitted parameter values.

To summarise, the procedure that reliability engineers might follow is:

1. Fit a model, e.g. the λ-Weibull model to failure data for new units. These models are accelerated-time models, so stress-testing results can easily be transformed back to the real time-scale;
2. Do some checks to be satisfied that the model fits the data, and compute errors or confidence intervals on the reliability-growth parameter ξ as described above;
3. Study the cost benefit of the reliability gain that the model suggests is possible, and so decide whether to burn-in the existing product, or to re-engineer the production process;
4. If using burn-in, the mean residual life formula (9) can be helpful in computing the optimum period.

8. Conclusions

A large class of new lifetime distributions has been introduced, with the λ distributions a useful special case. A crucial issue for manufacturers is how
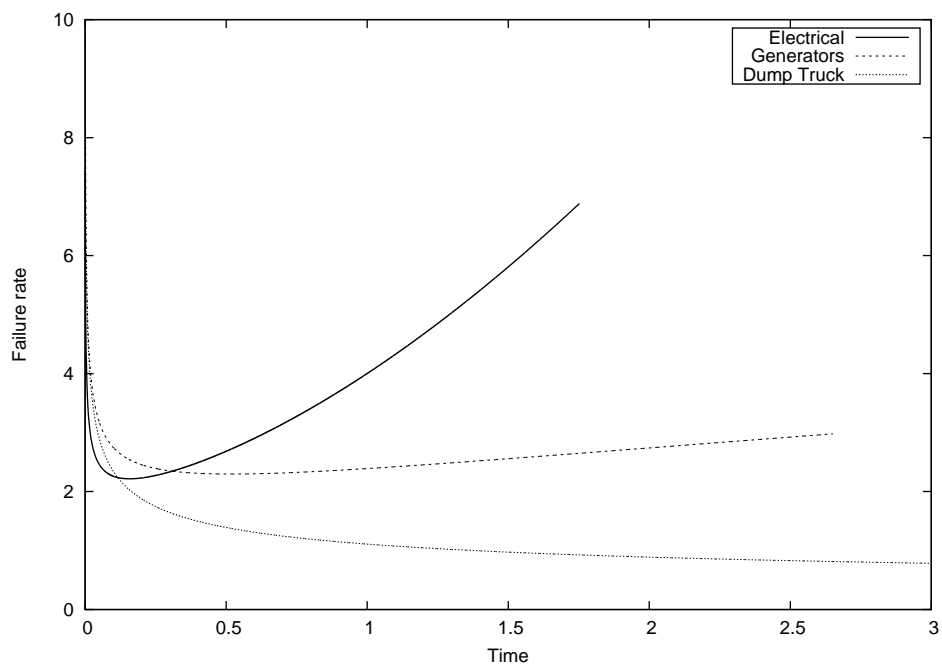
Figure 5: Estimated hazard functions for the 3 datasets studied. The curves end when the survival function becomes less than 0.001.

Dataset	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\xi}$	Bias	naïve CI	Boot CI
Electrical	2.488 (.741)	-.76	1.374 (.251)	-.035 (.008)	(.961, 1.965)	(1.016, 1.744)
Generators	1.519 (1.073)	.179	1.433 (.723)	-.085 (.014)	(.535, 3.834)	(.880, 2.136)
Dump truck	0.794 (.138)	1.474	.788 (.526)	.264 (.012)	(.213, 2.914)	(.548, 1.261)

Table 1: The λ-Weibull model fitted to the 3 datasets from Bebbington *et al* (2006) mentioned in the text. The bracketed numbers are standard errors. CI denotes 95% confidence interval, and Boot CI is the bias-corrected 95% bootstrap confidence interval for ξ.

serious the problem of early failure due to flawed components is, so-called 'infant mortality'. The new distributions allow this to be quantified and estimated as the proportionate or percentage increase in mean lifetime that could be achieved (reliability growth).

Estimation of the increase ξ is perhaps more useful than testing for the BT property. Indeed, a small admixture of flawed components will not produce a BT distribution, but as shown in figure 1 the failure rate can still be initially higher than the Weibull distribution would give. The gain ξ can be computed even in these situations.

The estimation of the gain from eliminating flawed components must of course be hedged about with caveats. The statistical error on the parameter ξ is fairly well quantifiable as a 95% confidence interval, where the bootstrap can be used to improve on the naïve large-sample estimate. However, there are other non-statistical causes of uncertainty. An admixture of frail components is not the only possible cause of an initially DFR distribution of failure time although the most likely, as mentioned in the introduction.

We have also to assume that the parent distribution, e.g. Weibull, would be a good fit in the absence of frail components. All this comes under the heading of 'model risk'. Still, $\hat{\xi}$ could at least be taken as an upper bound on what could be achieved by tighter quality control in manufacturing or engineering out defects.

Further work could include the derivation of more survival distributions using the general method, as illustrated in the appendix, and exploration of the discrete case in more detail. A vast number of new distributions can be generated, of which some should be useful.

21

[1] Almiki, S. J., Nadarajah, S., Modifications of the Weibull distribution: A review, Reliability Engineering and System Safety 2014, **124**, 32-55.

[2] Assis, E.M.; Borges, E.P., de Melo, S.A.B.V. Generalized q-Weibull model and the bathtub curve. International Journal of Quality and Reliability Management. 2013, 30, 720736.

[3] Bebbington, M., Lai, C-D, Zitikis, R., Useful periods for lifetime distributions with bathtub-shaped hazard rate function, IEEE Transactions on Reliability 2006, **55**, 245-251.

[4] Birnbaum, Z. W. and Saunders, S. C., A statistical model for life-length of materials, Journal of the American Statistical Association 1958, **53**, 151-160.

[5] Bo H., Cui, W. and Du, X. (2016) An additive modified Weibull distribution, Reliability Engineering and System safety, **145**, 28-37.

[6] Chen, Z., A new two-parameter lifetime distribution with bathtub shape or increasing failure function, Statistics and Probability Letters 2000, **49**, 155-161.

[7] Coetzee, J.L., Reliability degradation and the equipment replacement problem, in Proceedings of the International Conference of maintenance Societies (ICOMS-96), Melbourne 1996, paper 21.

[8] Cox, C., Chu, H., Schneider, M. F. and Muñoz, A., Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution, Statistics in Medicine 2007, **26**, 4352-4374.

[9] Dhillon, B.S., Life Distributions, IEEE Transactions on reliability, 1981, R-30, 457-460.

[10] Duchateau, L. and Janssen, P., The Frailty Model, 2008, Springer, Berlin.

[11] Feng, Z. and McCulloch, C. E. . Statistical inference using maximum likelihood estimation and the generalized likelihood ratio when the true parameter is on the boundary of the parameter space, Statistics and Probability Letters 1992 **13**, 325-332.

[12] Gupta, R.D. and Kundu, D., Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions, Biometrical Journal, 2001 **43**, 117-130.

[13] Jiang, R., Murthy, D. N. P. (1998) Mixture of Weibull distributions - parametric characterization of failure rate function. Applied Stochastic Models and Data Analysis 1998, **14**, 4765.

[14] Johnson, N.L., Kemp, A. W., Kotz, S., Univariate Discrete Distributions, 3rd ed., 2005, Wiley, New York.

[15] Johnson, N.L., Kotz, S., Balakrishnan, N., Continuous Univariate Distributions, 2nd ed., 1995, Wiley, New York.

[16] Klutke, G., Kiessler, P. C., Wortman, M. A., A critical look at the bathtub curve, IEEE Transactions on Reliability 2003 **52**, 125-129.

[17] Lai, C-D., Xie, M., Stochastic ageing and Dependence for Reliability, 2006, Springer, Berlin.

[18] Lai, C. D. Constructions and applications of lifetime distributions, Applied Stochastic Models in Business and Industry 2012, **29**, 127-140.

[19] Lawless, J.F., Statistical Models and Methods for Lifetime data, 2nd. ed., 2003, Wiley, New York.

[20] Méndez-González, L., Rodriguez-Picón, L. A., Valles-Rosales D. J., Romero-López R. and Quezeda-Carreón, A. E., Reliability analysis for electronic devices using beta-Weibull distribution, Quality and Reliability Engineering International 2017 **33**, 2521-2530.

[21] Nadarajah, S., Bathtub-shaped failure rate functions, Quality and Quantity 2009, **43**, 855-863.

[22] Pham, H. and Lai, C-D., On Recent Generalizations of theWeibull Distribution, IEEE Transactions on Reliability 2007 **56**, 454-458.

[23] Prataviera, F., Ortega, E. M. M., Cordeiro, G. M. and Pescim, R. R., A new generalized odd log-logistic flexible Weibull regression model with applications in repairable systems, Reliability Engineering and Systems safety, 2018, **176**, 13-26.

[24] Proschan, F., Theoretical explanation of observed decreasing failure rate, Technometrics 1963, **5**, 375-383.

[25] Wang, K. S., Hsu, F. S. and Liu, P. P., Modeling the bathtub shape hazard rate function in terms of reliability, Reliability Engineering and Systems Safety 2002, **75**, 397-406.

[26] Xu, M., Droguett, E. L., Lins, I. D., and das Chagas Moura, M. On the q-Weibull distribution for reliability applications: An adaptive hybrid artificial bee colony algorithm for parameter estimation. Reliability Engineering & System Safety 2017, **158**, 93 105.

[27] Yan, Y., Leemis, L.M., The logistic-exponential survival distribution, Naval Research Logistics 2008, **65**, 252-264.

[28] Yuan, T., Bae, S. J., Zhu, X., A Bayesian approach to degradation-based burn-in optimization for display products exhibiting two-phase degradation patterns, Reliability Engineering and System safety 2016, **155**, 55-63.

Appendix: applying the F^λ method to the exponentiated exponential and Weibull distributions

To illustrate the general method of generating new survival distributions, it is applied to the exponentiated-exponential distribution of Gupta and Kundu (2001), which has distribution function $F_\kappa(t) = (1 - \exp(-\alpha t))^\kappa$ for $\alpha > 0, \kappa > 0$, and pdf $f_\kappa(t) = \alpha\kappa \exp(-\alpha t)(1 - \exp(-\alpha t))^{\kappa-1}$. The pdf $f_\kappa(t)$ could be transformed using $u(t) = F_\kappa^\lambda(t)$. In this case from (5) the distribution function is

$$G(t) = \frac{\lambda(1 - \exp(-\alpha t))^\kappa - (1 - \exp(-\alpha t))^{\lambda\kappa}}{\lambda - 1},$$

and from (4) the pdf $g(t)$ follows. This can be written as a negative mixture of exponentiated exponential distributions

$$g(t) = \frac{\lambda}{\lambda - 1}f_\kappa(t) - \frac{1}{\lambda - 1}f_{\kappa\lambda}(t),$$

from which most of its properties such as moments can be derived from those of the parent distribution. For all values of λ the pdf is a negative weight mixture, and if $\lambda = 1$, L'Hospital's rule or (4) gives the distribution with pdf

$$g(t) = -\alpha\kappa^2 \exp(-\alpha t)(1 - \exp(-\alpha t))^{\kappa-1} \ln(1 - \exp(-\alpha t)),$$

and distribution function

$$G(t) = (1 - \exp(-\alpha t))^\kappa + \kappa(1 - \exp(-\alpha t))^\kappa \ln(1 - \exp(-\alpha t)).$$

As λ increases, $g(t)$ resembles $f(t)$ more closely, and as $\lambda \to \infty$, $g(t) \to f(t)$.

The new distributions have hazard 2α in the tail, rather than α as do the parent distributions. Whereas the exponentiated exponential distribution resembles the gamma distribution in having a hazard function that increases or decreases to a constant, the new distributions have a bathtub shape if $\kappa > 1$ and $\lambda\kappa < 1$, i.e. if the original distribution is IFR and λ is small enough.

When the parent distribution is Weibull, with scale $\alpha > 0$ and shape parameter $\beta > 0$, from (5) we have that

$$G(t) = \frac{\lambda(1 - \exp(-(\alpha t)^\beta)) - (1 - \exp(-(\alpha t)^\beta))^\lambda}{\lambda - 1},$$

$$g(t) = \frac{\alpha\beta\lambda\{1 - (1 - \exp(-(\alpha t)^\beta))^{\lambda-1}\}(\alpha t)^{\beta-1}\exp(-(\alpha t)^\beta)}{\lambda - 1}.$$

The moments of this distribution cannot be easily expressed. On expanding out the exponentials for $\alpha t \ll 1$, it can be seen that the distribution is initially DFR if $\beta\lambda < 1$, hence it is BT if $\beta\lambda < 1$, $\beta > 1$. The hazard function in the tail is $2\alpha\beta(\alpha t)^{\beta-1}$, twice that of the parent Weibull distribution.